

PRACTICE QUESTIONS

AWS

SAP-C03

AWS Certified Solutions Architect - Professional

SHARED BY [u/Crisp-Nebula-8790](#) on aiexamcerts

EDITION	Free Set 2
QUESTIONS	10
SHARED	October 15, 2026
EXAM	AWS SAP-C03, AWS Certified Solutions Architect - Professional

Shared by [u/Crisp-Nebula-8790](#) on aiexamcerts

10 practice questions for the AWS SAP-C03 exam, AWS Certified Solutions Architect - Professional, each followed by its answer and a short explanation. More free sets and the full practice exam:

www.aiexamcerts.com/aws/sap-c03. Spotted a mistake? Email support@aiexamcerts.com with the exam code SAP-C03 and the question number.

1. A company runs a shared network for 30 teams, and their VPC IP address ranges must never overlap. Which feature provides a central pool from which each team allocates its own CIDR blocks?
 - A AWS Cloud Map namespaces shared by every team so that services can discover each other
 - B Route 53 Resolver inbound endpoints
 - C Amazon VPC IP Address Manager with pools shared through AWS RAM
 - D AWS Global Accelerator static IP addresses

Answer: C

IP Address Manager (IPAM) tracks address space and allocates non-overlapping CIDR blocks from pools, and sharing pools through RAM lets each team draw from them. Cloud Map handles service discovery, Route 53 Resolver handles DNS queries, and Global Accelerator provides static anycast addresses for traffic entry. None of those prevent overlapping VPC ranges.

2. AWS Resource Access Manager can share a Transit Gateway with accounts in an organization so that those accounts can attach their VPCs to it.
 - A True
 - B False

Answer: A (True)

RAM supports sharing Transit Gateways, and shared accounts can create VPC attachments to the shared gateway. This keeps one central routing hub while each account owns its workloads. The statement is true.

3. An organization needs one place to define, deploy and monitor approved network security rules across many VPCs and accounts. Which service is designed for this?
 - A AWS Firewall Manager
 - B AWS Shield Advanced alone
 - C Amazon Inspector
 - D Amazon Cognito

Answer: A

Firewall Manager centrally applies WAF rules, security groups and Network Firewall policies across accounts that belong to the organization. Shield Advanced protects against DDoS attacks but does not manage rule sets across accounts. Inspector scans workloads for vulnerabilities, and Cognito handles application user sign-in.

4. AWS Transit Gateway uses _____ tables to decide which attachments can route traffic to each other.

Answer: route

Each attachment is associated with a route table, and propagation or static routes in that table determine reachability. Separate route tables are how hub designs isolate environments. The answer names the table type that controls this routing.

5. Which TWO are valid ways to monitor the quality of a generative AI application over time? (Choose two.)

- A Re-running Bedrock model evaluation jobs against a fixed test dataset after each change
- B Counting tokens only, and ignoring the content of responses
- C Sampling production responses for human review through a labeling workflow
- D Disabling invocation logs to reduce storage cost
- E Deleting failed responses before they are logged

Answer: A, C

A fixed evaluation dataset gives a repeatable quality baseline that can be compared after each prompt or model change, and sampled human review catches issues automated metrics miss. Counting tokens says nothing about quality. Disabling or editing logs removes the evidence needed to detect regressions.

6. A company must keep its model training data inside its VPC, with no path to the internet. Which SageMaker configuration supports this?
- A Training jobs in a public subnet with an internet gateway, downloading packages from public repositories
 - B Training in VPC mode with private subnets, VPC endpoints and network isolation enabled
 - C Training in the default SageMaker-managed network with internet access enabled
 - D Training with network isolation disabled and a NAT gateway for package downloads

Answer: B

VPC mode places training containers in private subnets, VPC endpoints reach the services they need, and network isolation blocks outbound traffic from the container. A public subnet with an internet gateway puts the job on an internet path. The default network with internet access and a NAT gateway path both leave a route to the internet.

7. A serverless pipeline must transform each PDF upload, extract its text, write the chunks to a vector index and retry failed items. Which design is best?
- A An EC2 cron job that polls S3 every minute
 - B An Amazon Data Firehose stream that delivers the PDFs straight into the vector index
 - C An S3 event notification that triggers a Lambda function or Step Functions workflow with a dead-letter queue for failed events
 - D A scheduled AWS Glue crawler only

Answer: C

S3 event notifications trigger processing as each file arrives, and a dead-letter queue captures failures for retry or investigation. Polling wastes compute and adds latency, and a Firehose stream does not extract text from PDFs or split them into chunks. A Glue crawler catalogs schemas and does not extract text or write vectors.

8. Amazon Bedrock Knowledge Bases requires you to fine-tune the foundation model on your documents before it can answer questions from them.
- A True
 - B False

Answer: B (False)

Knowledge Bases answers from retrieved document passages supplied in the prompt, so the foundation model does not need fine-tuning. Fine-tuning is a separate customization option with its own cost and purpose. The statement is false.

9. A nightly job must summarize two million documents with a Bedrock model, and the results are needed by morning. Which option minimizes cost?
- A Batch inference jobs that read from and write to S3
 - B Real-time on-demand calls from 200 parallel workers
 - C Provisioned Throughput purchased for a full year
 - D A fine-tuned model on a dedicated GPU instance that runs all day

Answer: A

Batch inference processes large offline volumes from S3 at a lower price than real-time calls, and the morning deadline suits an asynchronous job. Real-time calls cost more per request at this volume. Long commitments and always-on GPUs pay for capacity the job uses only overnight.

10. A production Lambda function is throttled during traffic peaks. Which action addresses this most directly?
- A Raise the function memory to the maximum size only, which speeds up each individual execution
 - B Remove the reserved concurrency setting
 - C Watch the Throttles metric and set reserved or provisioned concurrency to match demand
 - D Increase the function's ephemeral storage size

Answer: C

Throttles shows when invocations are rejected for lack of concurrency, and reserved or provisioned concurrency sets the capacity the function can use. Extra memory speeds execution but does not add concurrency. Removing reserved concurrency makes the problem harder to manage, and more ephemeral storage does not add concurrency.