

VCE4Plus



Everything you need to prepare, learn & pass your certification exam easily.

Pass Your Next Certification Exam Fast!

365 days free updates. First attempt guaranteed success.

Choose the version that fits your needs	PDF Version	Desktop Test Engine	Online Test Engine
Latest and Up-to-Date exam dumps with real exam questions answers.	✓	✓	✓
Get 12-Months free updates without any extra charges.	✓	✓	✓
Experience same exam environment before appearing in the certification exam.	✗	✓	✓
100% exam passing guarantee in the first attempt.	✓	✓	✓
20% discount on more than one license and 30% discount on 5+ license purchases.	✗	✓	✓
100% secure purchase on SSL.	✓	✓	✓
Completely private purchase without sharing your personal info with anyone.	✓	✓	✓

<http://www.vce4plus.com>

Accurate exam material ensure you pass for sure by your first attempt - VCE4Plus

Exam : **AWS-Certified-Data-Analytics-Specialty-KR**

Title : AWS Certified Data Analytics - Specialty (DAS -C01 Korean Version)

Vendor : Amazon

Version : DEMO

QUESTION NO: 1

한 금융 서비스 회사가 Amazon S3에 데이터 레이크 솔루션을 구축하고 있습니다. 회사는 일회성 쿼리 및 비즈니스 인텔리전스 보고서에 대한 사용자 요구 사항을 충족하기 위해 AWS의 분석 서비스를 사용할 계획입니다. 열의 일부에는 개인 식별 정보(PII)가 포함됩니다. 승인된 사용자만 일반 텍스트 PII 데이터를 볼 수 있어야 합니다.

이러한 요구 사항을 충족하는 가장 운영 효율적인 솔루션은 무엇입니까?

A. PII 데이터를 볼 수 있는 권한이 있는 사용자에게 액세스를 허용하도록 데이터 레이크의 각 S3 버킷에 대한 버킷 정책을 정의합니다. AWS Glue를 사용하여 데이터를 카탈로그화합니다. 두 개의 IAM 역할을 생성합니다. PII 열에 대한 액세스 권한이 있는 권한 정책을 하나의 역할에 연결합니다. 이러한 권한이 없는 정책을 다른 역할에 연결합니다.

B. AWS Lake Formation에 S3 위치를 등록합니다. 두 개의 IAM 역할을 생성합니다. Lake Formation 데이터 권한을 사용하여 한 역할의 모든 열에 선택 권한을 부여합니다. 다른 역할에 대해 PII가 아닌 데이터가 포함된 열에만 선택 권한을 부여합니다.

C. AWS Lake Formation에 S3 위치를 등록합니다. AWS Glue 작업을 생성하여 데이터에서 PII 열을 제거하고 다른 데이터 레이크 S3 버킷에 별도의 데이터 복사본을 생성하는 ETL 워크플로를 생성합니다. Lake Formation에 새 S3 위치를 등록합니다. 사용자에게 PII 데이터를 볼 수 있는 권한이 있는지 여부에 따라 각 데이터 레이크 데이터에 대한 권한을 사용자에게 부여합니다.

D. AWS Lake Formation에 S3 위치를 등록합니다. 두 개의 IAM 역할을 생성합니다. PII 열에 대한 액세스 권한이 있는 권한 정책을 하나의 역할에 연결합니다. 이러한 권한이 없는 정책을 다른 역할에 연결합니다. 각 다운스트림 분석 서비스에 대해 기본 보안 기능과 IAM 역할을 사용하여 PII 데이터를 보호합니다.

Answer: B

Explanation:

This solution meets the requirements because:

AWS Lake Formation is a fully managed service that allows you to build, secure, and manage data lakes on AWS¹. You can use Lake Formation to register your S3 locations as data sources and catalog your data using AWS Glue¹.

AWS Lake Formation provides fine-grained data permissions that enable you to control access to your data at the column or row level¹. You can use Lake Formation to create two IAM roles and grant them different Select permissions based on the PII status of the columns¹.

AWS Lake Formation integrates with various analytics services from AWS, such as Amazon Athena, Amazon Redshift, Amazon EMR, and Amazon QuickSight¹. You can use these services to query and visualize your data in S3 using the IAM roles and permissions defined by Lake Formation¹.

QUESTION NO: 2

전력 유틸리티 회사는 전력 소비에 대한 실시간 업데이트를 얻기 위해 수천 대의 스마트 미터를 배포하고 있습니다. 이 회사는 Amazon Kinesis Data Streams를 사용하여 스마트 미터에서 데이터 스트림을 수집하고 있습니다. 소비자 애플리케이션은 Kinesis Client Library(KCL)를 사용하여 a 스트림 데이터를 검색합니다. 회사에는 소비자 애플리케이션이 하나만 있습니다.

회사는 레코드가 스트림에 기록된 순간부터 소비자 애플리케이션에서 레코드를 읽을 때까지 평균 1초의 대기 시간을 관찰합니다. 회사는 이 대기 시간을 500밀리초로 줄여야 합니다.

어떤 솔루션이 이러한 요구 사항을 충족합니까?

- A. Kinesis Data Streams에서 향상된 팬아웃을 사용합니다.
- B. Kinesis 데이터 스트림의 샤드 수를 늘립니다.
- C. KCL 기본 설정을 재정의하여 전파 지연을 줄입니다.
- D. Amazon Kinesis Data Firehose를 사용하여 소비자를 개발합니다.

Answer: C

Explanation:

The KCL defaults are set to follow the best practice of polling every 1 second. This default results in average propagation delays that are typically below 1 second.

QUESTION NO: 3

소프트웨어 회사는 계측 데이터를 사용하여 오류를 감지하고 해결하여 응용 프로그램 복구 시간을 개선하려고 합니다. 회사는 오류율 및 응답 시간 급증과 같은 API 사용 이상 현상을 거의 실시간(NRT)으로 감지해야 합니다. 회사는 또한 데이터 분석가가 NRT에서 로그 분석을 위해 대시보드에 액세스할 수 있어야 합니다. 어떤 솔루션이 이러한 요구 사항을 충족합니까?

- A. Amazon Kinesis Data Firehose를 데이터 로깅을 위한 데이터 전송 계층으로 사용 Amazon Kinesis Data Analytics를 사용하여 NRT API 사용 이상 발견 Kinesis Data Firehose를 사용하여 검색, 로그를 위해 Amazon OpenSearch Service(Amazon Elasticsearch Service)에 로그 데이터 전달 분석 및 애플리케이션 모니터링 대시보드용 Amazon OpenSearch Service(Amazon Elasticsearch Service)의 OpenSearch 대시보드(Kibana)를 사용합니다.
- B. 데이터 로깅을 위한 데이터 전송 계층으로 Amazon Kinesis Data Analytics를 사용합니다. Amazon Kinesis Data Streams를 사용하여 NRT 모니터링 지표를 발견하십시오. Amazon Kinesis Data Firehose를 사용하여 검색, 로그 분석 및 애플리케이션 모니터링을 위해 Amazon OpenSearch Service(Amazon Elasticsearch Service)에 로그 데이터 전달 대시보드에 Amazon QuickSight 사용
- C. Amazon Kinesis Data Analytics를 데이터 로깅을 위한 데이터 전송 계층으로 사용하고 NRT 모니터링 메트릭을 발견합니다. Amazon Kinesis Data Firehose를 사용하여 검색, 로그 분석 및 애플리케이션 모니터링을 위해 Amazon OpenSearch Service(Amazon Elasticsearch Service)에 로그 데이터를 제공합니다. 대시보드에 Amazon OpenSearch Service(Amazon Elasticsearch Service)의 OpenSearch 대시보드(Kibana) 사용
- D. Amazon Kinesis Data Firehose를 데이터 로깅을 위한 데이터 전송 계층으로 사용 Amazon Kinesis Data Analytics를 사용하여 NRT 모니터링 지표 발견 Amazon Kinesis Data Streams를 사용하여 검색, 로그 분석을 위해 Amazon OpenSearch Service(Amazon Elasticsearch Service)에 로그 데이터 전달 및 애플리케이션 모니터링 대시보드에 Amazon QuickSight를 사용합니다.

Answer: C

QUESTION NO: 4

마케팅 회사가 클릭 스트림 데이터를 수집합니다. 회사는 데이터를 Amazon Kinesis Data Firehose로 보내고 Amazon S3에 데이터를 저장합니다. 회사는 여러 부서의 수백 명의 사용자가 사용할 일련의 대시보드를 구축하려고 합니다. 회사는 Amazon QuickSight를 사용하여 다음을 수행합니다. 이 대시보드 개발 회사는 리소스가 제한되어 있고 클릭스트림 활동에 대한 일일 업데이트를 확장하고 제공할 수 있는 솔루션을 원합니다. 가장 비용 효율적인 솔루션을 제공하는 옵션 조합은 무엇입니까? (2개 선택)

- A. Amazon Redshift를 사용하여 클릭스트림 데이터 저장 및 쿼리
- B. 직접 SQL 쿼리와 함께 QuickSight 사용
- C. Amazon Athena를 사용하여 Amazon S3에서 클릭스트림 데이터 쿼리
- D. S3 분석을 사용하여 클릭스트림 데이터 쿼리
- E. 매일 업데이트되는 QuickSight SPICE 엔진 사용

Answer: B,D

QUESTION NO: 5

미국에 본사를 둔 한 운동화 소매 회사가 글로벌 웹사이트를 개설했습니다. 모든 트랜잭션 데이터는 Amazon RDS에 저장되고 큐레이트된 과거 트랜잭션 데이터는 us-east-1 리전의 Amazon Redshift에 저장됩니다. 비즈니스 인텔리전스(BI) 팀은 운동화 트렌드에 대한 대시보드를 제공하여 사용자 경험을 향상하고자 합니다.

BI 팀은 Amazon QuickSight를 사용하여 웹 사이트 대시보드를 렌더링하기로 결정합니다. 개발 중에 일본의 한 팀은 ap-northeast-1에서 Amazon QuickSight를 프로비저닝했습니다. 팀이 Amazon QuickSight를 ap-northeast-1에서 us-east-1의 Amazon Redshift로 연결하는 데 어려움을 겪고 있습니다.

이 문제를 해결하고 요구 사항을 충족하는 솔루션은 무엇입니까?

- A. Amazon Redshift 콘솔에서 교차 리전 스냅샷 구성을 선택하고 대상 리전을 ap-northeast-1로 설정합니다. 스냅샷에서 Amazon Redshift 클러스터를 복원하고 ap-northeast-1에서 시작된 Amazon QuickSight에 연결합니다.
- B. Amazon QuickSight가 Amazon Redshift의 데이터에 액세스할 수 있도록 Amazon QuickSight VPC에서 Amazon Redshift VPC로 VPC 엔드포인트를 생성합니다.
- C. 문자열에 지역 정보가 포함된 Amazon Redshift 엔드포인트 연결 문자열을 생성하고 Amazon QuickSight에서 이 연결 문자열을 사용하여 Amazon Redshift에 연결합니다.
- D. ap-northeast-1의 Amazon QuickSight 서버에 대해 적절한 IP 주소 범위에서 액세스 권한을 부여하는 인바운드 규칙을 사용하여 us-east-1에서 Amazon Redshift에 대한 새 보안 그룹을 생성합니다.

Answer: B

QUESTION NO: 6

온라인 소매 회사는 Amazon Redshift를 사용하여 과거 판매 트랜잭션을 저장합니다. 회사는 PCI DSS(Payment Card Industry Data Security Standard)를 준수하기 위해 클러스터의 미사용 데이터를 암호화해야 합니다. 기업 거버넌스 정책에서는 온프레미스 HSM(하드웨어 보안 모듈)을 사용하여 암호화 키를 관리해야 합니다.

어떤 솔루션이 이러한 요구 사항을 충족합니까?

- A. AWS CloudHSM Classic을 사용하여 암호화 키를 생성하고 관리합니다. 키 관리를 위해 CloudHSM Classic을 사용하는 옵션으로 VPC에서 Amazon Redshift 클러스터를 시작합니다.
- B. VPC를 생성하고 VPC와 온프레미스 네트워크 간에 VPN 연결을 설정합니다. 온프레미스 HSM에 대한 HSM 연결 및 클라이언트 인증서를 생성합니다. 온프레미스 HSM을 사용하여 키를 저장하는 옵션으로 VPC에서 클러스터를 시작합니다.
- C. 온프레미스 HSM에 대한 HSM 연결 및 클라이언트 인증서를 생성합니다. 클러스터를 수정하여 기존의 암호화되지 않은 클러스터에서 HSM 암호화를 활성화합니다. VPN을 사용하여 온프레미스 네트워크에서 Amazon Redshift 클러스터가 상주하는 VPC에 연결합니다.

D. AWS CloudHSM에서 온프레미스 HSM의 복제본을 생성합니다. CloudHSM을 사용하여 키를 저장하는 옵션을 사용하여 VPC에서 클러스터를 시작합니다.

Answer: B

QUESTION NO: 7

한 기계 회사가 센서로부터 데이터를 수집하려고 합니다. 데이터 분석 전문가는 거의 실시간으로 데이터를 집계하고 해당 데이터를 영구 데이터 저장소에 저장하는 솔루션을 구현해야 합니다. 데이터는 중첩된 JSON 형식으로 저장되어야 하며 한 자릿수 밀리초의 지연 시간으로 데이터 저장소에서 쿼리되어야 합니다.

어떤 솔루션이 이러한 요구 사항을 충족합니까?

A. Amazon Kinesis Data Streams를 사용하여 센서로부터 데이터를 수신합니다. Amazon Kinesis Data Analytics를 사용하여 스트림을 읽고, 데이터를 집계하고, 데이터를 AWS Lambda 함수로 보냅니다. Amazon DynamoDB에 데이터를 저장하도록 Lambda 함수를 구성합니다.

B. Amazon Kinesis Data Firehose를 사용하여 센서로부터 데이터를 수신합니다. Amazon Kinesis Data Analytics를 사용하여 데이터를 집계합니다. AWS Lambda 함수를 사용하여 Kinesis Data Analytics에서 데이터를 읽고 Amazon S3에 데이터를 저장합니다.

C. Amazon Kinesis Data Firehose를 사용하여 센서로부터 데이터를 수신합니다. AWS Lambda 함수를 사용하여 캡처 중에 데이터를 집계합니다. Kinesis Data Firehose의 데이터를 Amazon DynamoDB에 저장합니다.

D. Amazon Kinesis Data Firehose를 사용하여 센서로부터 데이터를 수신합니다. AWS Lambda 함수를 사용하여 캡처 중에 데이터를 집계합니다. Amazon S3에 데이터를 저장합니다.

Answer: C

Explanation:

This solution meets the requirements because:

Amazon Kinesis Data Firehose is a fully managed service that can capture, transform, and load streaming data into AWS data stores, such as Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Amazon DynamoDB1. It can receive data from sensors and other sources and deliver it to a destination with near-real time latency.

AWS Lambda is a serverless compute service that can run code in response to events and automatically manage the underlying compute resources2. It can be used to perform custom transformations on the data during capture by Kinesis Data Firehose3. It can aggregate the data according to the desired logic and output format.

Amazon DynamoDB is a fully managed NoSQL database service that supports key-value and document data models4. It can store nested JSON data as document attributes and provide single-digit millisecond latency for queries. It can be used as a persistent data store for the aggregated sensor data.

QUESTION NO: 8

대규모 소매업체가 Amazon S3 데이터 레이크 아키텍처로 성공적으로

마이그레이션했습니다. 이 회사의 마케팅 팀은 Amazon Redshift 및 Amazon QuickSight를 사용하여 데이터를 분석하고 인사이트를 도출 및 시각화합니다. 마케팅 팀이 실행 가능한 최신 정보를 확보할 수 있도록 데이터 분석가는 전날의 테라바이트급 업데이트를 사용하여 Amazon Redshift의 야간 새로 고침을 구현합니다.

첫 번째 야간 새로 고침 후 사용자는 새로 고침 이전에 올바르게 실행되었던 가장 인기 있는 대시보드의 절반이 현재 훨씬 느리게 실행되고 있다고 보고합니다. Amazon CloudWatch는 알림을 표시하지 않습니다.

성능 저하의 가장 가능성 있는 원인은 무엇입니까?

- A. 대시보드는 비효율적인 SQL 쿼리로 인해 어려움을 겪고 있습니다.
- B. 대시보드에서 실행 중인 쿼리에 비해 클러스터 크기가 작습니다.
- C. 야간 데이터 새로 고침으로 인해 지속적인 사용자 워크로드로 인해 Amazon Redshift가 자동으로 달을 수 없는 느린 트랜잭션이 발생합니다.
- D. 야간 데이터 새로 고침으로 인해 지속적인 사용자 워크로드로 인해 Amazon Redshift가 자동으로 수행할 수 없는 진공 작업이 필요한 대시보드 테이블이 남았습니다.

Answer: D

Explanation:

<https://github.com/awsdocs/amazon-redshift-developer-guide/issues/21>

QUESTION NO: 9

회사에서 판매 데이터 대시보드의 데이터 로드 시간을 개선하려고 합니다. 데이터는 .csv 파일로 수집되어 날짜별로 분할된 Amazon S3 버킷에 저장되었습니다. 그런 다음 데이터는 빈번한 분석을 위해 Amazon Redshift 데이터 웨어하우스에 로드됩니다. 데이터 볼륨은 하루 최대 500GB입니다.

데이터 로딩 성능을 향상시키는 솔루션은 무엇입니까?

- A. .csv 파일을 압축하고 INSERT 문을 사용하여 데이터를 Amazon Redshift로 수집합니다.
- B. 대용량 .csv 파일을 분할한 다음 COPY 명령을 사용하여 데이터를 Amazon Redshift로 로드합니다.
- C. Amazon Kinesis Data Firehose를 사용하여 데이터를 Amazon Redshift로 수집합니다.
- D. 정렬되지 않은 키 순서로 .csv 파일을 로드하고 Amazon Redshift에서 테이블을 진공 청소기로 청소합니다.

Answer: B

Explanation:

https://docs.aws.amazon.com/redshift/latest/dg/c_loading-data-best-practices.html

QUESTION NO: 10

한 미디어 회사는 애플리케이션에서 생성된 로그 데이터에 대한 분석을 수행해 왔습니다. 최근 실행 중인 동시 분석 작업 수가 증가했으며 새 작업 수가 증가함에 따라 기존 작업의 전반적인 성능이 감소하고 있습니다. 분할된 데이터는 Amazon S3 One Zone-Infrequent Access(S3 One Zone-IA)에 저장되고 일관된 보기가 활성화된 EMRFS(EMR 파일 시스템)를 사용하여 Amazon EMR 클러스터에서 분석 처리가 수행됩니다. 데이터 분석가는 EMR 작업 노드가 Amazon S3의 객체를 나열하는 데 더 오래 걸린다고 판단했습니다.

Amazon S3에서 로그 데이터에 액세스하는 성능을 가장 높일 수 있는 작업은 무엇입니까?

- A. 해시 함수를 사용하여 임의의 문자열을 생성하고 Amazon S3에 로그 데이터를 저장할 때 개체 접두사 시작 부분에 추가합니다.
- B. 수명 주기 정책을 사용하여 로그 데이터에 대해 S3 스토리지 클래스를 S3 Standard로 변경합니다.
- C. 공유 Amazon DynamoDB 테이블의 읽기 용량 단위(RCU)를 늘립니다.
- D. 느리게 실행되는 EMR 클러스터를 다른 가용 영역에 재배포합니다.

Answer: C

Explanation:

<https://docs.aws.amazon.com/emr/latest/ManagementGuide/emrfs-metadata.html>

QUESTION NO: 11

회사에는 Amazon Kinesis Client Library(KCL)를 사용하여 Kinesis 데이터 스트림에서 레코드를 읽는 애플리케이션이 있습니다.

성공적인 마케팅 캠페인 후 애플리케이션 사용량이 크게 증가했습니다. 결과적으로 데이터 분석가는 데이터 스트림에서 일부 샤드를 분할해야 했습니다. 샤드가 분할되면 애플리케이션에서 산발적으로 ExpiredIteratorExceptions 오류가 발생하기 시작했습니다. 데이터 분석가는 이 문제를 해결하기 위해 무엇을 해야 합니까?

- A. 스트림 레코드를 처리하는 스레드 수를 늘립니다.
- B. 스트림의 Amazon DynamoDB 테이블에 할당된 프로비저닝된 읽기 용량 단위를 늘립니다.
- C. 스트림의 Amazon DynamoDB 테이블에 할당된 프로비저닝된 쓰기 용량 단위를 늘립니다.
- D. 스트림의 Amazon DynamoDB 테이블에 할당된 프로비저닝된 쓰기 용량 단위를 줄입니다.

Answer: C

QUESTION NO: 12

데이터 분석가는 애플리케이션에서 생성된 여러 데이터 세트에서 데이터 시각화를 위해 Amazon QuickSight를 사용하고 있습니다. 각 애플리케이션은 별도의 Amazon S3 버킷 내에 파일을 저장합니다. AWS Glue 데이터 카탈로그는 Amazon S3의 모든 애플리케이션 데이터에서 중앙 카탈로그로 사용됩니다. 새 애플리케이션은 별도의 S3 버킷에 데이터를 저장합니다. 새 애플리케이션 데이터 소스를 포함하도록 카탈로그를 업데이트한 후 데이터 분석가는 Amazon Athena 테이블에서 새 Amazon QuickSight 데이터 소스를 생성했지만 SPICE로 가져오기에 실패했습니다.

데이터 분석가는 이 문제를 어떻게 해결해야 합니까?

- A. Amazon QuickSight 콘솔 내에서 AWS Glue 데이터 카탈로그에 대한 권한을 편집합니다.
- B. Amazon QuickSight 콘솔 내에서 새 S3 버킷에 대한 권한을 편집합니다.
- C. AWS Glue 콘솔 내에서 AWS Glue 데이터 카탈로그에 대한 권한을 편집합니다.
- D. S3 콘솔 내에서 새 S3 버킷에 대한 권한을 편집합니다.

Answer: B

QUESTION NO: 13

비디오 스트리밍 웹사이트를 운영하는 회사는 사용자 행동을 분석하여 사용자에게 실시간으로 추천하고자 합니다. Clickstream 데이터는 Amazon Kinesis Data Streams로 전송되고 참조 데이터는 Amazon S3에 저장됩니다. 회사는 표준 SQL queries를 사용할 수 있는 솔루션을 원합니다. 또한 솔루션은 권장 사항을 제시하는 동안 미리 계산된 참조 데이터를 조회하는 방법을 제공해야 합니다. 어떤 솔루션이 이러한 요구 사항을 충족합니까?

- A. AWS Glue Python 셸 작업을 사용하여 Kinesis Data Streams에서 들어오는 데이터 처리 Boto3 라이브러리를 사용하여 Amazon Redshift에 데이터 쓰기
- B. AWS Glue 스트리밍 및 Scale을 사용하여 Kinesis Data Streams에서 들어오는 데이터 처리 AWS Glue 커넥터를 사용하여 Amazon Redshift에 데이터 쓰기
- C. Amazon Kinesis Data Analytics를 사용하여 참조 데이터를 기반으로 인애플리케이션 테이블 생성 Kinesis Data Streams에서 들어오는 데이터 처리 데이터 스트림을 사용하여

결과를 Amazon Redshift에 기록

D. Amazon Kinesis Data Analytics를 사용하여 참조 데이터를 기반으로 인애플리케이션 테이블 생성 Kinesis Data Streams에서 들어오는 데이터 처리 Amazon Kinesis Data Firehose 전송 스트림을 사용하여 결과를 Amazon Redshift에 쓰기

Answer: D

QUESTION NO: 14

회사에서 데이터 및 분석 플랫폼의 비용을 최적화하려고 합니다. 회사는 여러 가지를 섭취하고 있습니다.

다양한 데이터 소스의 Amazon S3에 있는 .csv 및 JSON 파일. 들어오는 데이터는 매일 50GB가 될 것으로 예상됩니다. 이 회사는 Amazon Athena를 사용하여 Amazon S3의 원시 데이터를 직접 쿼리하고 있습니다. 대부분의 쿼리는 지난 12개월 동안의 데이터를 집계하며 5년보다 오래된 데이터는 드물게 쿼리됩니다. 일반적인 쿼리는 약 500MB의 데이터를 스캔하고 1분 이내에 결과를 반환할 것으로 예상됩니다. 원시 데이터는 규정 준수 요구 사항을 위해 무기한 보관해야 합니다.

회사의 요구 사항을 충족하는 솔루션은 무엇입니까?

- A.** AWS Glue ETL 작업을 사용하여 데이터를 압축, 분할 및 열 데이터 형식으로 변환합니다. Athena를 사용하여 처리된 데이터 세트를 쿼리합니다. 객체 생성 후 5년 후에 처리된 데이터를 Amazon S3 Standard-Infrequent Access(S3 Standard-IA) 스토리지 클래스로 이동하도록 수명 주기 정책을 구성합니다. 객체 생성 후 7일 동안 장기 보관을 위해 원시 데이터를 Amazon S3 Glacier로 이동하도록 두 번째 수명 주기 정책을 구성합니다.
- B.** AWS Glue ETL 작업을 사용하여 데이터를 분할하고 행 기반 데이터 형식으로 변환합니다. Athena를 사용하여 처리된 데이터 세트를 쿼리합니다. 객체 생성 후 5년이 지나면 데이터를 Amazon S3 Standard-Infrequent Access(S3 Standard-IA) 스토리지 클래스로 이동하도록 수명 주기 정책을 구성합니다. 객체 생성 후 7일 동안 장기 보관을 위해 원시 데이터를 Amazon S3 Glacier로 이동하도록 두 번째 수명 주기 정책을 구성합니다.
- C.** AWS Glue ETL 작업을 사용하여 데이터를 압축, 분할 및 열 데이터 형식으로 변환합니다. Athena를 사용하여 처리된 데이터 세트를 쿼리합니다. 객체에 마지막으로 액세스한 지 5년 후에 처리된 데이터를 Amazon S3 Standard-Infrequent Access(S3 Standard-IA) 스토리지 클래스로 이동하도록 수명 주기 정책을 구성합니다. 객체에 마지막으로 액세스한 날짜로부터 7일 후에 장기 보관을 위해 원시 데이터를 Amazon S3 Glacier로 이동하도록 두 번째 수명 주기 정책을 구성합니다.
- D.** AWS Glue ETL 작업을 사용하여 데이터를 분할하고 행 기반 데이터 형식으로 변환합니다. Athena를 사용하여 처리된 데이터 세트를 쿼리합니다. 객체에 마지막으로 액세스한 지 5년 후에 데이터를 Amazon S3 Standard-Infrequent Access(S3 Standard-IA) 스토리지 클래스로 이동하도록 수명 주기 정책을 구성합니다. 객체에 마지막으로 액세스한 날짜로부터 7일 후에 장기 보관을 위해 원시 데이터를 Amazon S3 Glacier로 이동하도록 두 번째 수명 주기 정책을 구성합니다.

Answer: A

QUESTION NO: 15

한 전자상거래 회사는 Amazon Aurora PostgreSQL을 사용하여 실시간 트랜잭션 데이터를 처리 및 저장하고 데이터 웨어하우스 솔루션으로 Amazon Redshift를 사용합니다.

PostgreSQL 데이터베이스의 새 데이터로 Redshift 클러스터를 업데이트하기 위해 야간 ETL 작업이 구현되었습니다. 비즈니스가 빠르게 성장함에 따라 Redshift 클러스터의 규모와

비용도 커졌습니다. 회사의 데이터 분석 팀은 비용 절감을 위해 기록 데이터를 보관하고 가장 최근 12개월의 데이터만 Amazon Redshift에 보관하는 솔루션을 만들어야 합니다. 또한 데이터 분석가는 PostgreSQL의 실시간 트랜잭션 데이터, Redshift의 현재 데이터 및 보관된 기록 데이터의 데이터를 효과적으로 결합하는 분석 쿼리를 실행할 수 있어야 합니다. 이러한 요구 사항을 충족하는 작업 조합은 무엇입니까? (3개를 선택하세요.)

- A. PostgreSQL 데이터베이스에서 실시간 트랜잭션 데이터를 쿼리하도록 Amazon Redshift Federated Query 기능을 구성합니다.
- B. PostgreSQL 데이터베이스에서 실시간 트랜잭션 데이터를 쿼리하도록 Amazon Redshift Spectrum을 구성합니다.
- C. UNLOAD 명령을 사용하여 12개월이 지난 데이터를 Amazon S3에 복사하도록 월별 작업을 예약한 다음 Redshift 클러스터에서 해당 데이터를 삭제합니다. Amazon S3의 기록 데이터에 액세스하도록 Amazon Redshift Spectrum을 구성합니다.
- D. UNLOAD 명령을 사용하여 12개월이 지난 데이터를 Amazon S3 Glacier 유연한 검색으로 복사하도록 월간 작업을 예약한 다음 Redshift 클러스터에서 해당 데이터를 삭제합니다. S3 Glacier 유연한 검색을 통해 기록 데이터에 액세스하도록 Redshift Spectrum을 구성합니다.
- E. Amazon Redshift에서 다양한 소스의 실시간, 현재 및 기록 데이터를 결합하는 후기 바인딩 보기를 생성합니다.
- F. Amazon Redshift에서 다양한 소스의 실시간, 현재 및 기록 데이터를 결합하는 구체화된 뷰를 생성합니다.

Answer: A,C,E

QUESTION NO: 16

회사는 Amazon Elasticsearch Service(Amazon ES)를 사용하여 웹사이트 클릭스트림 데이터를 저장하고 분석합니다. 이 회사는 Amazon Kinesis Data Firehose를 사용하여 매일 1TB의 데이터를 수집하고 하루 분량의 데이터를 Amazon ES 클러스터에 저장합니다. 이 회사는 Amazon ES 인덱스에 대한 쿼리 성능이 매우 느리고 인덱스에 쓰기를 시도할 때 때때로 Kinesis Data Firehose에서 오류가 표시됩니다. Amazon ES 클러스터에는 단일 인덱스를 실행하는 10개의 노드와 3개의 전용 마스터 노드가 있습니다. 각 데이터 노드에는 1.5TB의 Amazon EBS 스토리지가 연결되어 있고 클러스터는 1,000개의 샤드로 구성됩니다. 경우에 따라 클러스터 로그에서 JVMMemoryPressure 오류가 발견됩니다.

Amazon ES의 성능을 향상시키는 솔루션은 무엇입니까?

- A. Amazon ES 마스터 노드의 메모리를 늘립니다.
- B. Amazon ES 데이터 노드의 수를 줄입니다.
- C. 인덱스의 Amazon ES 샤드 수를 줄입니다.
- D. 인덱스에 대한 Amazon ES 샤드 수를 늘립니다.

Answer: C

Explanation:

<https://aws.amazon.com/premiumsupport/knowledge-center/high-jvm-memory-pressure-elasticsearch/>

QUESTION NO: 17

전자상거래 회사는 Amazon RDS에 고객 구매 데이터를 저장합니다. 회사는 기록 데이터를 저장하고 분석할 수 있는 솔루션을 원합니다. 가장 최근 6개월의 데이터는 분석 워크로드에 대해 자주 쿼리됩니다. 이 데이터의 크기는 수 테라바이트입니다. 한 달에 한 번 지난 5년

동안의 기록 데이터에 액세스할 수 있어야 하며 최신 데이터와 결합됩니다. 회사는 성능과 비용을 최적화하려고 합니다.

이러한 요구 사항을 충족하는 스토리지 솔루션은 무엇입니까?

- A. RDS 데이터베이스의 읽기 전용 복제본을 생성하여 가장 최근 6개월 간의 데이터를 저장합니다. 기록 데이터를 Amazon S3에 복사합니다. Amazon S3 및 Amazon RDS에서 데이터의 AWS Glue 데이터 카탈로그를 생성합니다. Amazon Athena를 사용하여 기록 쿼리를 실행합니다.
- B. ETL 도구를 사용하여 가장 최근 6개월의 데이터를 Amazon Redshift 클러스터로 점진적으로 로드합니다. 이 클러스터에 대해 더 자주 쿼리를 실행합니다. 기록 데이터에 대한 쿼리를 실행하기 위해 RDS 데이터베이스의 읽기 복제본을 생성합니다.
- C. Amazon RDS에서 Amazon S3로 데이터를 점진적으로 복사합니다. Amazon S3에서 데이터의 AWS Glue 데이터 카탈로그를 생성합니다. Amazon Athena를 사용하여 데이터를 쿼리합니다.
- D. Amazon RDS에서 Amazon S3로 데이터를 점진적으로 복사합니다. Amazon Redshift에서 최근 6개월 데이터를 로드하고 저장합니다. 모든 기록 데이터에 연결하도록 Amazon Redshift Spectrum 테이블을 구성합니다.

Answer: D

Explanation:

Section: (none)

QUESTION NO: 18

마케팅 회사는 워크로드에 Amazon EMR 클러스터를 사용하고 있습니다. 회사는 마스터 노드에 로그인하여 클러스터에 타사 라이브러리를 수동으로 설치합니다. 데이터 분석가는 수동 프로세스를 대체할 자동화된 솔루션을 만들어야 합니다.

이러한 요구 사항을 충족할 수 있는 옵션은 무엇입니까? (두 가지를 선택하세요.)

- A. 필요한 설치 스크립트를 Amazon S3에 배치하고 사용자 지정 부트스트랩 작업을 사용하여 실행합니다.
- B. 필요한 설치 스크립트를 Amazon S3에 배치하고 Amazon EMR의 Apache Spark를 통해 실행합니다.
- C. 기존 EMR 마스터 노드에 필요한 타사 라이브러리를 설치합니다. 해당 마스터 노드에서 AMI를 생성하고 해당 사용자 지정 AMI를 사용하여 EMR 클러스터를 다시 생성합니다.
- D. Amazon DynamoDB 테이블을 사용하여 필요한 애플리케이션 목록을 저장합니다. DynamoDB Streams로 AWS Lambda 함수를 트리거하여 소프트웨어를 설치합니다.
- E. Amazon Linux로 Amazon EC2 인스턴스를 시작하고 필요한 타사 라이브러리를 인스턴스에 설치합니다. AMI를 생성하고 해당 AMI를 사용하여 EMR 클러스터를 생성합니다.

Answer: A,E

Explanation:

<https://aws.amazon.com/about-aws/whats-new/2017/07/amazon-emr-now-supports-launching-clusters-with-custom-amazon-linux-amis/>

https://docs.aws.amazon.com/de_de/emr/latest/ManagementGuide/emr-plan-bootstrap.html

QUESTION NO: 19

회사는 Amazon Kinesis SDK를 사용하여 Kinesis Data Streams에 데이터를 씁니다. 규정 준수 요구 사항에 따르면 회전할 수 있는 키를 사용하여 미사용 데이터를 암호화해야 합니다.

회사는 최소한의 코딩 작업으로 이 암호화 요구 사항을 충족하고자 합니다.

이러한 요구 사항을 어떻게 충족할 수 있습니까?

- A. AWS KMS에서 고객 마스터 키(CMK)를 생성합니다. CMK에 별칭을 할당합니다. AWS 암호화 SDK를 사용하여 키 별칭을 제공하여 데이터를 암호화하고 해독합니다.
- B. AWS KMS에서 고객 마스터 키(CMK)를 생성합니다. CMK에 별칭을 할당합니다. CMK 별칭을 KMS 마스터 키로 사용하여 Kinesis 데이터 스트림에서 서버 측 암호화를 활성화합니다.
- C. AWS KMS에서 고객 마스터 키(CMK)를 생성합니다. 데이터를 암호화하고 해독하는 AWS Lambda 함수를 생성합니다. 함수의 환경 변수에서 KMS 키 ID를 설정합니다.
- D. Kinesis Data용 기본 KMS 키를 사용하여 Kinesis 데이터 스트림에서 서버 측 암호화 활성화

Answer: B

Streams.

QUESTION NO: 20

공유 작업 공간 회사의 데이터 엔지니어링 팀은 공간 예약 시스템에서 생성된 모든 웹로그에 대한 중앙 집중식 로깅 시스템을 구축하려고 합니다. 이 회사는 웹 사이트에서 공유 공간 예약 요청을 처리하는 Amazon EC2 인스턴스 플릿을 보유하고 있습니다. 데이터 엔지니어링 팀은 거의 실시간 검색 엔진을 제공할 서비스로 모든 웹로그를 수집하려고 합니다. 팀은 로깅 시스템의 유지 관리 및 운영을 관리하기를 원하지 않습니다.

데이터 엔지니어링 팀이 AWS 내에서 웹 로깅 시스템을 효율적으로 설정할 수 있는 솔루션은 무엇입니까?

- A. 웹로그를 CloudWatch 로그로 스트리밍하고 Amazon Kinesis 데이터 스트림을 CloudWatch에 구독하도록 Amazon CloudWatch 에이전트를 설정합니다. 웹로그의 최종 목적지로 Amazon Elasticsearch Service를 선택합니다.
- B. 웹로그를 CloudWatch 로그로 스트리밍하고 Amazon Kinesis Data Firehose 전송 스트림을 CloudWatch로 구독하도록 Amazon CloudWatch 에이전트를 설정합니다. 웹로그의 최종 목적지로 Amazon Elasticsearch Service를 선택합니다.
- C. 웹로그를 CloudWatch 로그로 스트리밍하고 Amazon Kinesis 데이터 스트림을 CloudWatch에 구독하도록 Amazon CloudWatch 에이전트를 설정합니다. Splunk를 웹로그의 최종 대상으로 구성합니다.
- D. 웹로그를 CloudWatch 로그로 스트리밍하고 Amazon Kinesis Firehose 전송 스트림을 CloudWatch로 구독하도록 Amazon CloudWatch 에이전트를 설정합니다. 웹로그의 최종 목적지로 Amazon DynamoDB를 구성합니다.

Answer: B

Explanation:

https://docs.aws.amazon.com/AmazonCloudWatch/latest/logs/CWL_ES_Stream.html

QUESTION NO: 21

제조 회사는 Amazon S3를 사용하여 데이터를 저장합니다.

1. 이 회사는 AWS Lake Formation을 사용하여 이러한 데이터 자산에 대한 세분화된 수준의 보안을 제공하려고 합니다. 데이터는 Apache Parquet 형식입니다. 회사는 컨설턴트가 데이터 레이크를 구축할 마감일을 정했습니다.

컨설턴트는 이러한 요구 사항을 충족하는 가장 비용 효율적인 솔루션을 어떻게 만들어야

합니까?

- A. Lake Formation 청사진을 실행하여 데이터를 Lake Formation으로 이동합니다. Lake Formation에 데이터가 있으면 Lake Formation에 대한 권한을 적용합니다.
- B. 데이터 카탈로그를 생성하려면 기존 Parquet 데이터에서 AWS Glue 크롤러를 실행합니다. Amazon S3 경로를 등록한 다음 Lake Formation을 통해 권한을 적용하여 세분화된 수준의 보안을 제공합니다.
- C. Amazon EC2 인스턴스에 Apache Ranger를 설치하고 Amazon EMR과 통합합니다. Ranger 정책을 사용하여 Amazon S3의 기존 데이터 자산에 대한 역할 기반 액세스 제어를 생성합니다.
- D. 다양한 사용자 및 그룹에 대해 여러 IAM 역할을 생성합니다. Amazon S3의 다양한 데이터 자산에 IAM 역할을 할당하여 테이블 기반 및 열 기반 액세스 제어를 생성합니다.

Answer: A

Explanation:

<https://aws.amazon.com/blogs/big-data/building-securing-and-managing-data-lakes-with-aws-lake-formation/>