CloudNinja.pro

PROFESSIONAL · SAP-C03

AWS Certified Solutions Architect – Professional

SAP-C03 Cheat Sheet

CloudNinja · cloudninja.pro · Free — no signup required

Every in-scope service, architecture pattern and decision rule for SAP-C03, organised by the five exam domains: cloud architecture and AI, security and governance, cost, resilience and migration, and operational excellence. Assumes 2+ years of hands-on AWS architecture.

EXAM SNAPSHOT

- **75 questions** — 65 scored + 10 unscored
- **180 minutes** total time
- **720 / 1000** to pass
- **$300 USD** exam fee

- **Architecture & AI** — 25%
- **Security & Governance** — 25%
- **Cost-Optimized Design** — 15%
- **Resilience & Migration** — 15%
- **Ops Excellence & Automation** — 20%

FREE SAP-C03 practice exam →

SAP-C03 replaces SAP-C02 on 17 November 2026 (16 November 2026 is the last day to sit SAP-C02). Five domains instead of four; passing score 720 instead of 750. New: Bedrock and AgentCore, durable workflows, VPC Lattice, Cloud WAN, IPv6, hybrid compute, Clean Rooms, Verified Access and Permissions, Security Lake, post-quantum crypto, Cost Categories, Data Lifecycle Manager, Resilience Hub, FIS, AWS Transform. Removed services are listed on the last page. DEPRECATED = closed to new customers, RETIRED = gone: know them, but do not pick them for a new design.

1 · Architecture Design & Implementation (25%)

1.1 Event-Driven Messaging

Service	Ordering, delivery & when to pick it
Amazon SQS	Standard: nearly unlimited throughput, at-least-once, best-effort order. FIFO: ordered per *message group ID*, exactly-once processing with deduplication ID (5-minute window). Visibility timeout, DLQ with redrive, long polling, delay queues (max 15 min).
Amazon SNS	Pub/sub fan-out to SQS, Lambda, HTTPS, email, SMS, Firehose. Message filtering policies. FIFO topics pair with FIFO queues for ordered fan-out.
Amazon EventBridge	Event bus with content-based rules, schema registry, archive and replay, Pipes (source → filter → enrich → target), Scheduler. **Cross-account / cross-Region**: a rule targets another account's or Region's event bus; the receiving bus needs a resource policy.
Amazon Kinesis Data Streams	Ordered per *partition key* (shard), replayable (retention 24 h – 365 days), multiple consumers, enhanced fan-out. On-demand or provisioned mode.
Amazon Managed Streaming for Apache Kafka (MSK)	Managed Kafka (provisioned or serverless). Pick when the scenario already uses Kafka APIs or needs Kafka Connect / exactly-once Kafka transactions.
Amazon MQ	Managed ActiveMQ or RabbitMQ. Lift existing JMS/AMQP/STOMP/MQTT/OpenWire apps without rewriting messaging code. ActiveMQ active/standby spans two AZs.

- **Idempotency** — at-least-once delivery means consumers must tolerate duplicates: idempotency key in DynamoDB with a conditional write, or Powertools for Lambda idempotency utility.
- **Large payloads** — claim-check pattern: payload in S3, pointer in the message (SQS max message 256 KB).
- **Real-time push** — API Gateway **WebSocket APIs** keep connections; backend posts to `@connections`. Pub/sub with SNS or EventBridge behind it.

Exam Tip: 'Strict order per customer' = SQS FIFO with message group ID = customer ID, or Kinesis with partition key = customer ID. 'Replay last week's events' = Kinesis retention or EventBridge archive. 'Legacy JMS app, no code changes' = Amazon MQ (the SQS JMS library supports queues only, not topics).

1.2 Durable, Multi-Step Workflows

Option	Limits	Pick when
Step Functions Standard	Up to 1 year, exactly-once, billed per state transition	Long-running orchestration across AWS services, human approval with `.waitForTaskToken`, visual audit history, 220+ service integrations.
Step Functions Express	Up to 5 min, at-least-once (async), billed by executions, duration and memory	High-volume, short event processing. No `.waitForTaskToken`.
Lambda durable functions	Durable execution up to 1 year; each invocation still ≤ 15 min	Workflow written as ordinary code in Lambda. Steps are checkpointed and skipped on replay; waits and callbacks suspend without compute charges. SDK for Node.js/TypeScript, Python, Java, Go, .NET.

Exam Tip: 'Wait days for a human, pay nothing while waiting, keep the logic in code' = Lambda durable functions (callback). 'Visual workflow across many AWS services with an approval step' = Step Functions Standard + task token. 'Millions of 2-second events' = Express.

1.3 Compute, Containers & Ingress

Service	Key facts
Amazon EC2 + Auto Scaling	Families M/T (general), C (compute), R/X (memory), I/D (storage), P/G/Inf/Trn (accelerated); Graviton (g suffix) for price-performance. Target tracking, step, scheduled, predictive scaling; warm pools; lifecycle hooks.
AWS Application Auto Scaling	Scaling for non-EC2 targets: ECS services, DynamoDB, Aurora replicas, Lambda provisioned concurrency, SageMaker AI endpoints, Spot Fleet, MSK storage.
AWS Lambda	Max 15 min, up to 10,240 MB, reserved concurrency (cap) vs provisioned concurrency (pre-warmed), SnapStart, event source mappings for SQS/Kinesis/MSK/DynamoDB Streams.
AWS Fargate	Serverless compute for ECS and EKS. Fargate Spot for interruptible tasks (ECS).
Amazon ECS	AWS-native orchestration. Service auto scaling, deployment circuit breaker with rollback, blue/green. **ECS Service Connect** = service discovery + traffic metrics between ECS services.
Amazon EKS	Managed Kubernetes. Auto Mode, managed node groups, Karpenter / Cluster Autoscaler, Horizontal Pod Autoscaler, **AWS Load Balancer Controller** (Ingress → ALB, Service → NLB), Pod Identity.
Amazon ECR	Image scanning (basic or enhanced with Amazon Inspector), lifecycle policies to expire old images, cross-Region/account replication, pull-through cache, image signing.
AWS Elastic Beanstalk	PaaS for web apps on EC2. Deployment policies: all-at-once, rolling, rolling with batch, immutable, traffic splitting, blue/green (swap URLs).
Amazon VPC Lattice	Application networking across VPCs and accounts without peering or TGW; works with overlapping CIDRs. Service networks shared with AWS RAM, IAM **auth policies**, access logs. Targets: EC2, IPs, Lambda, ECS, EKS, ALB.

Ingress / edge need (Elastic Load Balancing, API Gateway, CloudFront)	Pick
HTTP routing by path/host/header, WAF, OIDC auth	Application Load Balancer
TCP/UDP, static IP per AZ, extreme throughput, PrivateLink front end	Network Load Balancer
Inline third-party firewall appliances	Gateway Load Balancer (GENEVE)
Managed API: throttling, usage plans, API keys, authorizers, caching	API Gateway REST API (HTTP API = cheaper, fewer features)
Header rewrites / redirects at the edge, sub-millisecond	CloudFront Functions (viewer events only)
Edge logic with network calls or origin events	Lambda@Edge

1.4 Multi-Tenant SaaS Isolation

Model	How it is enforced
Silo	Dedicated resources (account, cluster, table, DB) per tenant; tenant-specific KMS key. Highest isolation and cost; for regulated/premium tenants.
Pool	Shared resources keyed by tenant ID. Enforce outside app code: IAM `dynamodb:LeadingKeys` = `${aws:PrincipalTag/TenantId}` via session tags, PostgreSQL **row-level security**, S3 prefix per tenant with ABAC.
Bridge	Mix: shared web tier, siloed data tier (or the reverse).

- **Tenant context** — Cognito custom attribute or group in the JWT; the API layer passes it as a session tag when assuming a tenant-scoped role.
- **Noisy neighbours** — API Gateway usage plans per tenant tier, Kubernetes resource quotas, separate queues per tier.
- **Per-tenant cost** — tenant tags, split cost allocation data for EKS/ECS, Cost Categories.

1.5 Generative AI & ML Integration

Capability	What to know
Amazon Bedrock	Serverless access to foundation models through one API (Converse / InvokeModel). On-demand, batch (about 50% cheaper, best-effort completion), Provisioned Throughput. Prompt caching cuts cost and latency for repeated context.
Cross-Region inference	Inference profiles route requests across Regions. *Geographic* profile stays inside a geography (US, EU, APAC) for residency; *global* profile routes worldwide at about 10% lower price. No extra routing charge; SCPs must allow the destination Regions.
Bedrock Knowledge Bases	Managed RAG: ingest from S3 and other sources, chunk, embed, store in a vector store (for example OpenSearch Serverless or Aurora PostgreSQL with pgvector). Metadata filtering for per-user or per-tenant retrieval.
Bedrock AgentCore	Runtime (session-isolated serverless agent hosting), Gateway (turns APIs and Lambda into MCP tools), Identity (inbound auth + token vault for outbound OAuth), Memory, Code Interpreter, Browser, Observability, Policy. Works with any framework or model.
Multi-account GenAI	Central AI platform account, model access governed with SCPs, application inference profiles for per-team cost tracking, invocation logging to a central account.
SageMaker AI endpoints	Real-time (steady low latency), serverless (intermittent traffic, cold starts), asynchronous (large payloads, long processing, scale to zero), batch transform (offline). Inference components to pack models on one endpoint.

Amazon Bedrock Agents (Agents Classic) DEPRECATED is in maintenance mode and restricted for new customers. New agent designs use Bedrock AgentCore.

Exam Tip: 'Agent must call 14 internal REST APIs as tools' = AgentCore Gateway. 'Act on the user's behalf in a SaaS app' = AgentCore Identity (OAuth token vault). 'Grounded answers from our documents' = Knowledge Base. 'Overnight summarisation, cheapest' = Bedrock batch inference. 'EU data must stay in the EU' = geographic (EU) inference profile, not global.

1.6 Hybrid Cloud Patterns

Requirement	Answer
Encrypted link up in hours	AWS Site-to-Site VPN (two IPsec tunnels per connection; aggregate bandwidth with ECMP on Transit Gateway).
Consistent latency, high bandwidth	AWS Direct Connect (dedicated or hosted). Add MACsec (dedicated) or IPsec VPN over DX for encryption.
Maximum resiliency	Two DX locations, two connections each; VPN as backup over the internet.
Many VPCs + on-premises + Regions	Transit Gateway with DX Gateway, or Cloud WAN for a policy-driven global network.
Hybrid DNS	Route 53 Resolver inbound endpoints (on-prem → AWS) and outbound endpoints + forwarding rules (AWS → on-prem), rules shared with AWS RAM. Private hosted zones for split-horizon.

Hybrid compute	Control plane / hardware	Pick when
Amazon EKS Hybrid Nodes	AWS-managed EKS control plane in the Region; your own servers as nodes	Reuse on-prem hardware; reliable link to the Region.

Hybrid compute	Control plane / hardware	Pick when
Amazon EKS Anywhere	Customer-managed control plane and nodes on your infrastructure	Air-gapped / disconnected sites.
Amazon ECS Anywhere	ECS control plane in the Region; external instances registered with SSM Agent	Run ECS tasks on existing on-prem servers or VMs.
AWS Outposts	AWS-owned and AWS-managed racks or servers at your site, extension of a Region AZ	Same AWS APIs on premises (EC2, EBS, ECS/EKS, RDS on Outposts), local processing, needs a service link to the Region.
AWS Local Zones	AWS infrastructure in a metro area, extension of a parent Region	Low latency to end users in a city without owning a facility; some data residency needs.

Data sync / legacy	Use
AWS DataSync	Online transfer and scheduled sync (NFS, SMB, HDFS, object) to S3, EFS, FSx. Bandwidth throttling, verification, filters.
AWS Storage Gateway	S3 File Gateway (NFS/SMB onto S3), FSx File Gateway, Volume Gateway (iSCSI), Tape Gateway (virtual tapes). Low-latency local cache.
AWS Transfer Family	Managed SFTP, FTPS, FTP and AS2 endpoints backed by S3 or EFS for partner file exchange.
API Gateway + Lambda	Modernise legacy SOAP/XML behind a REST/JSON facade (Lambda transforms; REST API mapping templates for simple cases). Strangler fig in front of a monolith.

1.7 Network Architecture

Connectivity option	Use
VPC peering	1:1, non-transitive, no overlapping CIDRs. Few VPCs, lowest cost (no hourly charge).
AWS Transit Gateway	Regional hub; route tables per attachment for segmentation; inter-Region peering; per-attachment hourly + per-GB data processing.
AWS Cloud WAN	Global network from one *core network policy*: segments are isolated routing domains, segment sharing, `isolate-attachments`, `require-attachment-acceptance`, tag-based attachment policies, service insertion for inspection.
AWS PrivateLink	Expose one service (behind an NLB) to many consumer VPCs and accounts; overlapping CIDRs are fine; consumer-initiated only.
VPC endpoints	Gateway endpoints (S3, DynamoDB, free) and interface endpoints (hourly + per GB). Endpoint policies form part of a data perimeter.
Amazon VPC Lattice	Service-to-service (L7) connectivity with auth policies across accounts; no routing tables to manage.

- **Amazon VPC IP Address Manager (IPAM)** — pools per Region/OU with allocation rules shared through RAM; prevents overlapping CIDRs; tracks usage.
- **IPv6 / dual-stack** — VPCs and subnets get an Amazon-provided /56 and /64; private IPv6 egress through an **egress-only internet gateway**; IPv6-only subnets reach IPv4 services through **DNS64 + NAT64** on the NAT gateway. Reduces public IPv4 charges.
- **Traffic inspection** — AWS Network Firewall (stateful, Suricata rules) in a central inspection VPC, Gateway Load Balancer for third-party appliances, AWS WAF for L7 on CloudFront/ALB/API Gateway, security groups (stateful, allow-only) and NACLs (stateless, allow + deny), VPC Flow Logs.
- **Global routing** — Route 53 policies: simple, weighted, latency, failover, geolocation, geoproximity, multivalue, IP-based. AWS Global Accelerator: two static anycast IPs, TCP/UDP, fast Regional failover.

Exam Tip: 'Prod and dev must never talk, both need shared services, across 6 Regions, least effort' = Cloud WAN segments. 'Overlapping CIDRs, service-to-service, per-request IAM auth' = VPC Lattice. 'Expose one SaaS endpoint to hundreds of customer VPCs' = PrivateLink. 'IPv6-only pods must call an IPv4 API' = DNS64 + NAT64.

1.8 Data, Storage & Databases

Lakehouse & analytics

Service	Role
Amazon S3 + Apache Iceberg	Open table format on S3: ACID transactions, row-level upserts and deletes (MERGE INTO), time travel, schema evolution. Queried by Athena, EMR, Glue, Redshift.
AWS Glue	Serverless Spark ETL, Data Catalog (central metadata), crawlers, job bookmarks for incremental loads, **Glue Data Quality** rulesets.
AWS Lake Formation	Fine-grained permissions on catalog tables: LF-Tags (ABAC), data filters for row-, column- and cell-level security, cross-account grants without copying data.
Amazon Athena	Serverless SQL on S3; pay per TB scanned (Parquet + partitions cut cost); workgroups for cost limits.
Amazon Redshift	Data warehouse (provisioned or serverless). Spectrum queries S3; data sharing across accounts; zero-ETL from Aurora, RDS and DynamoDB.
Amazon EMR	Spark/Hive/Presto/Flink on EC2, EKS or serverless. Spot for task nodes.
Kinesis Data Streams / Amazon Data Firehose	Streams for real-time custom consumers; Firehose for managed delivery to S3, Redshift, OpenSearch, Splunk, HTTP (formerly Kinesis Data Firehose).
Amazon MSK / Amazon AppFlow	Kafka streaming; AppFlow pulls SaaS data (Salesforce, SAP, ServiceNow) into S3/Redshift on a schedule or event.
Amazon OpenSearch Service	Search, log analytics, vector search; OpenSearch Serverless collections.
Amazon Quick Sight	BI dashboards, part of Amazon Quick (renamed from Amazon QuickSight).
AWS Clean Rooms	Analyse combined data with partners without sharing or copying raw data. Analysis rules (aggregation, list, custom), minimum aggregation thresholds, join requirements, only the designated member receives results; differential privacy option.

Database selection

Need	Pick
Relational, managed, HA	RDS Multi-AZ (one standby) or Multi-AZ DB cluster (two readable standbys); read replicas (cross-Region allowed). RDS Proxy for connection pooling.
Cloud-scale MySQL/PostgreSQL	Aurora (6 copies across 3 AZs, up to 15 replicas, fast failover). Aurora Serverless v2 for variable load. Aurora Global Database: cross-Region replication typically under 1 s, managed failover.
Key-value / document at any scale	DynamoDB (on-demand or provisioned). Global tables for multi-Region active-active. DAX for microsecond reads.
Redis-compatible durable primary database	Amazon MemoryDB (Multi-AZ transaction log, microsecond reads).
Cache in front of a database	Amazon ElastiCache (Valkey, Redis OSS, Memcached); serverless option.
MongoDB / Cassandra / graph	Amazon DocumentDB / Amazon Keyspaces / Amazon Neptune.
Time series	Amazon Timestream for InfluxDB. Timestream for LiveAnalytics **DEPRECATED** (closed to new customers 20 Jun 2025).

Need	Pick
Analytics / OLAP	Amazon Redshift; ad-hoc on S3 with Athena.

Storage

Service	Key facts
Amazon S3	Standard, Intelligent-Tiering, Standard-IA, One Zone-IA, Glacier Instant / Flexible / Deep Archive. Versioning, replication (CRR/SRR, RTC), Object Lock, Access Points, presigned URLs, Transfer Acceleration. S3 Express One Zone for single-digit-ms access.
Amazon EBS	gp3 (default; IOPS and throughput set independently), io2 Block Express (highest durability/IOPS), st1/sc1 HDD. Snapshots, snapshot archive tier, snapshot lock, Fast Snapshot Restore.
Amazon EFS	Multi-AZ NFS for Linux; lifecycle to IA and Archive; elastic throughput.
Amazon FSx	Windows File Server (SMB, AD), Lustre (HPC, S3-linked), NetApp ONTAP (multi-protocol, SnapMirror), OpenZFS.

Exam Tip: 'Analysts see only their country, never the ID column' = Lake Formation data filter. 'Corrections to trades → upserts in the lake' = Apache Iceberg. 'Measure campaign overlap with a partner, aggregates only' = Clean Rooms. 'Redis API but must not lose data' = MemoryDB, not ElastiCache.

Database HA, scaling & caching

Goal	Pattern
Survive an AZ failure	RDS Multi-AZ / Multi-AZ DB cluster; Aurora replicas in other AZs (failover target); DynamoDB and S3 are multi-AZ by default.
Scale reads	RDS read replicas (cross-Region allowed); Aurora up to 15 replicas behind the reader endpoint with replica auto scaling; caching.
Survive a Region failure	Aurora Global Database (managed cross-Region failover), DynamoDB global tables (active-active), cross-Region RDS read replica (promote), S3 CRR.
Connection storms from Lambda	RDS Proxy (pooling, faster failover, IAM auth).
Cache strategy	Lazy loading (cache-aside, TTL) for read-heavy data; write-through for freshness; TTL plus jitter to avoid stampedes; DAX for DynamoDB microsecond reads; CloudFront for HTTP responses.

Storage selection

Need	Pick
Shared Linux file system across many instances or containers	Amazon EFS
Windows SMB shares with Active Directory	FSx for Windows File Server
HPC scratch linked to an S3 data set	FSx for Lustre
NetApp features (SnapMirror, multi-protocol, dedup) after migration	FSx for NetApp ONTAP
On-prem apps need NFS/SMB access to S3 with a local cache	S3 File Gateway
Replace a physical tape library	Tape Gateway → S3 Glacier Deep Archive
Retain for 7 years at lowest cost	S3 Glacier Deep Archive + Object Lock compliance mode

2 · Security, Compliance & Governance (25%)

2.1 Identity Architecture

Service / feature	Use
IAM Identity Center	Workforce SSO across accounts. External IdP via SAML 2.0, automatic provisioning via SCIM, permission sets, ABAC with user attributes.
Amazon Cognito	User pools: customer sign-in, OAuth 2.0 / OIDC tokens, social and SAML federation. Identity pools: exchange tokens for temporary AWS credentials.
EKS Pod Identity	Map a Kubernetes service account to an IAM role through the EKS API; no per-cluster OIDC trust in each role (simpler than IRSA at scale).
IAM Roles Anywhere	Temporary AWS credentials for on-prem or other-cloud workloads using X.509 certificates from your CA (trust anchor + profile). Replaces long-term access keys.
AWS Verified Access	VPN-less access to internal apps; every request evaluated against Cedar policies using identity and device trust providers; logs every attempt.
Amazon Verified Permissions	Fine-grained application authorization with Cedar policies in a policy store; policy templates; Cognito or OIDC identity source; `IsAuthorized(WithToken)`. Per-tenant policy stores for isolation.
IAM Access Analyzer	External access findings, unused access findings, policy validation and custom policy checks in CI.
AWS Directory Service	AWS Managed Microsoft AD (trusts with on-prem AD), AD Connector (proxy, no directory data in AWS), Simple AD.

- **Policy evaluation** — explicit deny wins; an action must be allowed by every applicable layer: SCP, RCP, permissions boundary, session policy, identity- or resource-based policy.
- **Least privilege across accounts** — roles with trust policies; **external ID** for third parties (confused deputy); session policies to scope down; ABAC with `aws:PrincipalTag` / `aws:ResourceTag`.
- **Data perimeter** — trusted identities (`aws:PrincipalOrgID`), trusted resources (`aws:ResourceOrgID`), expected networks (`aws:SourceVpce`) in SCPs, RCPs and VPC endpoint policies.

> **Exam Tip:** 'Servers in another cloud need S3 access without access keys' = IAM Roles Anywhere. 'Contractors, no VPN, device posture per request' = Verified Access. 'Tenant admins define their own sharing rules' = Verified Permissions. '30 EKS clusters, trust policies hit size limits with IRSA' = Pod Identity.

2.2 Encryption & Key Management

Service	Key facts
AWS KMS	Customer managed keys with key policies and grants; automatic rotation (configurable period); multi-Region keys; AWS managed keys cannot be used cross-account. Envelope encryption for data over 4 KB.
Post-quantum (KMS)	**ML-DSA** signing keys (FIPS 140-3 Level 3 HSMs; RAW messages up to 4 KB, larger with EXTERNAL_MU). **ML-KEM** key encapsulation (ENCAPSULATE_DECAPSULATE).
AWS CloudHSM	Single-tenant FIPS-validated HSM cluster under your exclusive control; custom key store for KMS when regulators require dedicated hardware.
AWS Certificate Manager (ACM) / AWS Private CA	Public certificates free, auto-renewed (DNS validation); CloudFront certificates in us-east-1. Private CA for internal mTLS. TLS 1.3 security policies on ALB, NLB, CloudFront.
Secrets Manager / Parameter Store	Secrets Manager: automatic rotation, cross-account resource policy, replication. Parameter Store: config and SecureString, cheaper, no built-in rotation.

2.3 Multi-Account Governance

Building block	Role
AWS Organizations	OUs, consolidated billing, policies: SCPs (cap principals), **RCPs** (cap access to resources, including from outside the org), tag policies, backup policies, AI opt-out policies. Delegated administrators keep the management account clean.
AWS Control Tower	Landing zone with Log Archive and Audit accounts, controls (preventive = SCP/RCP, detective = Config, proactive = CloudFormation Hooks), **Account Factory** (and Account Factory Customization / AFT) for automated account provisioning, Region deny control.
AWS Resource Access Manager (RAM)	Share subnets, Transit Gateways, Resolver rules, IPAM pools, VPC Lattice service networks, Aurora clusters, Cloud WAN.
CloudTrail organization trail	One trail for all accounts to the Log Archive bucket; CloudTrail Lake for SQL queries; data events for S3, Lambda, Bedrock.
AWS Config	Rules, conformance packs, organization aggregator, remediation with SSM Automation.
Amazon GuardDuty	Threat detection from CloudTrail, VPC Flow Logs, DNS logs, EKS, RDS, Lambda, S3, malware protection; org-wide via delegated admin.
AWS Security Hub	Unified security findings and exposure view; **Security Hub CSPM** runs the compliance standards (AWS Foundational Best Practices, CIS, PCI DSS, NIST).
Amazon Security Lake	Central security data lake in your account: CloudTrail, VPC Flow Logs, Route 53, Security Hub and third-party logs normalised to **OCSF** in Parquet; subscribers query with Athena or SIEM tools.
Amazon Inspector + SSM Patch Manager	Continuous vulnerability scanning of EC2, ECR images, Lambda; patch baselines and maintenance windows across accounts.
AWS Shield / AWS WAF	DDoS protection (Advanced adds response team and cost protection) and L7 rules (managed rule groups, rate-based rules, Bot Control).

Exam Tip: 'Stop S3 buckets and KMS keys from being shared outside the org, even by bucket owners' = RCP with `aws:PrincipalOrgID` (exempt `aws:PrincipalIsAWSService`). 'Normalise security logs from AWS and third parties into one queryable lake' = Security Lake. 'Self-service accounts with a baseline' = Control Tower Account Factory (+ customizations).

2.4 Compliance Architecture

- **Data residency** — SCP / Control Tower Region deny with `aws:RequestedRegion`; Region-specific deployments; geographic Bedrock inference profiles; Outposts or Local Zones when data must stay in a location without a Region.
- **Audit trails** — CloudTrail org trail to a Log Archive bucket with S3 Object Lock (compliance mode) and log file validation; Bedrock model invocation logging to S3/CloudWatch in a central account.
- **Evidence** — **AWS Artifact** for AWS compliance reports and agreements (SOC, ISO, PCI, BAA); Config conformance packs and Security Hub CSPM for your own controls.
- **Risk assessment** — threat modelling (STRIDE), Well-Architected Tool reviews, mitigations mapped to preventive, detective and responsive controls.

AWS Audit Manager DEPRECATED is in maintenance mode and cannot be set up in new accounts since 30 Apr 2026. It is not on the SAP-C03 in-scope list; use Config conformance packs, Security Hub CSPM and AWS Artifact for evidence.

2.5 Security & Responsible AI Controls

Need	Control
Block harmful content, PII, denied topics, prompt attacks	**Bedrock Guardrails**; `ApplyGuardrail` API also covers models outside Bedrock (e.g. on SageMaker AI).
Agent acts on behalf of a user	**AgentCore Identity**: inbound authorization of the caller, OAuth 2.0 token vault for outbound calls (3-legged OAuth for user-delegated access).

Need	Control
Human approval before an AI-proposed change	**Step Functions** `.waitForTaskToken` approval step (Standard workflow), then automation runs the approved change.
Record every prompt and response	Bedrock model invocation logging + CloudTrail data events, central bucket with Object Lock.

2.6 Landing Zone Reference & Detection Matrix

OU / account	Purpose
Security OU	Log Archive account (org trail, Config history, Security Lake) and Audit/Security Tooling account (delegated admin for GuardDuty, Security Hub, Inspector, Access Analyzer). No workloads.
Infrastructure OU	Network account (Transit Gateway or Cloud WAN, IPAM, egress and inspection VPCs, Resolver endpoints) and shared services, shared through RAM.
Workloads OU	Prod and non-prod sub-OUs; SCPs and RCPs tighten towards prod.
Sandbox / Policy Staging / Suspended	Experiments with budget caps; test new SCPs before rollout; accounts waiting for closure with deny-all.
Deployments OU	Central CI/CD tooling accounts (CodePipeline, StackSets administration).

Question	Service
Who called which API, when?	AWS CloudTrail
What did this resource look like last Tuesday? Is it compliant?	AWS Config
Is something malicious happening right now?	Amazon GuardDuty
Which images and instances have exploitable CVEs?	Amazon Inspector
One prioritised view of findings and exposures	AWS Security Hub
Keep all security logs, normalised, for our SIEM	Amazon Security Lake
Which resources are shared outside our organization?	IAM Access Analyzer
Metrics, logs and alarms on performance	Amazon CloudWatch

3 · Cost-Optimized Architecture Design (15%)

3.1 Pricing Models & Rightsizing

Option	Use
Compute Savings Plans	$/hour commitment across EC2 (any family/Region), Fargate and Lambda. Best default for steady compute.
EC2 Instance Savings Plans	Deeper discount, locked to a family in one Region.
Reserved Instances	Standard or convertible, 1 or 3 years; zonal RIs also reserve capacity. RDS, ElastiCache, OpenSearch, Redshift reserved nodes.
Database Savings Plans	Commitment across database services (Aurora, RDS, DynamoDB, ElastiCache, DocumentDB, Neptune, Keyspaces, Timestream, OpenSearch, DMS); up to 35% off.
Spot / **Fargate Spot**	Up to 90% off for interruptible work; diversify instance types; capacity-optimised allocation.
On-Demand Capacity Reservations	Guaranteed capacity in an AZ, no term; combine with Savings Plans for the discount.

- **Graviton** instances: better price-performance for most Linux workloads.
- **AWS Compute Optimizer**: rightsizing for EC2, Auto Scaling groups, EBS, Lambda, ECS on Fargate, RDS and Aurora.
- **Step Functions**: Standard billed per state transition; Express billed by executions, duration and memory — high-volume short workflows are far cheaper on Express.
- **AI/ML cost**: Bedrock batch inference (about 50% off on-demand), prompt caching, smaller or distilled models, global cross-Region inference profile (about 10% cheaper), SageMaker AI serverless or asynchronous endpoints that scale to zero, Savings Plans for SageMaker AI.

3.2 Data Transfer & Network Cost

Situation	Cheaper design
NAT gateway charges for S3/DynamoDB traffic	Gateway VPC endpoints (no charge).
Cross-AZ chatter	AZ-aware routing, keep producer and consumer in the same AZ, NLB cross-zone off where appropriate.
Heavy origin egress to users	CloudFront caching + **Origin Shield** (extra caching layer, fewer origin fetches); data transfer from AWS origins to CloudFront is free.
Two VPCs exchange large volumes	VPC peering (no data processing charge) instead of Transit Gateway (per-GB processing).
Many consumers of one service	PrivateLink (per endpoint-hour + per GB) vs. full network connectivity.
Multi-Region replication	Replicate only what DR needs: S3 CRR with filters, DynamoDB global tables only for tables that need active-active.

3.3 Cost Allocation & Visibility

Tool	Use
Cost allocation tags	Activate user-defined and AWS-generated tags in Billing. Enforce with tag policies (standard values), SCPs with `aws:RequestTag` (prevent untagged creation) and Config `required-tags` (detect).
AWS Cost Categories	Rules map accounts, tags, services and charge types to business units. **Split charge rules** allocate shared costs proportionally, by fixed percentage or evenly.

Tool	Use
Split cost allocation data	Allocates shared EC2 node costs to ECS tasks or EKS pods by CPU and memory; aggregate by namespace, cluster or workload in CUR 2.0.
AWS Cost and Usage Report (CUR 2.0)	Line-item billing data via Data Exports to S3; query with Athena, visualise in Quick Sight.
AWS Cost Explorer	Analysis and forecasting, RI and Savings Plans recommendations and coverage.
AWS Budgets / Cost Anomaly Detection	Actual or forecast thresholds with actions (apply IAM policy/SCP, stop EC2/RDS); ML-based anomaly alerts per service, account or cost category.
Bedrock token usage	IAM principal attribution (cost per calling role/user, with principal tags as cost allocation tags) or application inference profiles with tags per team.

3.4 Storage, Database & Observability Cost

- **S3** — Intelligent-Tiering for unknown patterns; lifecycle to IA/Glacier for known patterns; Glacier Deep Archive for 7-year retention; EFS lifecycle to IA/Archive.
- **EBS** — gp2 → gp3 (cheaper, IOPS independent of size); io2 only when IOPS/durability demand it; delete unattached volumes.
- **DynamoDB** — on-demand for spiky/unknown; provisioned + auto scaling (+ reserved capacity) for steady; Standard-IA table class for rarely read data.
- **Aurora** — Serverless v2 for variable load; **I/O-Optimized** when I/O is above about 25% of spend; reserved instances or Database Savings Plans.
- **Backups** — AWS Backup lifecycle to cold storage; **Amazon Data Lifecycle Manager** policies for EBS snapshots and AMIs by tag, up to 4 schedules, cross-Region copy with its own retention, **snapshot archive tier** for monthly/yearly snapshots (90-day minimum, restore up to 72 h; not for daily snapshots).
- **Observability** — log retention per log group, CloudWatch Logs Infrequent Access class, subscription filters to drop noise, export old logs to S3 Glacier, embedded metric format or metric filters instead of high-cardinality custom metrics.

> **Exam Tip:** 'Charge a shared EKS cluster back to teams' = split cost allocation data. 'Spread shared networking and support costs across business units' = Cost Categories split charge rules. '7-year monthly EBS snapshots, restore in days is fine' = DLM archive tier. 'Prevent untagged resources' = SCP (preventive), not tag policies alone (reporting).

4 · Resilience, Migration & Business Continuity (15%)

4.1 Migration Strategy

7 Rs	Meaning / typical tool
Rehost	Lift and shift to EC2 — **AWS Transform MGN** (previously AWS Application Migration Service): continuous block replication, test and cutover launches.
Relocate	Move a VMware environment as is (VMware Cloud on AWS).
Replatform	Small changes to use managed services: self-managed DB → RDS/Aurora, app → ECS/Fargate or Elastic Beanstalk.
Refactor	Re-architect: microservices, serverless, event-driven, strangler fig.
Repurchase / Retire / Retain	Move to SaaS / switch off / keep on premises for now.

Service	Use
AWS Transform	Agentic migration and modernisation. Migrations (incl. VMware): discovery (inventory + network connections), dependency-based wave planning, landing zone, network migration to VPCs, server rehost with Transform MGN. Also .NET, mainframe (COBOL → Java refactoring) and other modernisation jobs.
AWS DMS + AWS SCT	Full load + CDC with minimal downtime; SCT (or DMS Schema Conversion) for heterogeneous engines (Oracle → Aurora PostgreSQL); data validation.
AWS DataSync	Online file and object migration with throttling, verification, filters and path mapping.
AWS Data Transfer Terminal	Book a secure AWS facility, bring your own storage devices and upload over high-speed fibre (two 100 Gbps links). Replaces most offline Snow use cases.
AWS Transfer Family	Keep partner SFTP/AS2 flows working after migration.

AWS Migration Hub DEPRECATED (incl. Strategy Recommendations, Refactor Spaces, Orchestrator) closed to new customers on 7 Nov 2025 — use AWS Transform. **AWS Mainframe Modernization** DEPRECATED runtimes are closed to new customers — refactor with AWS Transform for mainframe. **AWS Snowball Edge** DEPRECATED closed to new customers 7 Nov 2025; Snowmobile and Snowcone RETIRED — use DataSync or Data Transfer Terminal.

- **Strangler fig** — API Gateway or ALB routes a growing share of paths to new services until the monolith is retired.
- **Database decomposition** — one database per service; DMS or CDC streams keep data in sync during the transition.
- **Mainframe** — refactor (Transform for mainframe → Java on EKS/ECS, Db2 → Aurora PostgreSQL) or replatform to an emulator partner.

4.2 Disaster Recovery

Strategy	RPO	RTO	Mechanics & cost
Backup & restore	Hours	Hours	AWS Backup cross-Region copies; IaC to rebuild. Cheapest.
Pilot light	Minutes	Tens of minutes	Data replicated (Aurora Global, DynamoDB global tables, S3 CRR); compute off until failover.
Warm standby	Seconds	Minutes	Scaled-down full stack running; scale out on failover.
Multi-site active/active	Near zero	Near zero	Full capacity in 2+ Regions; Route 53 / Global Accelerator / ARC routing. Most expensive.

- **AWS Elastic Disaster Recovery** — continuous block-level replication of servers into a low-cost staging area; RPO seconds, RTO minutes; drills without impact.
- **Amazon Application Recovery Controller (ARC)** — routing controls, readiness checks, zonal shift and zonal autoshift.
- **Route 53 health checks + failover routing**; **Global Accelerator** for instant failover on static IPs.
- **Multi-cluster containers** — EKS or ECS clusters in 2+ AZs/Regions, GitOps to keep them identical, Route 53 / Global Accelerator across clusters, ECR replication.
- **AWS Resilience Hub** — describe the app (CloudFormation, Terraform state, EKS, Resource Groups, myApplications), set a resiliency policy (RTO/RPO for application, infrastructure, AZ and Region disruptions), run assessments, get recommended alarms, SOPs (SSM runbooks) and FIS tests, gate CI/CD on the result.

4.3 Business Continuity & Data Protection

Need	Control
Central backup for every account	AWS Backup + **backup policies** in Organizations, cross-account and cross-Region copy to a dedicated backup account.
Backups nobody (incl. root) can delete	**AWS Backup Vault Lock** compliance mode (immutable after the grace time).
WORM objects	**S3 Object Lock** — compliance mode (no one can shorten) vs governance mode (privileged bypass).
Protect EBS snapshots from ransomware	**EBS snapshot lock** — compliance mode (optional 1–72 h cooling-off) or governance mode; Recycle Bin for accidental deletion.
Prove failover works	**AWS Fault Injection Service** experiment templates with stop conditions (CloudWatch alarms) and a steady-state hypothesis; AZ power-interruption and Region connectivity scenarios.

Exam Tip: '1,800 VMware VMs, unknown dependencies, rebuild VLANs as VPCs' = AWS Transform. 'Which apps miss their RTO/RPO, and block bad releases' = Resilience Hub. 'Ransomware-proof backups across 40 accounts' = Backup policies + cross-account copy + Vault Lock compliance. 'Chaos test with automatic stop' = FIS with stop-condition alarms.

5 · Operational Excellence & Automation (20%)

5.1 Infrastructure as Code & Policy as Code

Tool	Key facts
AWS CloudFormation	Change sets, drift detection, stack policies, DeletionPolicy/UpdateReplacePolicy, nested stacks, cross-stack exports (same account and Region only).
CloudFormation StackSets	Deploy to many accounts and Regions; *service-managed* permissions with automatic deployment to new accounts in target OUs; concurrency and failure tolerance.
AWS CDK	Imperative code (TypeScript, Python, Java, C#, Go) synthesised to CloudFormation; constructs L1–L3; **Aspects** check or modify every construct at synth time; CDK Pipelines.
CloudFormation Hooks	Proactive checks before create/update/delete of stacks, resources, change sets or Cloud Control API calls. **Guard Hooks** (Guard rules in S3) or Lambda Hooks. WARN mode first, then FAIL. Control Tower proactive controls are Hooks.
Detective / corrective	AWS Config rules + auto-remediation, SCPs for hard limits.
Secrets in IaC	Dynamic references `{{resolve:secretsmanager:...}}` / `{{resolve:ssm-secure:...}}` ; never plain parameters.

5.2 CI/CD & DevSecOps

Service	Role
AWS CodePipeline	Orchestration; manual approval actions; **cross-account deploys**: artifact store per Region encrypted with a customer managed KMS key that target accounts can use, deploy actions assume roles in target accounts; build once, promote the same artifact.
AWS CodeBuild	Managed builds, buildspec, VPC access, reports; run scanners and tests.
AWS CodeDeploy	EC2/on-prem (in-place, blue/green), Lambda and ECS (canary, linear, all-at-once) with CloudWatch alarm rollback.
AWS CodeCommit	Managed Git. Back to general availability for new customers since November 2025.
Security in the pipeline	Amazon Inspector / ECR enhanced scanning (block on critical CVEs), image signing with AWS Signer and verification at deploy (admission control on EKS).
Amazon Bedrock Evaluations	Automatic or LLM-as-a-judge evaluation of models, prompts and knowledge bases (correctness, faithfulness, harmfulness) as a release gate for AI changes.

- **Deployment strategies** — rolling, blue/green, canary, linear, feature flags (AppConfig); automatic rollback on alarms; ECS deployment circuit breaker.

5.3 Observability

Service	Use
Amazon CloudWatch	Metrics, alarms (composite, anomaly detection), Logs + Logs Insights, dashboards, **cross-account observability** (monitoring account), Container Insights for ECS/EKS.
AWS X-Ray / AWS Distro for OpenTelemetry	Distributed tracing; ADOT = AWS-supported OpenTelemetry collector and SDKs for vendor-neutral metrics and traces.
CloudWatch Application Signals	Auto-instrumented service metrics, service map, **SLOs** with burn-rate alarms.
CloudWatch Synthetics / RUM	Canaries for synthetic transactions; real user monitoring of page load, errors and sessions in the browser.

Service	Use
Log aggregation	Cross-account subscription filters to a central destination (Firehose → S3, OpenSearch); org trail; S3 archive tiers for old logs.
AI/ML monitoring	Bedrock CloudWatch metrics (invocations, latency, tokens, throttles) and invocation logs; AgentCore Observability for agent traces; SageMaker Model Monitor for data and model drift.

5.4 Operational Automation & Resilient Systems

- **Health and capacity** — CloudWatch alarms → EventBridge → SSM Automation runbooks; **Service Quotas** usage alarms and automated increase requests.
- **AWS Systems Manager** — Automation (rate control: concurrency + error threshold, multi-account/Region), Run Command, Patch Manager, State Manager, Session Manager, Parameter Store, OpsCenter.
- **Safety constraints** — circuit breakers, retries with exponential backoff and jitter, bulkheads (separate pools/queues per dependency), rate limits (API Gateway throttling, SQS buffering), timeouts.
- **Scaling** — target tracking (default), step (large sudden changes), scheduled (known events), predictive (recurring daily/weekly patterns); scale on SQS backlog per instance.
- **Self-healing** — ELB health checks on Auto Scaling groups, ECS task replacement, Lambda retries + destinations/DLQ, Route 53 health checks.

Exam Tip: 'Block non-compliant resources in every account, test first' = Guard Hook (WARN → FAIL) deployed by service-managed StackSets. 'Same artifact through dev → prod across accounts, no long-term keys' = CodePipeline cross-account roles + CMK-encrypted artifact store. 'Page loads slow but backend healthy' = CloudWatch RUM. 'SLO alert before the error budget runs out' = Application Signals burn-rate alarm.

6 · Scenario → Answer Patterns

If the question says…	Think
'one sign-in for 200 accounts with Okta'	IAM Identity Center + SAML + SCIM
'restrict all accounts to two Regions'	SCP / Control Tower Region deny (`aws:RequestedRegion`)
'no resource shared outside the organization'	RCP + SCP data perimeter with `aws:PrincipalOrgID` / `aws:ResourceOrgID`
'hundreds of VPCs, segmentation, global'	Cloud WAN (or Transit Gateway route tables per Region)
'inspect all egress centrally'	Inspection/egress VPC with Network Firewall behind TGW or Cloud WAN service insertion
'expose service to customer VPCs, overlapping CIDRs'	PrivateLink (NLB endpoint service)
'multi-Region writes, active-active'	DynamoDB global tables (Aurora Global Database when one writer Region is acceptable)
'RPO seconds, RTO minutes for servers'	Elastic Disaster Recovery
'lift and shift 500 servers fast'	AWS Transform MGN (planning with AWS Transform)
'Oracle → Aurora PostgreSQL with minimal downtime'	SCT/DMS Schema Conversion + DMS full load + CDC
'500 TB, limited bandwidth, no Snowball'	DataSync with throttling, or Data Transfer Terminal
'RAG over company documents'	Bedrock Knowledge Bases (+ metadata filters per tenant)
'block PII and harmful content in all AI apps'	Bedrock Guardrails / ApplyGuardrail
'cheapest for nightly bulk LLM jobs'	Bedrock batch inference

If the question says…	Think
'chargeback by team on a shared cluster'	Split cost allocation data + Cost Categories
'unexpected spend spike alert'	Cost Anomaly Detection
'Kubernetes on own servers, AWS-managed control plane'	EKS Hybrid Nodes
'low latency for users in a metro area, no data centre'	Local Zones
'AWS APIs and RDS on the factory floor'	Outposts
'patch 2,000 instances within 7 days with reporting'	SSM Patch Manager + maintenance windows + compliance
'no SSH or bastion'	SSM Session Manager

7 · Removed from SAP-C03 & Well-Architected

These were listed for SAP-C02 and are no longer on the SAP-C03 in-scope list (AWS notes that concepts can still appear implicitly):

- **Categories**: all AWS IoT services; Amazon WorkSpaces and AppStream 2.0; Elastic Transcoder and Kinesis Video Streams.
- **Individual AI services** (replaced by Bedrock and SageMaker AI): Comprehend, Fraud Detector, Kendra, Lex, Personalize, Polly, Rekognition, Textract, Transcribe, Translate.
- **Individual services**: CodeGuru, Lightsail, Managed Blockchain, Pinpoint, SES, Amplify, App Runner, Batch, CodeArtifact, Data Exchange, Device Farm, Health Dashboard, License Manager, Proton, Service Catalog, Wavelength.
- **Explicitly out of scope**: Braket, Comprehend Medical, Transcribe Medical, IVS, Lightsail, RDS Custom for Oracle, DeepComposer, DeepRacer, Elemental MediaConvert / MediaLive, Ground Station.

Pillar	What the 'best' answer usually optimises
Operational Excellence	Automation, IaC, observability, small reversible changes.
Security	Least privilege, preventive guardrails, encryption, traceability.
Reliability	Multi-AZ first, tested recovery, self-healing, quotas managed.
Performance Efficiency	Managed and serverless services, right data store, edge.
Cost Optimization	Commitments for steady load, Spot/serverless for variable, allocation and accountability.
Sustainability	Graviton, rightsizing, managed services, lifecycle policies.

> **Exam Tip:** SAP-C03 answers are long and all plausible. Underline the constraint words — 'least operational overhead', 'most cost-effective', 'without code changes', 'cannot leave the EU' — and eliminate any option that violates one, even if it would work technically.

cloudninja.pro — Free AWS certification practice exams, study guides, and cheat sheets.
Pair this cheat sheet with our free SAP-C03 practice exam for the best results. Exam facts verified against the official SAP-C03 exam guide and AWS documentation, November 2026.