



아키텍처 다이어그램

서버리스 분석을 통한 추적성



서버리스 분석을 통한 추적성: 아키텍처 다이어그램

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon의 상표 및 트레이드 드레스는 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계와 관계없이 해당 소유자의 자산입니다.

Table of Contents

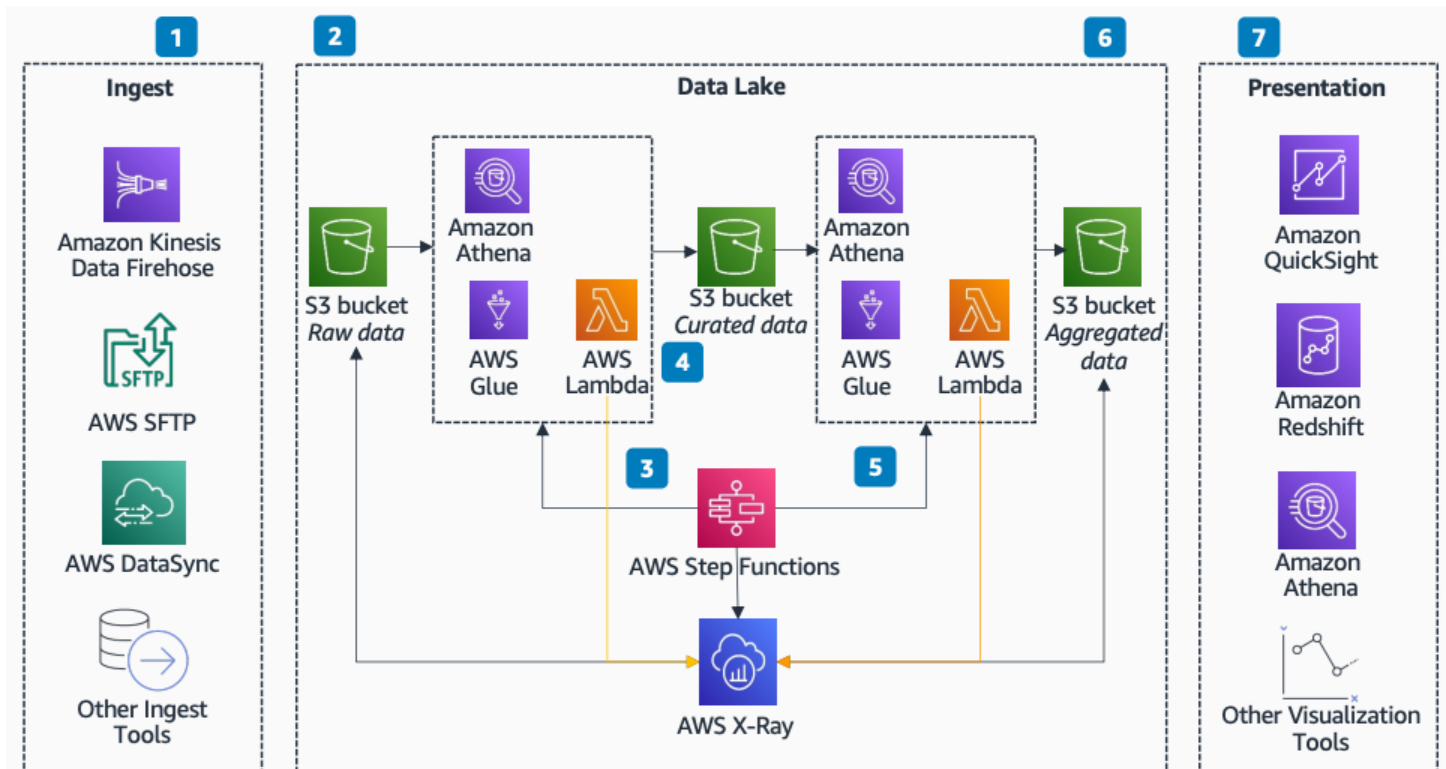
Data Lake 수집	1
AWS X-Ray를 사용하여 Data Lake 수집 및 처리 추적	1
참조 자료	2
다이어그램 기록	2
다운스트림 분석	4
다운스트림 분석을 사용한 Data Lake 수집 및 처리	4
참조 자료	5
다이어그램 기록	5
.....	vi

Data Lake 수집 및 처리 추적

게시일: 2021년 7월 15일([다이어그램 기록](#))

이 아키텍처는 AWS X-Ray를 사용하여 데이터 레이크 수집 및 처리를 추적하는 방법을 보여줍니다. 흐름의 주요 체크포인트에서 트레이스를 내보내 데이터 레이크 처리 수명 주기 문제를 진단하고 성능 병목 현상을 식별할 수 있습니다.

AWS X-Ray를 사용하여 Data Lake 수집 및 처리 추적



다음 단계에서는 아키텍처를 설명합니다.

1. 다양한 생산자가 [Amazon Simple Storage Service](#)에 구축된 데이터 레이크로 데이터를 전송합니다.
2. 데이터는 먼저 원시 데이터를 위해 Amazon S3 버킷에 저장됩니다. AWS X-Ray가 Amazon S3 이벤트 알림에 대한 추적 메시지를 활성화하므로 이는 첫 번째 추적입니다.
3. 데이터 레이크 큐레이션 블록은 분석을 위해 데이터를 준비합니다. 이러한 처리 블록을 [AWS Step Functions](#) 트리거합니다. AWS X-Ray는 Step Functions와 통합되어 성능 병목 현상을 식별하고 오류로 이어진 요청 문제를 해결하는 데 도움이 됩니다.

4. AWS X-Ray는 [AWS Lambda](#) 함수를 추적할 수 있습니다. Lambda는 X-Ray 데몬을 실행하고 함수 호출 및 결과에 대한 세부 정보가 포함된 세그먼트를 기록합니다. 추가 계측을 위해 X-Ray SDK를 함수와 함께 번들링하여 발신 호출을 기록하고 주식 및 메타데이터를 추가합니다.
5. 데이터 큐레이션은 다음 큐레이션 블록을 트리거하는 Step Functions 상태 머신으로 계속됩니다. 첫 번째 블록과 마찬가지로 X-Ray는 Step Functions 및 Lambda 함수를 캡처할 수 있습니다.
6. 데이터는 분석을 위해 준비된 집계된 버킷에 저장됩니다. 이는 데이터 레이크 수집 흐름의 마지막 트레이스입니다. 추적을 모든 키 체크포인트에서 내보내면 데이터 레이크 처리 수명 주기 문제를 진단할 수 있습니다.
7. [Amazon Quick Sight](#), Amazon Redshift, [Amazon Athena](#)와 같은 다양한 분석 도구를 사용하면 Amazon S3에 구축된 데이터 레이크의 데이터를 시각화할 수 있습니다.

참조 자료

자세한 내용은 다음 리소스를 참조하세요.

- [AWS 아키텍처 아이콘](#)
- [AWS 아키텍처 센터](#)
- [AWS Well-Architected](#)
- [빠른 버전](#)

다이어그램 기록

이 참조 아키텍처 다이어그램의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독하세요.

변경 사항	설명	날짜
최초 게시	참조 아키텍처 다이어그램이 처음 게시되었습니다.	2021년 7월 15일
최초 게시	참조 아키텍처 다이어그램이 처음 게시되었습니다.	2021년 7월 15일

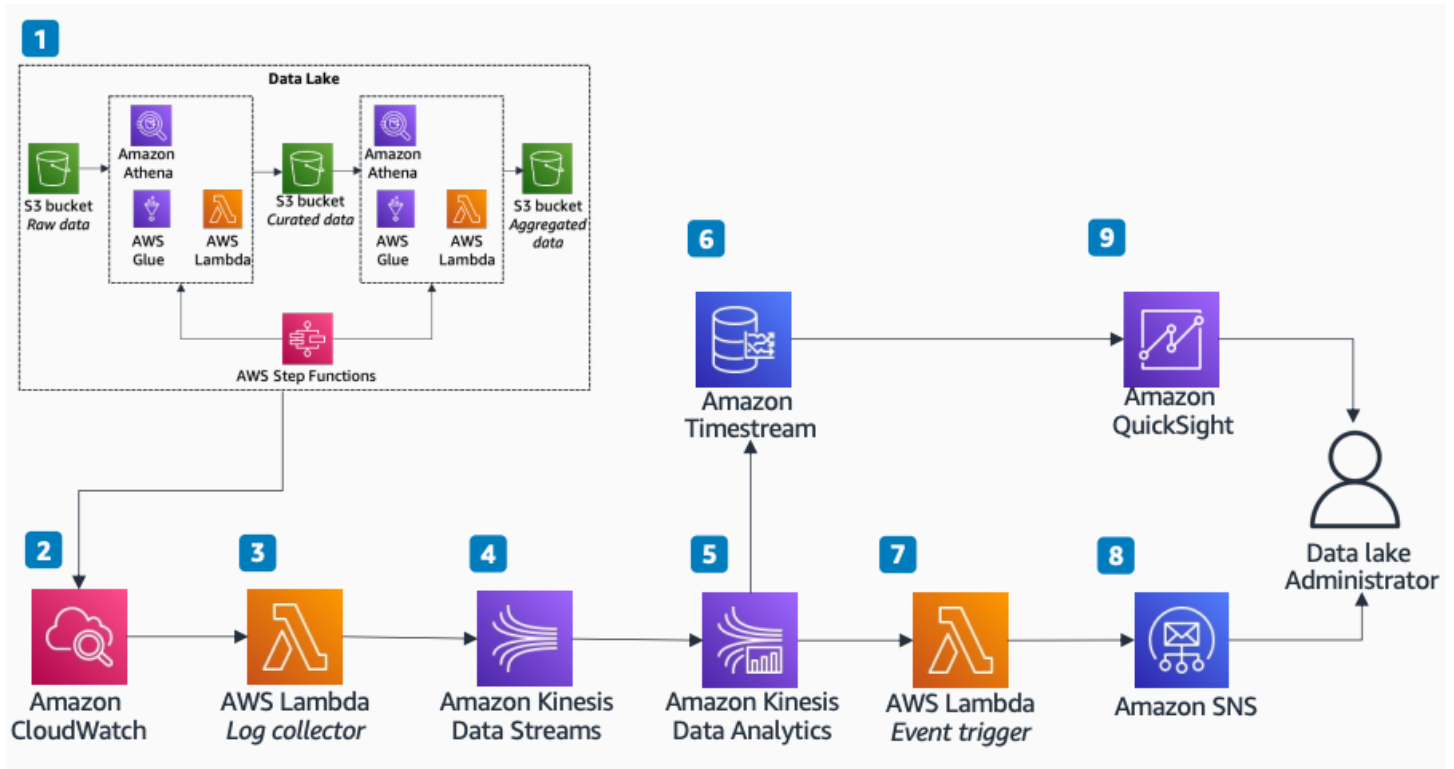
 Note

RSS 업데이트를 구독하려면 사용 중인 브라우저에 대해 RSS 플러그인이 활성화되어 있어야 합니다.

다운스트림 분석을 사용한 Data Lake 수집 및 처리

이 아키텍처는 다운스트림 분석을 사용하여 데이터 레이크 내의 데이터 처리 파이프라인을 모니터링하는 방법을 보여줍니다. 관련 로그, 지표 및 이벤트를 게시하여 비정상적인 동작을 감지하고 추세를 거의 실시간으로 시각화할 수 있습니다.

다운스트림 분석을 사용한 Data Lake 수집 및 처리



다음 단계에서는 아키텍처를 설명합니다.

1. 데이터 레이크 내의 데이터 처리 파이프라인은 데이터를 생성합니다. 이 단계에서는 원하는 수의 소스 파이프라인을 생성할 수 있습니다.
2. 데이터 파이프라인 구성 요소는 관련 로그, 지표 또는 이벤트를 [Amazon CloudWatch](#)에 게시합니다.
3. [AWS Lambda](#) 함수는 CloudWatch에서 데이터를 가져오고 사전 처리된 데이터를 [Amazon Kinesis Data Streams](#)로 푸시합니다.
4. Kinesis 데이터 스트림은 거의 실시간에 가까운 스트리밍 데이터를 위한 중앙 스토리지입니다.
5. Amazon Kinesis Data Analytics는 거의 실시간으로 데이터를 분석하고 추가 분석을 위해 집계된 지표를 Amazon Timestream 데이터베이스에 저장합니다. Kinesis Data Analytics 쿼리는 파이프라인에서 비정상적인 데이터 동작을 감지할 수도 있습니다.

6. Amazon Timestream은 시계열 분석을 지원하고 시계열을 기본 데이터 유형으로 정의합니다. 고급 집계, 창 함수, 배열 및 행과 같은 복잡한 데이터 유형을 지원합니다.
7. Amazon Kinesis Data Analytics 임계값 쿼리는 Lambda에 실시간에 가까운 분석을 전송합니다.
8. Lambda 함수는 Kinesis Data Analytics에서 처리된 데이터를 기반으로 [Amazon Simple Notification Service](#)를 사용하여 실시간 알림을 보낼 수 있습니다.
9. [Amazon Quick Sight](#) 대시보드는 성공한 파이프라인 수, 오류율, 오류 유형 분포 등과 같은 추세를 시각화합니다.

참조 자료

자세한 내용은 다음 리소스를 참조하세요.

- [AWS 아키텍처 아이콘](#)
- [AWS 아키텍처 센터](#)
- [AWS Well-Architected](#)

다이어그램 기록

이 참조 아키텍처 다이어그램의 업데이트에 대한 알림을 받으려면 RSS 피드를 구독하세요.

변경 사항	설명	날짜
최초 게시	참조 아키텍처 다이어그램이 처음 게시되었습니다.	2021년 7월 15일
최초 게시	참조 아키텍처 다이어그램이 처음 게시되었습니다.	2021년 7월 15일

Note

RSS 업데이트를 구독하려면 사용 중인 브라우저에 대해 RSS 플러그인이 활성화되어 있어야 합니다.

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.