Amazon-Web-Services

Exam Questions AIP-C01

AWS Certified Generative AI Developer - Professional

About Exambible

Your Partner of IT Exam

Found in 1998

Exambible is a company specialized on providing high quality IT exam practice study materials, especially Cisco CCNA, CCDA, CCNP, CCIE, Checkpoint CCSE, CompTIA A+, Network+ certification practice exams and so on. We guarantee that the candidates will not only pass any IT exam at the first attempt but also get profound understanding about the certificates they have got. There are so many alike companies in this industry, however, Exambible has its unique advantages that other companies could not achieve.

Our Advances

* 99.9% Uptime

All examinations will be up to date.

* 24/7 Quality Support

We will provide service round the clock.

* 100% Pass Rate

Our guarantee that you will pass the exam.

* Unique Gurantee

If you do not pass the exam at the first time, we will not only arrange FULL REFUND for you, but also provide you another exam of your claim, ABSOLUTELY FREE!

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 runs a Retrieval Augmented Generation (RAG) application that uses Amazon Bedrock Knowledge Bases to perform regulatory compliance queries. The application uses the RetrieveAndGenerateStream API. The application retrieves relevant documents from a knowledge base that contains more than 50,000 regulatory documents, legal precedents, and policy updates.
The RAG application is producing suboptimal responses because the initial retrieval often returns semantically similar but contextually irrelevant documents. The poor responses are causing model hallucinations and incorrect regulatory guidance. The company needs to improve the performance of the RAG application so it returns more relevant documents.
Which solution will meet this requirement with the LEAST operational overhead?

A. Deploy an Amazon SageMaker endpoint to run a fine-tuned ranking mode
B. Use an Amazon API Gateway REST API to route request
C. Configure the application to make requests through the REST API to rerank the results.
D. Use Amazon Comprehend to classify documents and apply relevance score
E. Integrate the RAG application??s reranking process with Amazon Textract to run document analysi
F. Use Amazon Neptune to perform graph-based relevance calculations.
G. Implement a retrieval pipeline that uses the Amazon Bedrock Knowledge Bases Retrieve API to perform initial document retrieva
H. Call the Amazon Bedrock Rerank API to rerank the result
I. Invoke the InvokeModelWithResponseStream operation to generate responses.
J. Use the latest Amazon reranker model through the reranking configuration within Amazon Bedrock Knowledge Base
K. Use the model to improve document relevance scoring and to reorder results based on contextual assessments.

Answer: D

NEW QUESTION 4
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 5
A company deploys multiple Amazon Bedrock–based generative AI (GenAI) applications across multiple business units for customer service, content generation, and document analysis. Some applications show unpredictable token consumption patterns. The company requires a comprehensive observability solution that provides real-time visibility into token usage patterns across multiple models. The observability solution must support custom dashboards for multiple stakeholder groups and provide alerting capabilities for token consumption across all the foundation models that the company??s applications use.
Which combination of solutions will meet these requirements with the LEAST operational overhead? (Select TWO.)

A. Use Amazon CloudWatch metrics as data sources to create custom Amazon QuickSight dashboards that show token usage trends and usage patterns across FMs.
B. Use CloudWatch Logs Insights to analyze Amazon Bedrock invocation logs for token consumption patterns and usage attribution by applicatio
C. Create custom queries to identify high-usage scenario
D. Add log widgets to dashboards to enable continuous monitoring.
E. Create custom Amazon CloudWatch dashboards that combine native Amazon Bedrock token and invocation CloudWatch metric
F. Set up CloudWatch alarms to monitor token usage thresholds.
G. Create dashboards that show token usage trends and patterns across the company??s FMs by using an Amazon Bedrock zero-ETL integration with Amazon Managed Grafana.
H. Implement Amazon EventBridge rules to capture Amazon Bedrock model invocation event
I. Route token usage data to Amazon OpenSearch Serverless by using Amazon Data Firehos
J. Use OpenSearch dashboards to analyze usage patterns.

Answer: CD

NEW QUESTION 6
A company has a recommendation system. The system's applications run on Amazon EC2 instances. The applications make API calls to Amazon Bedrock foundation models (FMs) to analyze customer behavior and generate personalized product recommendations.
The system is experiencing intermittent issues. Some recommendations do not match customer preferences. The company needs an observability solution to monitor operational metrics and detect patterns of operational performance degradation compared to established baselines. The solution must also generate alerts with correlation data within 10 minutes when FM behavior deviates from expected patterns.
Which solution will meet these requirements?

A. Configure Amazon CloudWatch Container Insights for the application infrastructur
B. Set up CloudWatch alarms for latency threshold
C. Add custom metrics for token counts by using the CloudWatch embedded metric forma
D. Create CloudWatch dashboards to visualize the data.
E. Implement AWS X-Ray to trace requests through the application component
F. Enable CloudWatch Logs Insights for error pattern detectio
G. Set up AWS CloudTrail to monitor all API calls to Amazon Bedroc
H. Create custom dashboards in Amazon QuickSight.
I. Enable Amazon CloudWatch Application Insights for the application resource
J. Create custom metrics for recommendation quality, token usage, and response latency by using the CloudWatch embedded metric format with dimensions for request types and user segment
K. Configure CloudWatch anomaly detection on the model metric
L. Establish log pattern analysis by using CloudWatch Logs Insights.
M. Use Amazon OpenSearch Service with the Observability plugi
N. Ingest model metrics and logs by using Amazon Kinesi
O. Create custom Piped Processing Language (PPL) queries to analyze model behavior pattern
P. Establish operational dashboards to visualizeanomalies in real time.

Answer: C

NEW QUESTION 7
A financial services company needs to build a document analysis system that uses Amazon Bedrock to process quarterly reports. The system must analyze financial data, perform sentiment analysis, and validate compliance across batches of reports. Each batch contains 5 reports. Each report requires multiple foundation model (FM) calls. The solution must finish the analysis within 10 seconds for each batch. Current sequential processing takes 45 seconds for each batch.
Which solution will meet these requirements?

A. Use AWS Lambda functions with provisioned concurrency to process each analysis type sequentiall
B. Configure the Lambda function timeouts to 10 second
C. Configure automatic retries with exponential backoff.
D. Use AWS Step Functions with a Parallel state to invoke separate AWS Lambdafunctions for each analysis type simultaneousl
E. Configure Amazon Bedrock client timeout
F. Use Amazon CloudWatch metrics to track execution time and model inference latency.
G. Create an Amazon SQS queue to buffer analysis request
H. Deploy multiple AWS Lambda functions with reserved concurrenc
I. Configure each Lambda function to process different aspects of each report sequentially and then combine the results.
J. Deploy an Amazon ECS cluster that runs containers that process each report sequentiall
K. Use a load balancer to distribute batch workload
L. Configure an auto-scaling policy based on CPU utilization.

Answer: B

NEW QUESTION 8
A company uses Amazon Bedrock to build a Retrieval Augmented Generation (RAG) system. The RAG system uses an Amazon Bedrock Knowledge Bases that is based on an Amazon S3 bucket as the data source for emergency news video content. The system retrieves transcripts, archived reports, and related documents from the S3 bucket.
The RAG system uses state-of-the-art embedding models and a high-performing retrieval setup. However, users report slow responses and irrelevant results, which cause decreased user satisfaction. The company notices that vector searches are evaluating too many documents across too many content types and over long periods of time.
The company determines that the underlying models will not benefit from additional fine- tuning. The company must improve retrieval accuracy by applying smarter constraints and wants a solution that requires minimal changes to the existing architecture.
Which solution will meet these requirements?

A. Enhance embeddings by using a domain-adapted model that is specifically trained on emergency news content for improved vector similarity.
B. Migrate to Amazon OpenSearch Servic
C. Use vector fields and metadata filters to define the scope of results retrieval.
D. Enable metadata-aware filtering within the Amazon Bedrock knowledge base by indexing S3 object metadata.
E. Migrate to an Amazon Q Business index to perform structured metadata filtering and document categorization during retrieval.

Answer: C

NEW QUESTION 9
A specialty coffee company has a mobile app that generates personalized coffee roast profiles by using Amazon Bedrock with a three-stage prompt chain. The prompt chain converts user inputs into structured metadata, retrieves relevant logs for coffee roasts, and generates a personalized roast recommendation for each customer.
Users in multiple AWS Regions report inconsistent roast recommendations for identical inputs, slow inference during the retrieval step, and unsafe recommendations such as brewing at excessively high temperatures. The company must improve the stability of outputs for repeated inputs. The company must also improve app performance and the safety of the app's outputs. The updated solution must ensure 99.5% output consistency for identical inputs and achieve inference latency of less than 1 second. The solution must also block unsafe or hallucinated recommendations by using validated safety controls.
Which solution will meet these requirements?

A. Deploy Amazon Bedrock with provisioned throughput to stabilize inference latenc
B. Apply Amazon Bedrock guardrails that have semantic denial rules to block unsafe output
C. Use Amazon Bedrock Prompt Management to manage prompts by using approval workflows.
D. Use Amazon Bedrock Agents to manage chainin
E. Log model inputs and outputs to Amazon CloudWatch Log
F. Use logs from Amazon CloudWatch to perform A/B testing for prompt versions.
G. Cache prompt results in Amazon ElastiCach
H. Use AWS Lambda functions to pre- process metadata and to trace end-to-end latenc
I. Use AWS X-Ray to identify and remediate performance bottlenecks.
J. Use Amazon Kendra to improve roast log retrieval accurac
K. Store normalized prompt metadata within Amazon DynamoD
L. Use AWS Step Functions to orchestrate multi-step prompts.

Answer: A

NEW QUESTION 10
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 10
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 14
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 17
A company is developing a generative AI (GenAI) application that analyzes customer service calls in real time and generates suggested responses for human customer service agents. The application must process 500,000 concurrent calls during peak hours with less than 200 ms end-to-end latency for each suggestion. The company uses existing architecture to transcribe customer call audio streams. The application must not exceed a predefined monthly compute budget and must maintain auto scaling capabilities.
Which solution will meet these requirements?

A. Deploy a large, complex reasoning model on Amazon Bedroc
B. Purchase provisioned throughput and optimize for batch processing.
C. Deploy a low-latency, real-time optimized model on Amazon Bedroc
D. Purchase provisioned throughput and set up automatic scaling policies.
E. Deploy a large language model (LLM) on an Amazon SageMaker real-time endpoint that uses dedicated GPU instances.
F. Deploy a mid-sized language model on an Amazon SageMaker serverless endpoint that is optimized for batch processing.

Answer: B

NEW QUESTION 19
An ecommerce company is developing a generative AI (GenAI) solution that uses Amazon Bedrock with Anthropic Claude to recommend products to customers. Customers report that some recommended products are not available for sale or are not relevant. Customers also report long response times for some recommendations.
The company confirms that most customer interactions are unique and that the solution recommends products not present in the product catalog.
Which solution will meet this requirement?

A. Increase grounding within Amazon Bedrock Guardrail
B. Enable automated reasoning check
C. Set up provisioned throughput.
D. Use prompt engineering to restrict model responses to relevant product
E. Use streaming inference to reduce perceived latency.
F. Create an Amazon Bedrock Knowledge Bases and implement Retrieval Augmented Generation (RAG). Set the PerformanceConfigLatency parameter to optimized.
G. Store product catalog data in Amazon OpenSearch Servic
H. Validate model recommendations against the catalo
I. Use Amazon DynamoDB for response caching.

Answer: C

NEW QUESTION 23
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 24
A financial services company needs to pre-process unstructured data such as customer transcripts, financial reports, and documentation. The company stores the unstructured data in Amazon S3 to support an Amazon Bedrock application.

The company must validate data quality, create auditable metadata, monitor data metrics, and customize text chunking to optimize foundation model (FM) performance.
Which solution will meet these requirements with the LEAST development effort?

A. Use Amazon SageMaker Data Wrangler to create a data flo
B. Configure Amazon CloudWatch metrics and alarms to monitor data qualit
C. Use a custom AWS Lambda function to pre-process the dat
D. Load processed data into Amazon Bedrock.
E. Set up an AWS Glue crawler to catalog data source
F. Create AWS Glue ETL jobs to run custom transformation script
G. Use AWS Glue Data Quality to validate and monitor data qualit
H. Load processed data into Amazon Bedrock.
I. Use Amazon Comprehend to extract entitie
J. Create an AWS Lambda function to chunk tex
K. Run Amazon Athena to query and validate data qualit
L. Load processed data into Amazon Bedrock.
M. Create an AWS Step Functions workflow to orchestrate data pre-processing task
N. Run custom code on Amazon EC2 instance
O. Use Amazon SageMaker Model Monitor to monitor data qualit
P. Load processed data into Amazon Bedrock.

Answer: B

NEW QUESTION 28
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 32
A company has a customer service application that uses Amazon Bedrock to generate personalized responses to customer inquiries. The company needs to establish a quality assurance process to evaluate prompt effectiveness and model configurations across updates. The process must automatically compare outputs from multiple prompt templates, detect response quality issues, provide quantitative metrics, and allow human reviewers to give feedback on responses. The process must prevent configurations that do not meet a predefined quality threshold from being deployed.
Which solution will meet these requirements?

A. Create an AWS Lambda function that sends sample customer inquiries to multiple Amazon Bedrock model configurations and stores responses in Amazon S3. Use Amazon QuickSight to visualize response pattern
B. Manually review outputs dail
C. Use AWS CodePipeline to deploy configurations that meet the quality threshold.
D. Use Amazon Bedrock evaluation jobs to compare model outputs by using custom prompt dataset
E. Configure AWS CodePipeline to run the evaluation jobs when prompt templates chang
F. Configure CodePipeline to deploy only configurations that exceed the predefined quality threshold.
G. Set up Amazon CloudWatch alarms to monitor response latency and error rates from Amazon Bedroc
H. Use Amazon EventBridge rules to notify teams when thresholds are exceede
I. Configure a manual approval workflow in AWS Systems Manager.
J. Use AWS Lambda functions to create an automated testing framework that samples production traffic and routes duplicate requests to the updated model versio
K. Use Amazon Comprehend sentiment analysis to compare result
L. Block deployment if sentiment scores decrease.

Answer: B

NEW QUESTION 37
A company is using Amazon Bedrock to build a customer-facing AI assistant that handles sensitive customer inquiries. The company must use defense-in-depth safety controls to block sophisticated prompt injection attacks. The company must keep audit logs of all safety interventions. The AI assistant must have cross-Region failover capabilities.
Which solution will meet these requirements?

A. Configure Amazon Bedrock guardrails with content filters set to high to protect against prompt injection attack
B. Use a guardrail profile to implement cross-Region guardrail inferenc
C. Use Amazon CloudWatch Logs with custom metrics to capture detailed guardrail intervention events.
D. Configure Amazon Bedrock guardrails with content filters set to hig
E. Use AWS WAF to block suspicious input

F. Use AWS CloudTrail to log API calls.
G. Deploy Amazon Comprehend custom classifiers to detect prompt injection attack
H. Use Amazon API Gateway request validatio
I. Use CloudWatch Logs to capture intervention events.
J. Configure Amazon Bedrock guardrails with custom content filters and word filters set to hig
K. Configure cross-Region guardrail replication for failove
L. Store logs in AWS CloudTrail for compliance auditing.

Answer: A

NEW QUESTION 42
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 46
A financial services company is developing a customer service AI assistant application that uses a foundation model (FM) in Amazon Bedrock. The application must provide transparent responses by documenting reasoning and by citing sources that are used for Retrieval Augmented Generation (RAG). The application must capture comprehensive audit trails for all responses to users. The application must be able to serve up to 10,000 concurrent users and must respond to each customer inquiry within 2 seconds.
Which solution will meet these requirements with the LEAST operational overhead?

A. Enable tracing for Amazon Bedrock Agent
B. Configure structured prompts that direct the FM to provide evidence presentation
C. Integrate Amazon Bedrock Knowledge Bases with data sources to enable RA
D. Configure the application to reference and cite authoritative conten
E. Deploy the application in a Multi-AZ architectur
F. Use Amazon API Gateway and AWS Lambda functions to scale the applicatio
G. Use Amazon CloudFront to provide low- latency delivery.
H. Enable tracing for Amazon Bedrock agent
I. Integrate a custom RAG pipeline with Amazon OpenSearch Service to retrieve and cite source
J. Configure structured prompts to present retrieved evidenc
K. Deploy the application behind an Amazon API Gateway REST AP
L. Use AWS Lambda functions and Amazon CloudFront to scale the application and to provide low latenc
M. Store logs in Amazon S3 and use AWS CloudTrail to capture audit trails.
N. Use Amazon CloudWatch to monitor latency and error rate
O. Embed model prompts directly in the application backend to cite source
P. Store application interactions with users in Amazon RDS for audits.
Q. Store generated responses and supporting evidence in an Amazon S3 bucke
R. Enable versioning on the bucket for audit
S. Use AWS Glue to catalog retrieved document
T. Process the retrieved documents in Amazon Athena to generate periodic compliance reports.

Answer: A

NEW QUESTION 51
A company is developing a generative AI (GenAI) application by using Amazon Bedrock. The application will analyze patterns and relationships in the company??s data. The application will process millions of new data points daily across AWS Regions in Europe, North America, and Asia before storing the data in Amazon S3. The application must comply with local data protection and storage regulations. Data residency and processing must occur within the same continent. The application must also maintain audit trails of the application??s decision-making processes and provide data classification capabilities.
Which solution will meet these requirements?

A. Deploy the application in each Region with local IAM policie
B. Use Amazon Bedrock cross-Region inference to distribute the workloa
C. Use Amazon CloudWatch to log AI decision-making processe
D. Manually track compliance certifications across Regions.
E. Use SCPs with AWS Organizations to manage location-specific permission
F. Use AWS CloudTrail immutable logs to audit decision-making processe
G. Import a custom model into Amazon Bedrock and deploy the model to each Region.
H. Use Amazon S3 Object Lock with Region-specific S3 bucket policie
I. Pre-process the data points within the Region based on geographic origin before sending the data points to Amazon Bedroc
J. Use Amazon Macie to classify the dat
K. Use AWS CloudTrail immutable logs to audit the decision-making processes.
L. Create separate AWS accounts for each Region with individual compliance framework
M. Use Amazon SageMaker AI with custom monitorin

N. Create manual compliance reports for each regulatory jurisdiction.

Answer: C

NEW QUESTION 53

.....

Relate Links

100% Pass Your AIP-C01 Exam with Exambible Prep Materials

https://www.exambible.com/AIP-C01-exam/

Contact us

We are proud of our high-quality customer service, which serves you around the clock 24/7.

Viste - https://www.exambible.com/