AWS Certified

Solutions Architect — Associate

SAA-C03 · QUICK REFERENCE CHEAT SHEET

26% **Domain 1** Design Secure Architectures	**24%** **Domain 2** Design Resilient Architectures	**30%** **Domain 3** Design High-Performance Architectures	**20%** **Domain 4** Design Cost-Optimized Architectures

65 scored + 20 unscored 85 Questions Total	**130 Minutes** Exam Duration	**720 / 1000** Minimum Passing Score	**Multiple Choice + Multiple Response** Question Types	**Associate Level** Prerequisites	**Scaled Score 100–1000** Scoring Range

DOMAIN 1 · DESIGN SECURE ARCHITECTURES (26%)

IAM · Encryption · VPC Security · Detective Controls

26%
of exam

IAM Core Concepts

Policy Evaluation	Explicit Deny → Allow → Implicit Deny. Evaluated across all applicable policies.
IAM Roles	Assumed via STS AssumeRole — grants temporary credentials. Preferred over access keys.
Permission Boundaries	IAM policy that sets maximum permissions an entity can have. Does not grant access itself.
SCPs	Service Control Policies — org-wide guardrails applied to OUs and accounts. Cannot grant permissions, only restrict.
IAM Access Analyzer	Identifies resources shared with external principals. Validates policies for overly permissive access.
Identity Federation	SAML 2.0 or OIDC federation. Web Identity Federation for mobile apps (Cognito recommended).
Condition Keys	aws:SourceIp, aws:RequestedRegion, aws:MultiFactorAuthPresent, aws:PrincipalTag.

Encryption & Secrets

Service	Key Feature	Use Case
AWS KMS	CMK managed keys; envelope encryption; key policies + grants	Encrypt EBS, S3, RDS, Lambda env vars
CloudHSM	Dedicated HSM; FIPS 140-2 Level 3; you manage keys	Custom key store; strict compliance (PCI)
ACM	Free TLS/SSL certs; auto-renewal; integrates with ALB/CloudFront	HTTPS for public-facing endpoints
Secrets Manager	Auto-rotate RDS/Redshift/custom secrets; cross-account access	Database credentials, API keys with rotation
SSM Parameter Store	Hierarchy /prod/db/password; SecureString via KMS; free tier	Config values, non-rotating secrets, feature flags

S3 Security

SSE-S3	AWS-managed keys (AES-256). Default encryption. Header: x-amz-server-side-encryption: AES256.
SSE-KMS	Customer-managed CMK. Audit via CloudTrail. Header: x-amz-server-side-encryption: aws:kms.
SSE-C	Customer-provided key. AWS encrypts but never stores the key. HTTPS required.
Client-Side Encryption	Encrypt before upload. AWS S3 Encryption Client or custom. Full client key control.
Bucket Policies	Resource-based policies. Deny HTTP (aws:SecureTransport: false), enforce encryption, restrict VPC.
Block Public Access	Account-level or bucket-level toggle. Overrides ACLs and bucket policies for public access.
Object Lock (WORM)	Compliance mode: nobody can delete (including root). Governance mode: only root can override.
Access Points	Per-team named endpoints with scoped policies. Simplify access control for shared buckets.

Threat Detection & Edge Protection

Service	What It Does	Key Detail
GuardDuty	ML-based threat detection on VPC Flow Logs, DNS, CloudTrail, S3 data events	No agents needed; enable per region; 30-day trial
Inspector	Automated CVE scanning for EC2 instances, Lambda functions, ECR container images	Agent required for EC2; agentless option available
Macie	S3 sensitive data discovery — PII, credentials, financial data using ML	Auto-discovers S3 buckets; creates findings
AWS WAF	Layer 7 firewall — IP rules, geo-blocking, rate limiting, SQL injection, XSS protection	Attach to ALB, CloudFront, API GW, AppSync
Shield Standard	Free automatic DDoS protection (L3/L4) for all AWS customers	Included at no cost; protects ELB, CloudFront, Route 53

Shield Advanced	$3,000/month. L7 DDoS protection + WAF + DDoS Response Team (DRT)	Cost protection; near real-time visibility; 24/7 support
Firewall Manager	Centrally manage WAF rules, SGs, Shield, NACLs across AWS Org accounts	Requires AWS Organizations; auto-remediate non-compliant
Network Firewall	Managed stateful/stateless inspection at VPC level. Suricata-compatible IDS/IPS rules	Deploy in each VPC or centralized with Transit GW

Compliance & Audit

CloudTrail	Records all API calls (who, what, when, from where). 90-day console history; S3+CloudWatch for long-term. Multi-region trail recommended.
AWS Config	Continuous resource inventory + compliance rules (managed or custom Lambda). Config Rules → auto-remediation via SSM.
Security Hub	Aggregates findings from GuardDuty, Inspector, Macie, Config, IAM AA into a single dashboard. CSPM.
AWS Artifact	On-demand access to AWS compliance reports: SOC 1/2/3, PCI DSS, ISO 27001, HIPAA eligibility.

DOMAIN 2 · DESIGN RESILIENT ARCHITECTURES (24%)

Multi-AZ · Multi-Region · Decoupling · DR Strategies

24%
of exam

High Availability Foundations

Multi-AZ	Synchronous replication across AZs for automatic failover. HA, not read scaling. RDS Multi-AZ, ALB, Aurora.
Multi-Region	Protects against entire region failure. RPO/RTO in minutes to near-zero. Aurora Global DB, Route 53 failover.
Elasticity	Automatically scale capacity to match demand. Use ASG + ALB for stateless tiers. Store state in ElastiCache/DynamoDB.
Health Checks	ELB health checks for ASG — use ELB type for web apps (not just EC2). Route 53 health checks drive DNS failover.

Disaster Recovery Strategies

Strategy	RTO	RPO	Cost	Description
Backup & Restore	Hours	Hours	★■■■	Backup to S3/Glacier. Restore from scratch. Cheapest.
Pilot Light	10–30 min	Minutes	★★■■	Core DB replicated; app servers stopped. Start on failover.
Warm Standby	Minutes	Minutes	★★★■	Scaled-down running copy. Scale up on failover.
Multi-Site Active-Active	~0	~0	★★★★	Full capacity in multiple regions. Route 53 distributes traffic.

■ *Exam tip: Multi-AZ = HA (synchronous); Read Replicas = performance (asynchronous). Pilot Light = DB runs, app servers stopped.*

Decoupling & Messaging

Service	Type	Key Characteristics
SQS Standard	Queue (pull)	At-least-once delivery, best-effort ordering, unlimited throughput. Visibility timeout default 30s.
SQS FIFO	Queue (pull)	Exactly-once processing, strict ordering. 3,000 msg/s batched, 300 msg/s unbatched.
SNS	Pub/Sub (push)	Fan-out to SQS, Lambda, HTTP, email, SMS. Filter policies per subscription.
EventBridge	Event bus	90+ AWS sources + custom events. Rules → targets. Event Archive + Replay. Pipes.
Kinesis Data Streams	Streaming	Real-time data streaming. 1 MB/s per shard. 24-hr default retention (up to 365 days).
Step Functions	Orchestration	Standard (1yr, exactly-once) vs Express (5min, at-least-once). Visual workflow.
Amazon MQ	Broker	Managed ActiveMQ/RabbitMQ. Lift-and-shift from on-prem messaging (AMQP, STOMP, MQTT).

Load Balancers

Type	Protocol	Key Features	Best For
ALB	HTTP/HTTPS/WebSocket	Path/host/header-based routing; Lambda & ECS targets; WAF; Cognito auth	Microservices, web apps, containers
NLB	TCP/UDP/TLS	Ultra-low latency; static IP / Elastic IP; preserves client IP; PrivateLink source	High-throughput, gaming, IoT, static IP needs
GWLB	IP (GENEVE 6081)	Transparent bump-in-the-wire; scales 3rd-party virtual appliances	Firewall, IDS/IPS, deep packet inspection

Auto Scaling & Fault Tolerance

Target Tracking	Maintain a CloudWatch metric at a target value (e.g., CPU at 60%). Recommended default policy.
Step Scaling	Respond to CloudWatch alarms with step adjustments. Good for fine-grained control.

Scheduled Scaling	Pre-scale for known traffic patterns (e.g., business hours, weekly peaks).
Predictive Scaling	ML-based forecast + proactive scale-out. Combine with Target Tracking for best results.
Cooldown Period	Default 300 seconds after scaling activity. Target Tracking manages this automatically.
Instance Refresh	Rolling replacement of instances to update AMI or launch template. Can set min healthy %.
Spot + On-Demand Mix	ASG Mixed Instances Policy: base On-Demand capacity + Spot for up to 90% cost savings.

DOMAIN 3 · DESIGN HIGH-PERFORMANCE ARCHITECTURES (30%)

Compute · Storage · Databases · Networking · Caching

30%
of exam

EC2 Instance Families

Family	Type	Use Case
t3/t4g	Burstable General Purpose	Dev/test, low-cost web servers, microservices
m6i / m7g	Balanced General Purpose	Application servers, small/medium databases, backend
c6i / c7g	Compute Optimized	HPC, batch, video encoding, gaming, scientific modeling
r6i / x2idn	Memory Optimized	In-memory DBs, SAP HANA, real-time big data analytics
i3 / i4i	Storage Optimized	NoSQL DBs, data warehousing, high IOPS local NVMe SSD
p3 / p4d	Accelerated (GPU)	ML training, deep learning, HPC, video rendering
inf1 / inf2	Accelerated (Inferentia)	High-throughput, low-cost ML inference at scale
hpc6a / hpc7g	HPC	Tightly coupled MPI workloads, computational fluid dynamics

Storage Performance

Service	Max IOPS / Throughput	Key Characteristics
EBS gp3	16,000 IOPS / 1,000 MB/s	Default SSD; decouple IOPS from size; 20% cheaper than gp2
EBS io2 Block Express	256,000 IOPS / 4,000 MB/s	Mission-critical DBs; 99.999% durability; multi-attach
EBS st1	500 MB/s throughput	Sequential big data; log processing; cannot be boot volume
EBS sc1	250 MB/s throughput	Lowest cost HDD; cold data; infrequent sequential access
Instance Store	Millions IOPS (NVMe)	Ephemeral; data lost on stop/terminate; highest raw performance
EFS	10+ GB/s / bursting	POSIX NFS; multi-AZ; auto-scaling; EFS IA for cost savings
FSx for Lustre	Hundreds GB/s	HPC & ML; integrates with S3 as data repository; sub-ms latency
FSx for Windows	2 GB/s / 350K IOPS	SMB protocol; Active Directory; DFS namespaces; shadow copies

Database Selection Guide

Service	Type	Sweet Spot
Aurora MySQL/PostgreSQL	Relational	Up to 5× MySQL, 3× PostgreSQL. 6 copies/3 AZs. Up to 15 read replicas. Global DB RPO ~1s.
RDS (MySQL, PG, SQL Server…)	Relational	Managed RDBMS. Multi-AZ standby. Up to 5 read replicas. Automated backups 1–35 days.
DynamoDB	Key-Value + Document	Single-digit ms at any scale. Serverless. Global Tables. DAX cache. TTL. PITR 35 days.
ElastiCache Redis	In-Memory Cache	Sorted sets, pub/sub, persistence, geo commands, multi-AZ, replication groups.
ElastiCache Memcached	In-Memory Cache	Simple caching, multithreaded, no persistence, horizontal scaling via sharding.
Redshift	OLAP Data Warehouse	Columnar storage, MPP. Redshift Spectrum queries S3. RA3 nodes separate compute/storage.
Neptune	Graph	Social networks, fraud detection, knowledge graphs. Supports Gremlin, SPARQL, openCypher.
DocumentDB	Document	MongoDB-compatible. 6 copies/3 AZs. Up to 15 read replicas. Auto-scales storage.
Timestream	Time-Series	IoT telemetry, DevOps metrics. Auto-tiers hot/warm/cold. Built-in time functions.
Keyspaces	Wide-Column	Apache Cassandra-compatible. Serverless. CQL. Pay-per-request or provisioned.

Caching & Performance Patterns

CloudFront	CDN with 400+ PoPs. Origins: S3, ALB, EC2, custom HTTP. OAC for S3 access control. Signed URLs/Cookies.
DAX (DynamoDB)	In-memory cache for DynamoDB. Microsecond read latency. Write-through. API-compatible with DynamoDB.
RDS Read Replicas	Async replication. Offload read traffic. Promote to standalone DB for DR. Cross-region replication supported.
Global Accelerator	Anycast static IPs → AWS edge → private network to endpoint. 60% faster routing. Instant failover.
API Gateway Cache	Cache responses by 0.5 GB–237 GB. TTL 0–3600s. Reduce backend calls for stable responses.
Lambda SnapStart	Pre-initialized execution environment (Java). Reduces cold start from seconds to milliseconds.

DOMAIN 4 · DESIGN COST-OPTIMIZED ARCHITECTURES (20%)

Compute Savings · Storage Tiering · Managed Services · Cost Tools

20%
of exam

EC2 Purchasing Options

Option	Discount vs On-Demand	Commitment	Best For
On-Demand	0%	None	Unpredictable workloads, dev/test, short-term
Reserved Instances (1yr)	~40%	1 year	Steady-state workloads with known instance type/region
Reserved Instances (3yr)	~60%	3 years	Long-running baseline workloads, maximum savings
Compute Savings Plans	up to 66%	1 or 3 yr	Flexible across EC2, Fargate, Lambda — most flexible
EC2 Instance Savings Plans	up to 72%	1 or 3 yr	Specific instance family + region. Highest EC2 discount.
Spot Instances	up to 90%	None	Stateless, fault-tolerant, flexible workloads and batch jobs
Dedicated Hosts	Varies	On-Demand/1yr/3yr	BYOL (SQL Server, Windows Server) per-socket licensing
Dedicated Instances	~10% premium	None	Single-tenant hardware, no BYOL benefit, compliance needs

S3 Storage Cost Optimization

Storage Class	Min Storage	Retrieval	Monthly Cost (approx)
S3 Standard	None	ms	Highest — frequent access
S3 Intelligent-Tiering	None	ms	Auto-moves between tiers; monitoring fee per obj
S3 Standard-IA	30 days	ms	~40% less than Standard; per-GB retrieval fee
S3 One Zone-IA	30 days	ms	~20% less than Standard-IA; single AZ only
S3 Glacier Instant Retrieval	90 days	ms	~68% less than Standard; rare access
S3 Glacier Flexible Retrieval	90 days	min–hrs	~85% less; bulk retrieval free; backups
S3 Glacier Deep Archive	180 days	hrs	Lowest cost (~$0.00099/GB/mo); 7–10yr retention

■ *Use S3 Lifecycle Policies to automatically transition objects between storage classes based on age or access patterns.*

Serverless & Managed Service Cost Benefits

Lambda	Pay only for invocation duration (100ms increments) + requests. No idle cost. Free tier: 1M requests/mo.
Fargate	Pay for vCPU + memory per second. No EC2 instance management. Spot Fargate for up to 70% savings.
Aurora Serverless v2	Scales by 0.5 ACU increments. Pay per ACU-hour. Ideal for variable/unpredictable DB workloads.
DynamoDB On-Demand	Pay per read/write request. No capacity planning. Use provisioned + auto-scaling for predictable workloads.
SQS / SNS / EventBridge	Pay per API call / event. No servers to manage. Use to decouple and reduce synchronous compute costs.
S3 + Athena	Replace expensive always-on servers with S3 storage + $5/TB Athena query. Ideal for reporting/analytics.

Cost Management Tools

Tool	Purpose	Key Action
Cost Explorer	Visualize and forecast AWS spend by service/tag/account	Identify top cost drivers; RI/SP purchase recommendations
AWS Budgets	Set cost, usage, RI/SP coverage thresholds and receive alerts	Alert via email/SNS; Budget Actions to auto-respond
Trusted Advisor	Best practice checks: cost, performance, security, fault tolerance	Underutilized EC2, idle load balancers, unattached EIPs

Compute Optimizer	ML-based right-sizing for EC2, ASG, Lambda, EBS, ECS on Fargate	Identifies over/under-provisioned resources with savings estimates
S3 Storage Lens	Org-wide S3 usage analytics, activity metrics, cost optimization	Find buckets with low access, multipart upload remnants
Cost Allocation Tags	Tag resources (project, team, env) to split costs in reports	Activate tags in Billing console for cost attribution

Networking Cost Optimization

VPC Endpoints	Gateway endpoints (S3, DynamoDB) are free. Interface endpoints eliminate NAT Gateway data costs for AWS services.
NAT Gateway	Charged per GB processed. Use VPC Endpoints for S3/DynamoDB. Consider NAT Instance for dev/test (cheaper, less HA).
Data Transfer	Free: same AZ (same ENI). AZ-to-AZ: $0.01/GB each way. Internet egress: $0.09/GB (CF cheaper). Use CloudFront.
CloudFront	Egress via CloudFront cheaper than direct EC2/ALB internet egress. Reduces origin load and data transfer costs.
Direct Connect	Consistent bandwidth; lower per-GB data transfer than internet. 1–10 Gbps dedicated connections.

NETWORKING · SERVERLESS · WELL-ARCHITECTED FRAMEWORK

VPC Deep Dive · Route 53 · Serverless Patterns · Architecture Best Practices

VPC Deep Dive

Concept	Detail
Subnets	Public subnet = IGW route + auto-assign public IP. Private subnet = no IGW route. One subnet per AZ minimum.
NAT Gateway	AZ-scoped; redundant within AZ. For multi-AZ resilience, deploy one NAT GW per AZ. Managed, scales to 45 Gbps.
VPC Peering	Non-transitive (A↔B, B↔C ≠ A↔C). No overlapping CIDRs. No bandwidth bottleneck. Works cross-account/region.
Transit Gateway	Hub-and-spoke for VPCs + on-prem. Transitive routing. Attach VPCs, VPNs, Direct Connect GWs. Route tables per attachment.
VPC Endpoints	Gateway (S3, DynamoDB): free, route table entry. Interface (PrivateLink): ENI in subnet, all services, per-hour + per-GB cost.
Security Groups	Stateful (return traffic auto-allowed). Instance level. Allow rules only. All rules evaluated. Default: deny all inbound.
NACLs	Stateless (must allow both directions). Subnet level. Allow + Deny rules. Evaluated in order (lowest # first). Default: allow all.
Flow Logs	Capture IP traffic metadata (not content). VPC/subnet/ENI level. Send to S3 or CloudWatch Logs. Use for troubleshooting.
VPN	Site-to-Site VPN: 1.25 Gbps per tunnel, two tunnels for HA. Client VPN: OpenVPN for remote users. Uses internet.
Direct Connect	Dedicated private circuit 1–100 Gbps. Consistent latency. DX Gateway for multi-VPC/multi-region access.

Route 53 Routing Policies

Policy	Use Case	Health Checks?
Simple	Single resource. No health check support. Returns all values for multi-value.	No
Weighted	A/B testing, gradual blue/green deployment. Weight 0 = no traffic.	Yes
Latency-Based	Route to AWS region with lowest latency for user. Not geographically closest.	Yes
Failover	Active/passive setup. Primary fails → secondary. Requires health check on primary.	Yes (required)
Geolocation	Route based on user country or continent. Default record for unmatched locations.	Yes
Geoproximity	Route based on geographic location with bias (+/-). Use Traffic Flow required.	Yes
Multi-Value Answer	Up to 8 healthy records returned randomly. Client-side load balancing. Not a replacement for ELB.	Yes
IP-Based	Route based on client IP CIDR. Use for known ISP/corporate network routing.	Yes

Serverless Architecture Patterns

API GW + Lambda	REST API: full features, caching, usage plans, API keys. HTTP API: 70% cheaper, OIDC/OAuth2. WebSocket: real-time.
Event-Driven Lambda	S3 → Lambda (object events), SQS → Lambda (polling), SNS → Lambda (push), EventBridge → Lambda (scheduled/rules).
Lambda Limitations	Max 15 min execution. 10 GB memory. 512 MB–10 GB /tmp. 1000 concurrent (default, soft). 250 MB deployment package.
Provisioned Concurrency	Pre-initialized Lambda environments. Eliminates cold starts. Use for latency-sensitive APIs. Costs more.

ECS Fargate	Serverless containers. No EC2 to manage. Task = running container. Service = desired count + load balancer integration.
App Runner	Container or source code → managed web service. Auto TLS, auto-scaling. Simpler than ECS for web apps.
Cognito	User Pools: user directory for authentication (sign-up/sign-in, MFA, federation). Identity Pools: swap tokens for AWS credentials.
AppSync	Managed GraphQL API. Real-time subscriptions via WebSocket. Resolvers for DynamoDB, Lambda, HTTP, RDS, OpenSearch.

AWS Well-Architected Framework — 6 Pillars

Pillar	Core Focus	Key Design Principles
Operational Excellence	Run & monitor systems, continuously improve	IaC (CloudFormation/CDK); small reversible changes; annotate docs; anticipate failure
Security	Protect data, systems & assets	Strong identity (IAM roles, least privilege); enable traceability (CloudTrail); encrypt everywhere
Reliability	Recover from failures, meet demand	Automatic recovery (CloudWatch Alarms → ASG); horizontal scaling; stop guessing capacity
Performance Efficiency	Use resources efficiently	Democratize advanced tech (managed services); go global in minutes; experiment frequently
Cost Optimization	Deliver value at lowest price	Consumption model (serverless/Spot); measure efficiency; avoid undifferentiated heavy lifting
Sustainability	Minimize environmental impact	Understand your impact; maximize utilization; adopt efficient tech; use managed services

MASTER QUICK REFERENCE — MOST FREQUENTLY TESTED

SAA-C03 · Know These Cold

Know These Cold

Service / Concept	Category	Remember This
Amazon S3	Object Storage	11 nines durability; unlimited capacity; storage classes; lifecycle policies; Glacier for archive
Amazon EC2	Compute	Instance families; purchasing options; ASG + ALB; key pairs; EBS-backed vs instance store
Auto Scaling Groups	Elasticity	Target Tracking (recommended); Step; Scheduled; Predictive. ELB health checks for web apps.
Amazon RDS Multi-AZ	HA Database	Synchronous standby replica — HA, NOT read scaling. Automatic failover. Same region only.
RDS Read Replicas	Performance	Asynchronous replication. Up to 5 (RDS) / 15 (Aurora). Cross-region supported. Can promote.
Amazon Aurora	Relational DB	6 copies across 3 AZs; auto-grows to 128 TB; Aurora Global DB RPO ~1s / RTO <1min
DynamoDB	NoSQL	Single-digit ms; serverless; DAX for µs reads; Global Tables; Streams → Lambda; PITR 35 days
ElastiCache Redis	Caching	Session store, leaderboards, pub/sub, Sorted Sets. Multi-AZ with auto-failover.
Amazon VPC	Networking	Public subnet = IGW route; Private = no IGW; SGs stateful (instance); NACLs stateless (subnet)
VPC Endpoints	Cost + Security	Gateway (S3/DynamoDB) free; Interface (PrivateLink) paid. Keeps traffic in AWS network.
Transit Gateway	Connectivity	Hub-and-spoke; transitive routing; attach VPCs + VPN + DX. Replaces VPC peering at scale.
Application Load Balancer	Load Balancing	HTTP/HTTPS; path/host routing; Lambda + ECS targets; WAF integration; Cognito auth
Network Load Balancer	Load Balancing	TCP/UDP; static IP / Elastic IP; ultra-low latency; PrivateLink source; preserves client IP
Amazon CloudFront	CDN	400+ PoPs; OAC for S3; Signed URLs/Cookies; WAF; Lambda@Edge; Cache Behaviors
Route 53	DNS	Latency / Failover / Weighted / Geolocation / Geoproximity / Multi-Value routing policies
AWS Lambda	Serverless Compute	Max 15 min; 10 GB memory; 1000 concurrent default; Provisioned Concurrency = no cold start
Amazon SQS	Decoupling	Standard (at-least-once) vs FIFO (exactly-once, ordered). DLQ after maxReceiveCount. Long polling.
Amazon SNS	Pub/Sub	Fan-out: SNS → multiple SQS queues. Push model. Filter policies per subscription.
Amazon EventBridge	Event Bus	90+ AWS sources; custom events; rules → targets; Archive + Replay; Pipes for streaming
AWS KMS	Encryption	CMK (Customer Managed Key); envelope encryption; key policies; CloudTrail auditable
AWS IAM	Identity	Explicit Deny wins; roles over keys; permission boundaries; SCPs for org guardrails
Amazon CloudWatch	Monitoring	Metrics (default 5min, detailed 1min); Alarms; Logs; Insights; Dashboards; Events
AWS CloudTrail	Audit	All API calls logged; 90-day console; S3 for long-term; multi-region trail recommended

AWS Config	Compliance	Resource inventory + compliance rules + auto-remediation. Track config changes over time.
AWS Direct Connect	Hybrid	Private dedicated circuit 1–100 Gbps. Consistent latency. DX Gateway for multi-VPC.
AWS Storage Gateway	Hybrid Storage	File GW (S3-backed NFS/SMB) / Volume GW (iSCSI/EBS) / Tape GW (VTL → Glacier)
Spot Instances	Cost	Up to 90% savings. Interruption 2-min notice. Use for fault-tolerant, batch, stateless workloads.
Savings Plans	Cost	Compute SP (EC2+Lambda+Fargate, most flexible) vs EC2 Instance SP (highest EC2 discount)
AWS Organizations	Governance	Consolidated billing; SCPs; OU hierarchy; AWS Config org aggregator; multi-account strategy
Disaster Recovery	Resilience	Backup/Restore (hours) → Pilot Light (30min) → Warm Standby (minutes) → Multi-Site (~0)

EXAM DAY REMINDERS

720/1000 to pass | 85 questions (65 scored + 20 unscored) | 130 minutes | No penalty for guessing — answer every question!

Multi-AZ = HA (synchronous replication) | Read Replicas = read performance (asynchronous) | Never confuse these!

SQS + Lambda = decoupling & async | SNS + SQS = fan-out pattern | EventBridge = event-driven architecture

Spot = fault-tolerant batch (up to 90% off) | Reserved/Savings Plans = predictable steady-state | On-Demand = short-term/unknown

VPC Endpoint Gateway (S3/DynamoDB) = FREE | Interface endpoints = per-hour + per-GB | Always use endpoints over NAT GW for AWS services

NLB = static IP, high-throughput TCP/UDP | ALB = HTTP rules, path/host routing | GWLB = 3rd-party appliances (bump-in-the-wire)

IAM Roles > Access Keys | Explicit Deny always wins | SCPs restrict but never grant | Permission Boundaries set max permissions

CloudFormation = IaC; CDK = IaC in code | StackSets = multi-account/region deploy | Drift detection monitors config changes