Amazon-Web-Services

Exam Questions AIP-C01

AWS Certified Generative AI Developer - Professional

NEW QUESTION 1
A company is developing a generative AI (GenAI) application that uses Amazon Bedrock foundation models. The application has several custom tool integrations. The application has experienced unexpected token consumption surges despite consistent user traffic.
The company needs a solution that uses Amazon Bedrock model invocation logging to monitor InputTokenCount and OutputTokenCount metrics. The solution must detect unusual patterns in tool usage and identify which specific tool integrations cause abnormal token consumption. The solution must also automatically adjust thresholds as traffic patterns change.
Which solution will meet these requirements?

A. Use Amazon CloudWatch Logs to capture model invocation log
B. Create CloudWatch dashboards for token metric
C. Configure static CloudWatch alarms with fixed thresholds for each tool integration.
D. Store model invocation logs in Amazon S3. Use AWS Glue and Amazon Athena to analyze token usage trends.
E. Use Amazon CloudWatch Logs to capture model invocation log
F. Create CloudWatch metric filters to extract tool-specific invocation pattern
G. Apply CloudWatch anomaly detection alarms that automatically adjust baselines for each tool??s token metrics.
H. Store model invocation logs in an Amazon S3 bucke
I. Use AWS Lambda to process logs in real tim
J. Manually update CloudWatch alarm thresholds based on trends identified by the Lambda function.

Answer: C

NEW QUESTION 2
An ecommerce company is developing a generative AI application that uses Amazon Bedrock with Anthropic Claude to recommend products to customers. Customers report that some recommended products are not available for sale on the website or are not relevant to the customer. Customers also report that the solution takes a long time to generate some recommendations.
The company investigates the issues and finds that most interactions between customers and the product recommendation solution are unique. The company confirms that the solution recommends products that are not in the company??s product catalog. The company must resolve these issues.
Which solution will meet this requirement?

A. Increase grounding within Amazon Bedrock Guardrail
B. Enable Automated Reasoningcheck
C. Set up provisioned throughput.
D. Use prompt engineering to restrict the model responses to relevant product
E. Use streaming techniques such as the InvokeModelWithResponseStream action to reduce perceived latency for the customers.
F. Create an Amazon Bedrock knowledge bas
G. Implement Retrieval Augmented Generation RA
H. Set the PerformanceConfigLatency parameter to optimized.
I. Store product catalog data in Amazon OpenSearch Servic
J. Validate the model??s product recommendations against the product catalo
K. Use Amazon DynamoDB to implement response caching.

Answer: C

NEW QUESTION 3
A company has a generative AI (GenAI) application that uses Amazon Bedrock to provide real-time responses to customer queries. The company has noticed intermittent failures with API calls to foundation models (FMs) during peak traffic periods.
The company needs a solution to handle transient errors and provide detailed observability into FM performance. The solution must prevent cascading failures during throttling events and provide distributed tracing across service boundaries to identify latency contributors. The solution must also enable correlation of performance issues with specific FM characteristics.
Which solution will meet these requirements?

A. Implement a custom retry mechanism with a fixed delay of 1 second between retrie
B. Configure Amazon CloudWatch alarms to monitor the application??s error rates and latency metrics.
C. Configure the AWS SDK with standard retry mode and exponential backoff with jitte
D. Use AWS X-Ray tracing with annotations to identify and filter service components.
E. Implement client-side caching of all FM response
F. Add custom logging statements in the application code to record API call durations.
G. Configure the AWS SDK with adaptive retry mod
H. Use AWS CloudTrail distributed tracing to monitor throttling events.

Answer: B

NEW QUESTION 4
A retail company is using Amazon Bedrock to develop a customer service AI assistant. Analysis shows that 70% of customer inquiries are simple product questions that a smaller model can effectively handle. However, 30% of inquiries are complex return policy questions that require advanced reasoning.
The company wants to implement a cost-effective model selection framework to automatically route customer inquiries to appropriate models based on inquiry complexity. The framework must maintain high customer satisfaction and minimize response latency.
Which solution will meet these requirements with the LEAST implementation effort?

A. Create a multi-stage architecture that uses a small foundation model (FM) to classify the complexity of each inquir
B. Route simple inquiries to a smaller, more cost-effective mode
C. Route complex inquiries to a larger, more capable mode
D. Use AWS Lambda functions to handle routing logic.
E. Use Amazon Bedrock intelligent prompt routing to automatically analyze inquirie
F. Route simple product inquiries to smaller models and route complex return policy inquiries to more capable larger models.
G. Implement a single-model solution that uses an Amazon Bedrock mid-sized foundation model (FM) with on-demand pricin
H. Include special instructions in model prompts to handle both simple and complex inquiries by using the same model.
I. Create separate Amazon Bedrock endpoints for simple and complex inquirie

J. Implement a rule-based routing system based on keyword detectio
K. Use on-demand pricing for the smaller model and provisioned throughput for the larger model.

Answer: B

NEW QUESTION 5
A company is building a legal research AI assistant that uses Amazon Bedrock with an Anthropic Claude foundation model (FM). The AI assistant must retrieve highly relevant case law documents to augment the FM??s responses. The AI assistant must identify semantic relationships between legal concepts, specific legal terminology, and citations. The AI assistant must perform quickly and return precise results.
Which solution will meet these requirements?

A. Configure an Amazon Bedrock knowledge base to use a default vector search configuratio
B. Use Amazon Bedrock to expand queries to improve retrieval for legal documents based on specific terminology and citations.
C. Use Amazon OpenSearch Service to deploy a hybrid search architecture that combines vector search with keyword searc
D. Apply an Amazon Bedrock reranker model to optimize result relevance.
E. Enable the Amazon Kendra query suggestion feature for end user
F. Use Amazon Bedrock to perform post-processing of search results to identify semantic similarity in the documents and to produce precise results.
G. Use Amazon OpenSearch Service with vector search and Amazon Bedrock Titan Embeddings to index and search legal document
H. Use custom AWS Lambda functions to merge results with keyword-based filters that are stored in an Amazon RDS database.

Answer: B

NEW QUESTION 6
A company needs a system to automatically generate study materials from multiple content sources. The content sources include document files (PDF files, PowerPoint presentations, and Word documents) and multimedia files (recorded videos). The system must process more than 10,000 content sources daily with peak loads of 500 concurrent uploads. The system must also extract key concepts from document files and multimedia files and create contextually accurate summaries. The generated study materials must support real-time collaboration with version control.
Which solution will meet these requirements?

A. Use Amazon Bedrock Data Automation (BDA) with AWS Lambda functions to orchestrate document file processin
B. Use Amazon Bedrock Knowledge Bases to process all multimedi
C. Store the content in Amazon DocumentDB with replicatio
D. Collaborate by using Amazon SNS topic subscription
E. Track changes by using Amazon Bedrock Agents.
F. Use Amazon Bedrock Data Automation (BDA) with foundation models (FMs) to process document file
G. Integrate BDA with Amazon Textract for PDF extraction and with Amazon Transcribe for multimedia file
H. Store the processed content in Amazon S3 with versioning enable
I. Store the metadata in Amazon DynamoD
J. Collaborate in real time by using AWS AppSync GraphQL subscriptions and DynamoDB.
K. Use Amazon Bedrock Data Automation (BDA) with Amazon SageMaker AI endpoints to host content extraction and summarization model
L. Use Amazon Bedrock Guardrails to extract content from all file type
M. Store document files in Amazon Neptune for time series analysi
N. Collaborate by using Amazon Bedrock Chat for real-time messaging.
O. Use Amazon Bedrock Data Automation (BDA) with AWS Lambda functions to process batches of content file
P. Fine-tune foundation models (FMs) in Amazon Bedrock to classify documents across all content type
Q. Store the processed data in Amazon ElastiCache (Redis OSS) by using Cluster Mode with shardin
R. Use Prompt management in Amazon Bedrock for version control.

Answer: B

NEW QUESTION 7
A company is designing an API for a generative AI (GenAI) application that uses a foundation model (FM) that is hosted on a managed model service. The API must stream responses to reduce latency, enforce token limits to manage compute resource usage, and implement retry logic to handle model timeouts and partial responses.
Which solution will meet these requirements with the LEAST operational overhead?

A. Integrate an Amazon API Gateway HTTP API with an AWS Lambda function to invoke Amazon Bedroc
B. Use Lambda response streaming to stream response
C. Enforce token limits within the Lambda functio
D. Implement retry logic for model timeouts by using Lambda and API Gateway timeout configurations.
E. Connect an Amazon API Gateway HTTP API directly to Amazon Bedroc
F. Simulate streaming by using client-side pollin
G. Enforce token limits on the fronten
H. Configure retry behavior by using API Gateway integration settings.
I. Connect an Amazon API Gateway WebSocket API to an Amazon ECS service that hosts a containerized inference serve
J. Stream responses by using the WebSocket protoco
K. Enforce token limits within Amazon EC
L. Handle model timeouts by using ECS task lifecycle hooks and restart policies.
M. Integrate an Amazon API Gateway REST API with an AWS Lambda function that invokes Amazon Bedroc
N. Use Lambda response streaming to stream response
O. Enforce token limits within the Lambda functio
P. Implement retry logic by using Lambda and API Gateway timeout configurations.

Answer: A

NEW QUESTION 8
A publishing company is developing a chat assistant that uses a containerized large language model (LLM) that runs on Amazon SageMaker AI. The architecture consists of an Amazon API Gateway REST API that routes user requests to an AWS Lambda function. The Lambda function invokes a SageMaker AI real-time endpoint that hosts the LLM.

Users report uneven response times. Analytics show that a high number of chats are abandoned after 2 seconds of waiting for the first token. The company wants a solution to ensure that p95 latency is under 800 ms for interactive requests to the chat assistant.
Which combination of solutions will meet this requirement? (Select TWO.)

A. Enable model preload upon container startu
B. Implement dynamic batching to process multiple user requests together in a single inference pass.
C. Select a larger GPU instance type for the SageMaker AI endpoin
D. Set the minimum number of instances to 0. Continue to perform per-request processin
E. Lazily load model weights on the first request.
F. Switch to a multi-model endpoin
G. Use lazy loading without request batching.
H. Set the minimum number of instances to greater than 0. Enable response streaming.
I. Switch to Amazon SageMaker Asynchronous Inference for all request
J. Store requests in an Amazon S3 bucke
K. Set the minimum number of instances to 0.

Answer: AD

NEW QUESTION 9
A company is using Amazon Bedrock to design an application to help researchers apply for grants. The application is based on an Amazon Nova Pro foundation model (FM). The application contains four required inputs and must provide responses in a consistent text format. The company wants to receive a notification in Amazon Bedrock if a response contains bullying language. However, the company does not want to block all flagged responses.
The company creates an Amazon Bedrock flow that takes an input prompt and sends it to the Amazon Nova Pro FM. The Amazon Nova Pro FM provides a response.
Which additional steps must the company take to meet these requirements? (Select TWO.)

A. Use Amazon Bedrock Prompt Management to specify the required inputs as variable
B. Select an Amazon Nova Pro F
C. Specify the output format for the respons
D. Add the prompt to the prompts node of the flow.
E. Create an Amazon Bedrock guardrail that applies the hate content filte
F. Set the filter response to bloc
G. Add the guardrail to the prompts node of the flow.
H. Create an Amazon Bedrock prompt route
I. Specify an Amazon Nova Pro F
J. Add the required inputs as variables to the input node of the flo
K. Add the prompt router to the prompts nod
L. Add the output format to the output node.
M. Create an Amazon Bedrock guardrail that applies the insults content filte
N. Set the filter response to detec
O. Add the guardrail to the prompts node of the flow.
P. Create an Amazon Bedrock application inference profile that specifies an Amazon Nova Pro F
Q. Specify the output format for the response in the descriptio
R. Include a tag for each of the input variable
S. Add the profile to the prompts node of the flow.

Answer: AD

NEW QUESTION 10
A medical company is creating a generative AI (GenAI) system by using Amazon Bedrock. The system processes data from various sources and must maintain end-to-end data lineage. The system must also use real-time personally identifiable information (PII) filtering and audit trails to automatically report compliance.
Which solution will meet these requirements?

A. Use AWS Glue Data Catalog to register all data sources and track lineag
B. Use Amazon Bedrock Guardrails PII filter
C. Enable AWS CloudTrail logging for all Amazon Bedrock API calls with Amazon S3 integratio
D. Use Amazon Macie to scan stored data for sensitive information and publish findings to Amazon CloudWatch Log
E. Create CloudWatch dashboards to visualize the findings and generate automated compliance reports.
F. Use AWS Config to track data source configurations and change
G. Use AWS WAF with custom rules to filter PII at the application layer before Amazon Bedrock processes the dat
H. Configure Amazon EventBridge to capture and route audit events to Amazon S3. Use Amazon Comprehend Medical with scheduled AWS Lambda functions to analyze stored outputs for compliance violations.
I. Use AWS DataSync to replicate data sources to track lineag
J. Configure Amazon Macie to scan Amazon Bedrock outputs for sensitive informatio
K. Use AWS Systems Manager Session Manager to log user interaction
L. Deploy Amazon Textract with AWS Step Functions workflows to identify and redact PII from generated reports.
M. Configure Amazon Athena to query data sources to analyze and report on data lineag
N. Use Amazon CloudWatch custom metrics to monitor PII exposure in Amazon Bedrock responses and establish AWS X-Ray tracing to generate an audit trai
O. Use an Amazon Rekognition Custom Labels model to detect sensitive information in the data that Amazon Bedrock processes.

Answer: A

NEW QUESTION 10
A healthcare company uses Amazon Bedrock to deploy an application that generates summaries of clinical documents. The application experiences inconsistent response quality with occasional factual hallucinations. Monthly costs exceed the company??s projections by 40%. A GenAI developer must implement a near real-time monitoring solution to detect hallucinations, identify abnormal token consumption, and provide early warnings of cost anomalies. The solution must require minimal custom development work and maintenance overhead.
Which solution will meet these requirements?

A. Configure Amazon CloudWatch alarms to monitor InputTokenCount and OutputTokenCount metrics to detect anomalie

B. Store model invocation logs in an Amazon S3 bucke
C. Use AWS Glue and Amazon Athena to identify potential hallucinations.
D. Run Amazon Bedrock evaluation jobs that use LLM-based judgments to detect hallucination
E. Configure Amazon CloudWatch to track token usag
F. Create an AWS Lambda function to process CloudWatch metric
G. Configure the Lambda function to send usage pattern notifications.
H. Configure Amazon Bedrock to store model invocation logs in an Amazon S3 bucke
I. Enable text output loggin
J. Configure Amazon Bedrock guardrails to run contextual grounding checks to detect hallucination
K. Create Amazon CloudWatch anomaly detection alarms for token usage metrics.
L. Use AWS CloudTrail to log all Amazon Bedrock API call
M. Create a custom dashboard in Amazon QuickSight to visualize token usage pattern
N. Use Amazon SageMaker Model Monitor to detect quality drift in generated summaries.

Answer: C

NEW QUESTION 15
A university recently digitized a collection of archival documents, academic journals, and manuscripts. The university stores the digital files in an AWS Lake Formation data lake.
The university hires a GenAI developer to build a solution to allow users to search the digital files by using text queries. The solution must return journal abstracts that are semantically similar to a user's query. Users must be able to search the digitized collection based on text and metadata that is associated with the journal abstracts. The metadata of the digitized files does not contain keywords. The solution must match similar abstracts to one another based on the similarity of their text. The data lake contains fewer than 1 million files.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use Amazon Titan Embeddings in Amazon Bedrock to create vector representations of the digitized file
B. Store embeddings in the OpenSearch Neural plugin for Amazon OpenSearch Service.
C. Use Amazon Comprehend to extract topics from the digitized file
D. Store the topics and file metadata in an Amazon Aurora PostgreSQL databas
E. Query the abstract metadata against the data in the Aurora database.
F. Use Amazon SageMaker AI to deploy a sentence-transformer mode
G. Use the model to create vector representations of the digitized file
H. Store embeddings in an Amazon Aurora PostgreSQL database that has the pgvector extension.
I. Use Amazon Titan Embeddings in Amazon Bedrock to create vector representations of the digitized file
J. Store embeddings in an Amazon Aurora PostgreSQL Serverless database that has the pgvector extension.

Answer: D

NEW QUESTION 20
A retail company has a generative AI (GenAI) product recommendation application that uses Amazon Bedrock. The application suggests products to customers based on browsing history and demographics. The company needs to implement fairness evaluation across multiple demographic groups to detect and measure bias in recommendations between two prompt approaches. The company wants to collect and monitor fairness metrics in real time. The company must receive an alert if the fairness metrics show a discrepancy of more than 15% between demographic groups. The company must receive weekly reports that compare the performance of the two prompt approaches.
Which solution will meet these requirements with the LEAST custom development effort?

A. Configure an Amazon CloudWatch dashboard to display default metrics from Amazon Bedrock API call
B. Create custom metrics based on model output
C. Set up Amazon EventBridge rules to invoke AWS Lambda functions that perform post-processing analysis on model responses and publish custom fairness metrics.
D. Create the two prompt variants in Amazon Bedrock Prompt Managemen
E. Use Amazon Bedrock Flows to deploy the prompt variants with defined traffic allocatio
F. Configure Amazon Bedrock guardrails to monitor demographic fairnes
G. Set up Amazon CloudWatch alarms on the GuardrailContentSource dimension by using InvocationsIntervened metrics to detect recommendation discrepancy threshold violations.
H. Set up Amazon SageMaker Clarify to analyze model output
I. Publish fairness metrics to Amazon CloudWatc
J. Create CloudWatch composite alarms that combine SageMaker Clarify bias metrics with Amazon Bedrock latency metrics.
K. Create an Amazon Bedrock model evaluation job to compare fairness between the two prompt variant
L. Enable model invocation logging in Amazon CloudWatc
M. Set up CloudWatch alarms for InvocationsIntervened metrics with a dimension for each demographic group.

Answer: B

NEW QUESTION 25
An ecommerce company operates a global product recommendation system that needs to switch between multiple foundation models (FM) in Amazon Bedrock based on regulations,
cost optimization, and performance requirements. The company must apply custom controls based on proprietary business logic, including dynamic cost thresholds, AWS Region-specific compliance rules, and real-time A/B testing across multiple FMs.
The system must be able to switch between FMs without deploying new code. The system must route user requests based on complex rules including user tier, transaction value, regulatory zone, and real-time cost metrics that change hourly and require immediate propagation across thousands of concurrent requests.
Which solution will meet these requirements?

A. Deploy an AWS Lambda function that uses environment variables to store routing rules and Amazon Bedrock FM ID
B. Use the Lambda console to update the environment variables when business requirements chang
C. Configure an Amazon API Gateway REST API to read request parameters to make routing decisions.
D. Deploy Amazon API Gateway REST API request transformation templates to implement routing logic based on request attribute
E. Store Amazon Bedrock FM endpoints as REST API stage variable
F. Update the variables when the system switches between models.
G. Configure an AWS Lambda function to fetch routing configurations from the AWS AppConfig Agent for each user reques

H. Run business logic in the Lambda function to select the appropriate FM for each reques
I. Expose the FM through a single Amazon API Gateway REST API endpoint.
J. Use AWS Lambda authorizers for an Amazon API Gateway REST API to evaluate routing rules that are stored in AWS AppConfi
K. Return authorization contexts based on business logi
L. Route requests to model-specific Lambda functions for each Amazon Bedrock FM.

Answer: C

NEW QUESTION 26
A wildlife conservation agency operates zoos globally. The agency uses various sensors, trackers, and audiovisual recorders to monitor animal behavior. The agency wants to launch a generative AI (GenAI) assistant that can ingest multimodal data to study animal behavior.
The GenAI assistant must support natural language queries, avoid speculative behavioral interpretations, and maintain audit logs for ethical research audits.
Which solution will meet these requirements?

A. Ingest raw videos into Amazon Rekognition to detect animal postures and expression
B. Use Amazon Data Firehose to stream sensor and GPS data into Amazon S3. Prompt an Amazon Bedrock FM using basic templates stored in AWS Systems Manager Parameter Stor
C. Use IAM for access contro
D. Use AWS CloudTrail for audit logging.
E. Use Amazon SageMaker Processing and Amazon Transcribe to pre-process multimodal dat
F. Ingest curated summaries into an Amazon Bedrock Knowledge Base
G. Apply Amazon Bedrock guardrails to restrict speculative output
H. Use AWS AppConfig to manage prompt template
I. Use AWS CloudTrail to log research activity for audits.
J. Use Amazon OpenSearch Serverless to index behavioral logs and telemetr
K. Use Amazon Comprehend to extract entitie
L. Use Amazon Bedrock to answer questions over indexed dat
M. Use IAM for access control and CloudTrail for audit logging.
N. Configure Amazon O Business to federate data across Amazon S3, Amazon Kinesis, and Amazon SageMaker Feature Stor
O. Use EventBridge for ingestion orchestratio
P. Use custom AWS Lambda functions to filter LLM outputs for ethical compliance.

Answer: B

NEW QUESTION 28
A financial technology company is using Amazon Bedrock to build an assessment system for the company??s customer service AI assistant. The AI assistant must provide financial recommendations that are factually accurate, compliant with financial regulations, and conversationally appropriate. The company needs to combine automated quality evaluations at scale with targeted human reviews of critical interactions.
What solution will meet these requirements?

A. Configure a pipeline in which financial experts manually score all responses for accuracy, compliance, and conversational qualit
B. Use Amazon SageMaker notebooks to analyze results to identify improvement areas.
C. Configure Amazon Bedrock evaluations that use Anthropic Claude Sonnet as a judge model to assess response accuracy and appropriatenes
D. Configure custom Amazon Bedrock guardrails to check responses for compliance with financial policie
E. Add Amazon Augmented AI (Amazon A2I) human reviews for flagged critical interactions.
F. Create an Amazon Lex bot to manage customer service interaction
G. Configure AWSLambda functions to check responses against a static compliance databas
H. Configure intents that call the Lambda function
I. Add an additional intent to collect end-user reviews.
J. Configure Amazon CloudWatch to monitor response patterns from the AI assistan
K. Configure CloudWatch alerts for potential compliance violation
L. Establish a team of human evaluators to review flagged interactions.

Answer: B

NEW QUESTION 31
A financial services company is developing a generative AI (GenAI) application that serves both premium customers and standard customers. The application uses AWS Lambda functions behind an Amazon API Gateway REST API to process requests. The company needs to dynamically switch between AI models based on which customer tier each user belongs to. The company also wants to perform A/B testing for new features without redeploying code. The company needs to validate model parameters like temperature and maximum token limits before applying changes.
Which solution will meet these requirements with the LEAST operational overhead?

A. Create AWS Systems Manager Parameter Store parameters for each configuratio
B. Use Lambda functions to poll for parameter update
C. Use Amazon EventBridge events to trigger redeployments when configurations change.
D. Store model configurations in Amazon DynamoDB table
E. Optimize access patterns to retrieve configurations according to customer tie
F. Configure Lambda functions to query DynamoDB at the beginning of each request to determine which model to use.
G. Use AWS AppConfig to manage model configuration
H. Use feature flags to perform A/B testin
I. Define JSON schema validation rules for model parameter
J. Configure Lambda functions to retrieve configurations by using the AWS AppConfig Agent.
K. Create an Amazon ElastiCache (Redis OSS) cluster to store model configuration
L. Set short TTL value
M. Run custom validation logic in Lambda function
N. Use Amazon CloudWatch metrics to monitor configuration usage.

Answer: C

NEW QUESTION 35
A company is building a video analysis platform on AWS. The platform will analyze a large video archive by using Amazon Rekognition and Amazon Bedrock. The platform must comply with predefined privacy standards. The platform must also use secure model I/O, control foundation model (FM) access patterns, and provide an audit of who accessed what and when.
Which solution will meet these requirements?

A. Configure VPC endpoints for Amazon Bedrock model API call
B. Implement Amazon Bedrock guardrails to filter harmful or unauthorized content in prompts and response
C. Use Amazon Bedrock trace events to track all agent and model invocations for auditing purpose
D. Export the traces to Amazon CloudWatch Logs as an audit record of model usag
E. Store all prompts and outputs in Amazon S3 with server-side encryption with AWS KMS keys (SSE-KMS).
F. Define access control by using IAM with attribute-based access control (ABAC) to map departments to specific permission
G. Configure VPC endpoints for Amazon Bedrock model API call
H. Use IAM condition keys to enforce specific GuardrailIdentifier and ModelId value
I. Configure AWS CloudTrail to capture management and data events for S3 objects and KMS key usage activitie
J. Enable S3 server access logging to record detailed file-level interactions with the video archive
K. Send all CloudTrail logs to AWS CloudTrail Lak
L. Set up Amazon CloudWatch alarms to detect and alert on unexpected activity from Amazon Bedrock, Amazon Rekognition, and AWS KMS.
M. Restrict access to services by using VPC endpoint policie
N. Use AWS Config to track resource changes and compliance with security rule
O. Use server-side encryption with AWS KMS keys (SSE-KMS) to encrypt data at res
P. Store the model??s I/O in separate Amazon S3 bucket
Q. Enable S3 server access logging to track file-level interactions.
R. Configure AWS CloudTrail Insights to analyze API call patterns across accounts and detect anomalous activity in Amazon Bedrock, Amazon Rekognition, Amazon S3, and AWS KM
S. Deploy Amazon Macie to scan and classify the video archiv
T. Use server-side encryption with AWS KMS keys (SSE-KMS) to encrypt all stored dat
. Configure CloudTrail to capture KMS API usage events for audit purpose
. Configure Amazon EventBridge rules to process CloudTrail Insights anomalies and Macie finding
. Use CloudWatch alarms to trigger automated notifications and security responses when potential security issues are detected.

Answer: B

NEW QUESTION 39
A financial services company uses an AI application to process financial documents by using Amazon Bedrock. During business hours, the application handles approximately 10,000 requests each hour, which requires consistent throughput.
The company uses the CreateProvisionedModelThroughput API to purchase provisioned throughput. Amazon CloudWatch metrics show that the provisioned capacity is unused while on-demand requests are being throttled. The company finds the following code in the application:
response = bedrock_runtime.invoke_model(modelId="anthropic.claude-v2", body=json.dumps(payload)
)
The company needs the application to use the provisioned throughput and to resolve the throttling issues.
Which solution will meet these requirements?

A. Increase the number of model units (MUs) in the provisioned throughput configuration.
B. Replace the model ID parameter with the ARN of the provisioned model that the CreateProvisionedModelThroughput API returns.
C. Add exponential backoff retry logic to handle throttling exceptions during peak hours.
D. Modify the application to use the invokeModelWithResponseStream API instead of the invokeModel API.

Answer: B

NEW QUESTION 42
A company developed a multimodal content analysis application by using Amazon Bedrock. The application routes different content types (text, images, and code) to specialized foundation models (FMs).
The application needs to handle multiple types of routing decisions. Simple routing based on file extension must have minimal latency. Complex routing based on content semantics requires analysis before FM selection. The application must provide detailed history and support fallback options when primary FMs fail.
Which solution will meet these requirements?

A. Configure AWS Lambda functions that call Amazon Bedrock FMs for all routing logi
B. Use conditional statements to determine the appropriate FM based on content type and semantics.
C. Create a hybrid solutio
D. Handle simple routing based on file extensions in application cod
E. Handle complex content-based routing by using an AWS Step Functions state machine with JSONata for content analysis and the InvokeModel API for specialized FMs.
F. Deploy separate AWS Step Functions workflows for each content type with routing logic in AWS Lambda function
G. Use Amazon EventBridge to coordinate between workflows when fallback to alternate FMs is required.
H. Use Amazon SQS with different SQS queues for each content typ
I. Configure AWS Lambda consumers that analyze content and invoke appropriate FMs based on message attributes by using Amazon Bedrock with an AWS SDK.

Answer: B

NEW QUESTION 47
A company is planning to deploy multiple generative AI (GenAI) applications to five independent business units that operate in multiple countries in Europe and the Americas.
Each application uses Amazon Bedrock Retrieval Augmented Generation (RAG) patterns with business unit-specific knowledge bases that store terabytes of unstructured data.
The company must establish well-architected, standardized components for security controls, observability practices, and deployment patterns across all the GenAI applications. The components must be reusable, versioned, and governed consistently.
Which solution will meet these requirements?

A. Configure Amazon API Gateway REST API endpoints for the GenAI application
B. Deploy common security, observability, and RAG patterns based on the AWS Well- Architected Generative AI Lens in standardized AWS CloudFormation template
C. Use CloudFormation Guard after deployment to validate policy compliance in each business unit.
D. Create standardized AWS CloudFormation templates to implement security, observability, and RAG patterns based on the AWS Well-Architected Generative AI Len
E. Establish a centralized repository for version contro
F. Integrate a CI/CD pipeline with CloudFormation Guard to enforce consistent and repeatable deployments across business units.
G. Use AWS Service Catalog to define standardized portfolios and versioned products for each business uni
H. Use the portfolios to enforce security, observability, and RAG patterns based on the AWS Well-Architected Generative AI Len
I. Require business units to use the Service Catalog console to deploy resources.
J. Document security controls, observability requirements, and RAG patterns based on the AWS Well-Architected Generative AI Lens in a shared design documen
K. Use Amazon Macie to enforce deploymen
L. Delegate implementation responsibility to each business unit.

Answer: B

NEW QUESTION 52
A finance company is developing an AI assistant to help clients plan investments and manage their portfolios. The company identifies several high-risk conversation patterns such as requests for specific stock recommendations or guaranteed returns. High-risk conversation patterns could lead to regulatory violations if the company cannot implement appropriate controls.
The company must ensure that the AI assistant does not provide inappropriate financial advice, generate content about competitors, or make claims that are not factually grounded in the company's approved financial guidance. The company wants to use Amazon Bedrock Guardrails to implement a solution.
Which combination of steps will meet these requirements? (Select THREE)

A. Add the high-risk conversation patterns to a denied topics guardrail.
B. Configure a content filter guardrail to filter prompts that contain the high-risk conversation patterns.
C. Configure a content filter guardrail to filter prompts that contain competitor names.
D. Add the names of competitors as custom word filter
E. Set the input and output actions to block.
F. Set a low grounding score threshold.
G. Set a high grounding score threshold.

Answer: ADF

NEW QUESTION 54
A company wants to select a new FM for its AI assistant. A GenAI developer needs to generate evaluation reports to help a data scientist assess the quality and safety of various foundation models FMs. The data scientist provides the GenAI developer with sample prompts for evaluation. The GenAI developer wants to use Amazon Bedrock to automate report generation and evaluation.
Which solution will meet this requirement?

A. Combine the sample prompts into a single JSON documen
B. Create an Amazon Bedrock knowledge base with the documen
C. Write a prompt that asks the FM to generate a response to each sample promp
D. Use the RetrieveAndGenerate API to generate a report for each model.
E. Combine the sample prompts into a single JSONL documen
F. Store the document in an Amazon S3 bucke
G. Create an Amazon Bedrock evaluation job that uses a judge mode
H. Specify the S3 location as input and a different S3 location as outpu
I. Run an evaluation job for each FM and select the FM as the generator.
J. Combine the sample prompts into a single JSONL documen
K. Store the document in an Amazon S3 bucke
L. Create an Amazon Bedrock evaluation job that uses a judge mode
M. Specify the S3 location as input and Amazon QuickSight as outpu
N. Run an evaluation job for each FM and select the FM as the evaluator.
O. Combine the sample prompts into a single JSON documen
P. Create an Amazon Bedrock knowledge base from the documen
Q. Create an Amazon Bedrock evaluation job that uses the retrieval and response generation evaluation typ
R. Specify an Amazon S3 bucket as the outpu
S. Run an evaluation job for each FM.

Answer: B

NEW QUESTION 59
A healthcare company is developing an application to process medical queries. The application must answer complex queries with high accuracy by reducing semantic dilution. The application must refer to domain-specific terminology in medical documents to reduce ambiguity in medical terminology. The application must be able to respond to 1,000 queries each minute with response times less than 2 seconds.
Which solution will meet these requirements with the LEAST operational overhead?

A. Use Amazon API Gateway to route incoming queries to an Amazon Bedrock agen
B. Configure the agent to use an Anthropic Claude model to decompose queries and an Amazon Titan model to expand querie
C. Create an Amazon Bedrock knowledge base to store the reference medical documents.
D. Configure an Amazon Bedrock knowledge base to store the reference medical document
E. Enable query decomposition in the knowledge bas
F. Configure an Amazon Bedrock flow that uses a foundation model and the knowledge base to support the application.
G. Use Amazon SageMaker AI to host custom ML models for both query decomposition and query expansio
H. Configure Amazon Bedrock knowledge bases to store the reference medical document
I. Encrypt the documents in the knowledge base.
J. Create an Amazon Bedrock agent to orchestrate multiple AWS Lambda functions to decompose querie
K. Create an Amazon Bedrock knowledge base to store the reference medical document

L. Use the agent??s built-in knowledge base capabilitie
M. Add deep research and reasoning capabilities to the agent to reduce ambiguity in the medical terminology.

Answer: B

NEW QUESTION 63
A medical device company wants to feed reports of medical procedures that used the company??s devices into an AI assistant. To protect patient privacy, the AI assistant must expose patient personally identifiable information (PII) only to surgeons. The AI assistant must redact PII for engineers. The AI assistant must reference only medical reports that are less than 3 years old.
The company stores reports in an Amazon S3 bucket as soon as each report is published. The company has already set up an Amazon Bedrock Knowledge Bases. The AI assistant uses Amazon Cognito to authenticate users.
Which solution will meet these requirements?

A. Enable Amazon Macie PII detection on the S3 bucke
B. Use an S3 trigger to invoke an AWS Lambda function that redacts PII from the report
C. Configure the Lambda function to delete outdated documents and invoke knowledge base syncing.
D. Invoke an AWS Lambda function to sync the S3 bucket and the knowledge base when a new report is uploade
E. Use a second Lambda function with Amazon Comprehend to redact PII for engineer
F. Use S3 Lifecycle rules to remove reports older than 3 years.
G. Set up an S3 Lifecycle configuration to remove reports that are older than 3 year
H. Schedule an AWS Lambda function to run daily syncs between the bucket and the knowledge bas
I. When users interact with the AI assistant, apply a guardrail configuration selected based on the user??s Cognito user group to redact PII from responses when required.
J. Create a second knowledge bas
K. Use Lambda and Amazon Comprehend to redact PII before syncing to the second knowledge bas
L. Route users to the appropriate knowledge base based on Cognito group membership.

Answer: C

NEW QUESTION 67
A medical company is building a generative AI (GenAI) application that uses Retrieval Augmented Generation (RAG) to provide evidence-based medical information. The application uses Amazon OpenSearch Service to retrieve vector embeddings. Users report that searches frequently miss results that contain exact medical terms and acronyms and return too many semantically similar but irrelevant documents. The company needs to improve retrieval quality and maintain low end-user latency, even as the document collection grows to millions of documents.
Which solution will meet these requirements with the LEAST operational overhead?

A. Configure hybrid search by combining vector similarity with keyword matching to improve semantic understanding and exact term and acronym matching.
B. Increase the dimensions of the vector embeddings from 384 to 1536. Use a post- processing AWS Lambda function to filter out irrelevant results after retrieval.
C. Replace OpenSearch Service with Amazon Kendr
D. Use query expansion to handle medical acronyms and terminology variants during pre-processing.
E. Implement a two-stage retrieval architecture in which initial vector search results are re- ranked by an ML model hosted on Amazon SageMaker.

Answer: A

NEW QUESTION 69
A financial services company is developing a real-time generative AI (GenAI) assistant to support human call center agents. The GenAI assistant must transcribe live customer speech, analyze context, and provide incremental suggestions to call center agents while a customer is still speaking. To preserve responsiveness, the GenAI assistant must maintain end-to-end latency under 1 second from speech to initial response display. The architecture must use only managed AWS services and must support bidirectional streaming to ensure that call center agents receive updates in real time.
Which solution will meet these requirements?

A. Use Amazon Transcribe streaming to transcribe call
B. Pass the text to Amazon Comprehend for sentiment analysi
C. Feed the results to Anthropic Claude on Amazon Bedrock by using the InvokeModel AP
D. Store results in Amazon DynamoD
E. Use a WebSocket API to display the results.
F. Use Amazon Transcribe streaming with partial results enabled to deliver fragments oftranscribed text before customers finish speakin
G. Forward text fragments to Amazon Bedrock by using the InvokeModelWithResponseStream AP
H. Stream responses to call center agents through an Amazon API Gateway WebSocket API.
I. Use Amazon Transcribe batch processing to convert calls to tex
J. Pass complete transcripts to Anthropic Claude on Amazon Bedrock by using the ConverseStream AP
K. Return responses through an Amazon Lex chatbot interface.
L. Use the Amazon Transcribe streaming API with an AWS Lambda function to transcribe each audio segmen
M. Call the Amazon Titan Embeddings model on Amazon Bedrock by using the InvokeModel AP
N. Publish results to Amazon SNS.

Answer: B

NEW QUESTION 71
A financial services company wants to develop an Amazon Bedrock application that gives analysts the ability to query quarterly earnings reports and financial statements. The financial documents are typically 5–100 pages long and contain both tabular data and text. The application must provide contextually accurate responses that preserve the relationship between financial metrics and their explanatory text. To support accurate and scalable retrieval, the application must incorporate document segmentation and context management strategies.
Which solution will meet these requirements?

A. Use a direct model invocation approach that uses Anthropic Claude to process each financial document as a single inpu
B. Use fine-tuned prompts that instruct the model to parse tables and text separately.
C. Use Amazon Bedrock Knowledge Bases to create a Retrieval Augmented Generation (RAG) application that retrieves relevant information from contextually chunked sections of financial document

D. Segment documents based on their structural layou
E. Include citations that reference the original source materials.
F. Deploy an Amazon Bedrock agent that has an action group that calls custom AWS Lambda functions to analyze financial document
G. Configure the Lambda functions to perform fixed-size chunking when a user submits a query about financial metrics.
H. Create one specialized Amazon Bedrock application that is optimized for structured dat
I. Create a second application that is optimized for unstructured dat
J. Configure each application to use a tailored chunking strategy that is suited to the application's content typ
K. Implement logic to link queries to the appropriate sources.

Answer: B

NEW QUESTION 73
A financial services company is developing a customer service AI assistant by using Amazon Bedrock. The AI assistant must not discuss investment advice with users. The AI assistant must block harmful content, mask personally identifiable information (PII), and maintain audit trails for compliance reporting. The AI assistant must apply content filtering to both user inputs and model responses based on content sensitivity.
The company requires an Amazon Bedrock guardrail configuration that will effectively enforce policies with minimal false positives. The solution must provide multiple handling strategies for multiple types of sensitive content.
Which solution will meet these requirements?

A. Configure a single guardrail and set content filters to high for all categorie
B. Set up denied topics for investment advice and include sample phrases to bloc
C. Set up sensitive information filters that apply the block action for all PII entitie
D. Apply the guardrail to all model inference calls.
E. Configure multiple guardrails by using tiered policie
F. Create one guardrail and set content filters to hig
G. Configure the guardrail to block PII for public interaction
H. Configure a second guardrail and set content filters to mediu
I. Configure the second guardrail to mask PII for internal us
J. Configure multiple topic-specific guardrails to block investment advice and set up contextual grounding checks.
K. Configure a guardrail and set content filters to medium for harmful conten
L. Set updenied topics for investment advice and include clear definitions and sample phrases to bloc
M. Configure sensitive information filters to mask PII in responses and to block financial information in input
N. Enable both input and output evaluations that use custom blocked messages for audits.
O. Create a separate guardrail for each use cas
P. Create one guardrail that applies a harmful content filte
Q. Create a guardrail to apply topic filters for investment advic
R. Create a guardrail to apply sensitive information filters to block PI
S. Use AWS Step Functions to chain the guardrails sequentially.

Answer: C

NEW QUESTION 74
A company uses Amazon Bedrock to generate technical content for customers. The company has recently experienced a surge in hallucinated outputs when the company??s model generates summaries of long technical documents. The model outputs include inaccurate or fabricated details. The company??s current solution uses a large foundation model (FM) with a basic one-shot prompt that includes the full document in a single input.
The company needs a solution that will reduce hallucinations and meet factual accuracy goals. The solution must process more than 1,000 documents each hour and deliver summaries within 3 seconds for each document.
Which combination of solutions will meet these requirements? (Select TWO.)

A. Implement zero-shot chain-of-thought (CoT) instructions that require step-by-step reasoning with explicit fact verification before the model generates each summary.
B. Use Retrieval Augmented Generation (RAG) with an Amazon Bedrock knowledge bas
C. Apply semantic chunking and tuned embeddings to ground summaries in source content.
D. Configure Amazon Bedrock guardrails to block any generated output that matches patterns that are associated with hallucinated content.
E. Increase the temperature parameter in Amazon Bedrock.
F. Prompt the Amazon Bedrock model to summarize each full document in one pass.

Answer: BC

NEW QUESTION 77
A company uses an organization in AWS Organizations with all features enabled to manage multiple AWS accounts. Employees use Amazon Bedrock across multiple accounts. The company must prevent specific topics and proprietary information from being included in prompts to Amazon Bedrock models. The company must ensure that employees can use only approved Amazon Bedrock models. The company wants to manage these controls centrally.
Which combination of solutions will meet these requirements? (Select TWO.)

A. Create an IAM permissions boundary for each employee's IAM rol
B. Configure the permissions boundary to require an approved Amazon Bedrock guardrail identifier to invoke Amazon Bedrock model
C. Create an SCP that allows employees to use only approved models.
D. Create an SCP that allows employees to use only approved model
E. Configure the SCP to require employees to specify a guardrail identifier in calls to invoke an approved model.
F. Create an SCP that prevents an employee from invoking a model if a centrally deployed guardrail identifier is not specified in a call to the mode
G. Create a permissions boundary on each employee's IAM role that allows each employee to invoke only approved models.
H. Use AWS CloudFormation to create a custom Amazon Bedrock guardrail that has a block filtering polic
I. Use stack sets to deploy the guardrail to each account in the organization.
J. Use AWS CloudFormation to create a custom Amazon Bedrock guardrail that has a mask filtering polic
K. Use stack sets to deploy the guardrail to each account in the organization.

Answer: CD

NEW QUESTION 80
A company uses AWS Lambda functions to build an AI agent solution. A GenAI developer must set up a Model Context Protocol (MCP) server that accesses user information. The GenAI developer must also configure the AI agent to use the new MCP server. The GenAI developer must ensure that only authorized users can access the MCP server.
Which solution will meet these requirements?

A. Use a Lambda function to host the MCP serve
B. Grant the AI agent Lambda functions permission to invoke the Lambda function that hosts the MCP serve
C. Configure the AI agent??s MCP client to invoke the MCP server asynchronously.
D. Use a Lambda function to host the MCP serve
E. Grant the AI agent Lambda functions permission to invoke the Lambda function that hosts the MCP serve
F. Configure the AI agent to use the STDIO transport with the MCP server.
G. Use a Lambda function to host the MCP serve
H. Create an Amazon API Gateway HTTP API that proxies requests to the Lambda functio
I. Configure the AI agent solution to use the Streamable HTTP transport to make requests through the HTTP AP
J. Use Amazon Cognito to enforce OAuth 2.1.
K. Use a Lambda layer to host the MCP serve
L. Add the Lambda layer to the AI agent Lambda function
M. Configure the agentic AI solution to use the STDIO transport to send requests to the MCP serve
N. In the AI agent??s MCP configuration, specify the Lambda layer ARN as the comman
O. Specify the user credentials as environment variables.

Answer: C

NEW QUESTION 82
A hotel company wants to enhance a legacy Java-based property management system (PMS) by adding AI capabilities. The company wants to use Amazon Bedrock Knowledge Bases to provide staff with room availability information and hotel-specific details. The solution must maintain separate access controls for each hotel that the company manages. The solution must provide room availability information in near real time and must maintain consistent performance during peak usage periods.
Which solution will meet these requirements?

A. Deploy a single Amazon Bedrock knowledge base that contains combined data for all hotel
B. Configure AWS Lambda functions to synchronize data from each hotel??s PMS database through direct API connection
C. Implement AWS CloudTrail logging with hotel- specific filters to audit access logs for each hotel??s data.
D. Create an Amazon EventBridge rule for each hotel that is invoked by changes to the PMS databas
E. Configure the rule to send updates to a centralized Amazon Bedrock knowledge base in a management AWS accoun
F. Configure resource-based policies to enforce hotel-specific access controls.
G. Implement one Amazon Bedrock knowledge base for each hotel in a multi-account structur
H. Use direct data ingestion to provide near real-time room availability informatio
I. Schedule regular synchronization for less critical information.
J. Build a centralized Amazon Bedrock Agents solution that uses multiple knowledge base
K. Implement AWS IAM Identity Center with hotel-specific permission sets to control staff access.

Answer: C

NEW QUESTION 83
A software company is using Amazon Q Business to build an AI assistant that allows employees to access company information and personal information by using natural language prompts. The company stores this information in an Amazon S3 bucket.
Each department in the company has a dedicated prefix in the S3 bucket. Each object name includes the S3 prefix of the department that it belongs to. Each department can belong to only a single group in AWS IAM Identity Center. Each employee belongs to a single department.
The company configures Amazon Q Business to access data stored in an S3 bucket as a data source. The company needs to ensure that the AI assistant respects access controls based on the user's IAM Identity Center group membership.
Which solution will meet this requirement with the LEAST operational overhead?

A. Create a JSON file named acl.json in each department folde
B. In each file, create access control entries that specify the IAM Identity Center group that should have access to that department's dat
C. Indicate the location of the JSON file in the Access Control section of the data source settings.
D. Create a single JSON file named acl.json at the top level of the S3 bucke
E. Add access control entries that map each department's S3 prefix to its corresponding IAM Identity Center grou
F. Indicate the location of the JSON file in the Access Control section of the data source settings.
G. For each IAM Identity Center group, create a separate permissions set that denies access to all prefixes in the S3 bucke
H. Add a StringNotEquals condition key to the permissions set for each group that specifies the department each group is associated wit
I. Attach the permissions sets to the Identity Center groups.
J. Create a metadata file named metadata.json at the top level of the S3 bucke
K. Add anAccessControlList object to the file that specifies the S3 path of each department's prefi
L. Specify the IAM Identity Center group that should have access to each department's prefi
M. Reference the file location in the data source metadata settings.

Answer: B

NEW QUESTION 87
A company is implementing a serverless inference API by using AWS Lambda. The API will dynamically invoke multiple AI models hosted on Amazon Bedrock. The company needs to design a solution that can switch between model providers without modifying or redeploying Lambda code in real time. The design must include safe rollout of configuration changes and validation and rollback capabilities.
Which solution will meet these requirements?

A. Store the active model provider in AWS Systems Manager Parameter Stor
B. Configure a Lambda function to read the parameter at runtime to determine which model to invoke.
C. Store the active model provider in AWS AppConfi
D. Configure a Lambda function to read the configuration at runtime to determine which model to invoke.

E. Configure an Amazon API Gateway REST API to route requests to separate Lambda function
F. Hardcode each Lambda function to a specific model provide
G. Switch the integration target manually.
H. Store the active model provider in a JSON file hosted on Amazon S3. Use AWS AppConfig to reference the S3 file as a hosted configuration sourc
I. Configure a Lambda function to read the file through AppConfig at runtime to determine which model to invoke.

Answer: B

NEW QUESTION 92
A company uses AWS Lake Formation to set up a data lake that contains databases and tables for multiple business units across multiple AWS Regions. The company wants to use a foundation model (FM) through Amazon Bedrock to perform fraud detection. The FM must ingest sensitive financial data from the data lake. The data includes some customer personally identifiable information (PII).
The company must design an access control solution that prevents PII from appearing in a production environment. The FM must access only authorized data subsets that have PII redacted from specific data columns. The company must capture audit trails for all data access.
Which solution will meet these requirements?

A. Create a separate dataset in a separate Amazon S3 bucket for each business unit and Region combinatio
B. Configure S3 bucket policies to control access based on IAM roles that are assigned to FM training instance
C. Use S3 access logs to track data access.
D. Configure the FM to authenticate by using AWS Identity and Access Management roles and Lake Formation permissions based on LF-Tag expression
E. Define business units and Regions as LF-Tags that are assigned to databases and table
F. Use AWS CloudTrail to collect comprehensive audit trails of data access.
G. Use direct IAM principal grants on specific databases and tables in Lake Formatio
H. Create a custom application layer that logs access requests and further filters sensitive columns before sending data to the FM.
I. Configure the FM to request temporary credentials from AWS Security Token Servic
J. Access the data by using presigned S3 URLs that are generated by an API that applies business unit and Regional filter
K. Use AWS CloudTrail to collect comprehensive audit trails of data access.

Answer: B

NEW QUESTION 94
A bank is developing a generative AI (GenAI)-powered AI assistant that uses Amazon Bedrock to assist the bank??s website users with account inquiries and financial guidance. The bank must ensure that the AI assistant does not reveal any personally identifiable information (PII) in customer interactions.
The AI assistant must not send PII in prompts to the GenAI model. The AI assistant must not respond to customer requests to provide investment advice. The bank must collect audit logs of all customer interactions, including any images or documents that are transmitted during customer interactions.
Which solution will meet these requirements with the LEAST operational effort?

A. Use Amazon Macie to detect and redact PII in user inputs and in the model response
B. Apply prompt engineering techniques to force the model to avoid investment advice topic
C. Use AWS CloudTrail to capture conversation logs.
D. Use an AWS Lambda function and Amazon Comprehend to detect and redact PI
E. Use Amazon Comprehend topic modeling to prevent the AI assistant from discussing investment advice topic
F. Set up custom metrics in Amazon CloudWatch to capture customer conversations.
G. Configure Amazon Bedrock guardrails to apply a sensitive information policy to detect and filter PI
H. Set up a topic policy to ensure that the AI assistant avoids investment advice topic
I. Use the Converse API to log model invocation
J. Enable delivery and image logging to Amazon S3.
K. Use regex controls to match patterns for PI
L. Apply prompt engineering techniques to avoid returning PII or investment advice topics to customer
M. Enable model invocation logging, delivery logging, and image logging to Amazon S3.

Answer: C

NEW QUESTION 95
A company is using Amazon Bedrock to develop an AI-powered application that uses a foundation model that supports cross-Region inference and provisioned throughput. The application must serve users in Europe and North America with consistently low latency. The application must comply with data residency regulations that require European user data to remain within Europe-based AWS Regions.
During testing, the application experiences service degradation when Regional traffic spikes reach service quotas. The company needs a solution that maintains application resilience and minimizes operational complexity.
Which solution will meet these requirements?

A. Deploy separate Amazon Bedrock instances in North American and European Region
B. Use a custom routing layer that directs traffic based on user locatio
C. Configure Amazon CloudWatch alarms to monitor Regional service usag
D. Use Amazon SNS to send email alerts to the company when usage approaches specified thresholds.
E. Use Amazon Bedrock cross-Region inference profiles by specifying geographical codes in profile IDs when the application calls the InvokeModel AP
F. Configure separate Amazon API Gateway HTTP APIs to direct European and North American users to the appropriate Regional endpoints.
G. Deploy a multi-Region Amazon API Gateway HTTP API and AWS Lambda functions that implement retry logic to handle throttlin
H. Configure the Lambda functions to call the foundation model in the nearest secondary Region when the application reaches service quotas in the primary Regio
I. Use intelligent routing to ensure compliance with data residency requirements.
J. Configure provisioned throughput for Amazon Bedrock in multiple Region
K. Implement failover logic in the application code to switch between Regions when throttling occur
L. Use AWS Global Accelerator to route traffic to the appropriate endpoints based on user location.

Answer: B

NEW QUESTION 99
A financial services company is creating a Retrieval Augmented Generation (RAG) application that uses Amazon Bedrock to generate summaries of market activities. The application relies on a vector database that stores a small proprietary dataset with a low index count. The application must perform similarity

searches. The Amazon Bedrock model??s responses must maximize accuracy and maintain high performance.
The company needs to configure the vector database and integrate it with the application. Which solution will meet these requirements?

A. Launch an Amazon MemoryDB cluster and configure the index by using the Flat algorith
B. Configure a horizontal scaling policy based on performance metrics.
C. Launch an Amazon MemoryDB cluster and configure the index by using the Hierarchical Navigable Small World (HNSW) algorith
D. Configure a vertical scaling policy based on performance metrics.
E. Launch an Amazon Aurora PostgreSQL cluster and configure the index by using the Inverted File with Flat Compression (IVFFlat) algorith
F. Configure the instance class to scale to a larger size when the load increases.
G. Launch an Amazon DocumentDB cluster that has an IVFFlat index and a high probe valu
H. Configure connections to the cluster as a replica se
I. Distribute reads to replica instances.

Answer: B

NEW QUESTION 100
A media company is launching a platform that allows thousands of users every hour to upload images and text content. The platform uses Amazon Bedrock to process the uploaded content to generate creative compositions.
The company needs a solution to ensure that the platform does not process or produce inappropriate content. The platform must not expose personally identifiable information (PII) in the compositions. The solution must integrate with the company's existing Amazon S3 storage workflow.
Which solution will meet these requirements with the LEAST infrastructure management overhead?

A. Enable the Enhanced Monitoring too
B. Use an Amazon CloudWatch alarm to filter traffic to the platfor
C. Use Amazon Comprehend PII detection to pre-process the dat
D. Create a CloudWatch alarm to monitor for Amazon Comprehend PII detection event
E. Create an AWS Step Functions workflow that includes an Amazon Rekognition image moderation step.
F. Use an Amazon API Gateway HTTP API with request validation templates to screen content before storing the uploaded content in Amazon S3. Use Amazon SageMaker AI to build custom content moderation models that process content before sending the processed content to Amazon Bedrock.
G. Create an Amazon Cognito user pool that uses pre-authentication AWS Lambda functions to run content moderation check
H. Use Amazon Textract to filter text content and Amazon Rekognition to filter image content before allowing users to upload content to the platform.
I. Create an AWS Step Functions workflow that uses built-in Amazon Bedrock guardrails to filter conten
J. Use Amazon Comprehend PII detection to pre-process the conten
K. Use Amazon Rekognition image moderation.

Answer: D

NEW QUESTION 103
A legal research company has a Retrieval Augmented Generation (RAG) application that uses Amazon Bedrock and Amazon OpenSearch Service. The application stores 768- dimensional vector embeddings for 15 million legal documents, including statutes, court rulings, and case summaries.
The company's current chunking strategy segments text into fixed-length blocks of 500 tokens. The current chunking strategy often splits contextually linked information such as legal arguments, court opinions, or statute references across separate chunks. Researchers report that generated outputs frequently omit key context or cite outdated legal information.
Recent application logs show a 40% increase in response times. The p95 latency metric exceeds 2 seconds. The company expects storage needs for the application to grow from 90 GB to 360 GB within a year.
The company needs a solution to improve retrieval relevance and system performance at scale.
Which solution will meet these requirements?

A. Increase the embedding vector dimensionality from 768 to 4,096 without changing the existing chunking or pre-processing strategy.
B. Replace dynamic retrieval with static, pre-written summaries that are stored in Amazon S3. Use Amazon CloudFront to serve the summaries to reduce compute demand and improve predictability.
C. Update the chunking strategy to use semantic boundaries such as complete legal arguments, clauses, or sections rather than fixed token limit
D. Regenerate vector embeddings to align with the new chunk structure.
E. Migrate from OpenSearch Service to Amazon DynamoD
F. Implement keyword-based indexes to enable faster lookups for legal concepts.

Answer: C

NEW QUESTION 108
An elevator service company has developed an AI assistant application by using Amazon Bedrock. The application generates elevator maintenance recommendations to support the company??s elevator technicians. The company uses Amazon Kinesis Data Streams to collect the elevator sensor data.
New regulatory rules require that a human technician must review all AI-generated recommendations. The company needs to establish human oversight workflows to review and approve AI recommendations. The company must store all human technician review decisions for audit purposes.
Which solution will meet these requirements?

A. Create a custom approval workflow by using AWS Lambda functions and Amazon SQS queues for human review of AI recommendation
B. Store all review decisions in Amazon DynamoDB for audit purposes.
C. Create an AWS Step Functions workflow that has a human approval step that uses the waitForTaskToken API to pause executio
D. After a human technician completes a review, use an AWS Lambda function to call the SendTaskSuccess API with the approval decisio
E. Store all review decisions in Amazon DynamoDB.
F. Create an AWS Glue workflow that has a human approval ste
G. After the human technician review, integrate the application with an AWS Lambda function that calls the SendTaskSuccess AP
H. Store all human technician review decisions in Amazon DynamoDB.
I. Configure Amazon EventBridge rules with custom event patterns to route AI recommendations to human technicians for revie
J. Create AWS Glue jobs to process human technician approval queue
K. Use Amazon ElastiCache to cache all human technician review decisions.

Answer: B

NEW QUESTION 113
A company is developing a customer support application that uses Amazon Bedrock foundation models (FMs) to provide real-time AI assistance to the company??s employees. The application must display AI-generated responses character by character as the responses are generated. The application needs to support thousands of concurrent users with minimal latency. The responses typically take 15 to 45 seconds to finish.
Which solution will meet these requirements?

A. Configure an Amazon API Gateway WebSocket API with an AWS Lambda integratio
B. Configure the WebSocket API to invoke the Amazon Bedrock InvokeModelWithResponseStream API and stream partial responses through WebSocket connections.
C. Configure an Amazon API Gateway REST API with an AWS Lambda integratio
D. Configure the REST API to invoke the Amazon Bedrock standard InvokeModel API and implement frontend client-side polling every 100 ms for complete response chunks.
E. Implement direct frontend client connections to Amazon Bedrock by using IAM user credentials and the InvokeModelWithResponseStream API without any intermediate gateway or proxy layer.
F. Configure an Amazon API Gateway HTTP API with an AWS Lambda integratio
G. Configure the HTTP API to cache complete responses in an Amazon DynamoDB table and serve the responses through multiple paginated GET requests to frontend clients.

Answer: A

NEW QUESTION 117
A GenAI developer is evaluating Amazon Bedrock foundation models (FMs) to enhance a Europe-based company's internal business application. The company has a multi-account landing zone in AWS Control Tower. The company uses Service Control Policies (SCPs) to allow its accounts to use only the eu-north-1 and eu-west-1 Regions. All customer data must remain in private networks within the approved AWS Regions.
The GenAI developer selects an FM based on analysis and testing and hosts the model in the eu-central-1 Region and the eu-west-3 Region. The GenAI developer must enable access to the FM for the company's employees. The GenAI developer must ensure that requests to the FM are private and remain within the same Regions as the FM.
Which solution will meet these requirements?

A. Deploy an AWS Lambda function that is exposed by a private Amazon API Gateway REST API to a VPC in eu-north-1. Create a VPC endpoint for the selected FM in eu- central-1 and eu-west-3. Extend existing SCPs to allow employees to use the F
B. Integrate the REST API with the business application.
C. Deploy the FM on Amazon EC2 instances in eu-north-1. Deploy a private Amazon API Gateway REST API in front of the EC2 instance
D. Configure an Amazon Bedrock VPC endpoin
E. Integrate the REST API with the business application.
F. Configure the FM to use cross-Region inference through a Europe-scoped endpoin
G. Configure an Amazon Bedrock VPC endpoin
H. Extend existing SCPs to allow employees to use the FM through inference profiles in Europe-based Regions where the FM is availabl
I. Use an inference profile to integrate Amazon Bedrock with the business application.
J. Deploy the FM in Amazon SageMaker in eu-north-1. Configure a SageMaker VPC endpoin
K. Extend existing SCPs to allow employees to use the SageMaker endpoin
L. Integrate the FM in SageMaker with the business application.

Answer: C

NEW QUESTION 121
A company is developing a customer communication platform that uses an AI assistant powered by an Amazon Bedrock foundation model (FM). The AI assistant summarizes customer messages and generates initial response drafts.
The company wants to use Amazon Comprehend to implement layered content filtering. The layered content filtering must prevent sharing of offensive content, protect customer privacy, and detect potential inappropriate advice solicitation. Inappropriate advice solicitation includes requests for unethical practices, harmful activities, or manipulative behaviors.
The solution must maintain acceptable overall response times, so all pre-processing filters must finish before the content reaches the FM.
Which solution will meet these requirements?

A. Use parallel processing with asynchronous API call
B. Use toxicity detection for offensive conten
C. Use prompt safety classification for inappropriate advice solicitatio
D. Use personally identifiable information (PII) detection without redaction.
E. Use custom classification to build an FM that detects offensive content and inappropriate advice solicitatio
F. Apply personally identifiable information (PII) detection as a secondary filter only when messages pass the custom classifier.
G. Deploy a multi-stage proces
H. Configure the process to use prompt safety classification first, then toxicity detection on safe prompts only, and finally personally identifiable information (PII) detection in streaming mod
I. Route flagged messages through Amazon EventBridge for human review.
J. Use toxicity detection with thresholds configured to 0.5 for all categorie
K. Use parallel processing for both prompt safety classification and personally identifiable information (PII) detection with entity redactio
L. Apply Amazon CloudWatch alarms to filter metrics.

Answer: D

NEW QUESTION 124
.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

AIP-C01 Practice Exam Features:

* AIP-C01 Questions and Answers Updated Frequently

* AIP-C01 Practice Questions Verified by Expert Senior Certified Staff

* AIP-C01 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* AIP-C01 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year