생성형 AI 보안 제어

Information Security for Workload Factory

NetApp
September 16, 2026

목차

생성형 AI 보안 제어

NetApp Workload Factory의 AI 제어 기능에 대해 알아보십시오.

NetApp Workload Factory는 모델 액세스, 데이터 범위 및 런타임 동작에 대한 보안 제어를 강화하면서 진단 속도를 높이는 생성형 AI 기능을 포함합니다.

Workload Factory는 기본적으로 AI 진단을 활성화합니다. Workload Factory **Administration** 설정에서 언제든지 생성형 AI 기능을 비활성화할 수 있습니다. "GenAI 기능 관리에 대해 자세히 알아보십시오."

보안 범위 AI 기능 개요

Workload Factory는 현재 다음과 같은 기능에 생성형 AI를 적용하고 있습니다.

- Ask Me 어시스턴트(엄선된 NetApp 문서와 함께 RAG를 사용하여 소스 기반 응답 제공)
- 지연 시간 분석
- EMS 이벤트 분석
- 오류 로그 분석

AI 진단을 위한 AWS Bedrock 개요

Workload Factory는 AI 추론을 위해 Amazon Bedrock을 사용합니다. 추론은 구성된 자격 증명 및 추론 프로필을 사용하여 AWS 계정에서 실행됩니다.

관련 정보

- "Amazon Bedrock에서 NetApp Workload Factory AI 진단 옵트인"
- "NetApp Workload Factory용 AI 툴 경계"
- "NetApp Workload Factory용 AI 모델 파라미터 및 청구"

Amazon Bedrock에서 NetApp Workload Factory AI 진단 옵트인

NetApp Workload Factory의 AI 진단은 Amazon Bedrock을 사용하여 이벤트를 분석하며, 이 기능은 기본적으로 활성화되어 있습니다. Workload Factory에서 선택한 Bedrock 모델을 호출하고 진단 데이터를 수집하려면 IAM 권한, 추론 프로필, 그리고 SSH 기능을 갖춘 연결된 Lambda 링크를 구성해야 합니다. 이러한 구성 요소 중 하나라도 누락되었거나 Bedrock 구성이 제거되면 AI 진단을 사용할 수 없게 됩니다.

AI 진단은 기본적으로 활성화되어 있습니다.

시작하기 전에

- 계정에 AWS Bedrock 추론 프로필이 포함되어 있는지 확인하십시오.

- 진단 데이터 수집을 위해 SSH 기능이 있는 연결된 Lambda 링크를 보유합니다.

단계

1. IAM 역할에서 Bedrock 권한을 부여합니다.
   - `bedrock:InvokeModel`
   - `bedrock:InvokeModelWithResponseStream`
   - `bedrock:ListInferenceProfiles`
   - `bedrock:GetInferenceProfile`
2. Workload Factory 설정에서 추론 프로파일 ARN을 구성하십시오.

필수 구성 요소 중 하나라도 누락된 경우 시스템은 누락된 구성 요소를 구체적으로 보고하고 AI 기능은 비활성화된 상태로 유지됩니다.

AI 진단 비활성화

AI 진단을 비활성화하려면:

- IAM 역할에서 Bedrock 권한을 제거하거나
- Workload Factory 설정에서 Bedrock 구성을 삭제합니다.

AI 진단을 즉시 사용할 수 없게 되며, 잔여 데이터 처리는 발생하지 않습니다.

지역 및 지역 간 추론 프로필

Bedrock 추론은 추론 프로필에 지정된 AWS 리전에서 실행됩니다. 교차 리전 추론 프로필을 사용하는 경우, AWS는 표준 AWS Bedrock 동작에 따라 프로필 구성에 있는 모든 리전으로 요청을 라우팅할 수 있습니다. 데이터는 AWS 인프라를 벗어나지 않습니다.

NetApp Workload Factory용 AI 툴 경계

NetApp Workload Factory는 AI 기능이 AWS 및 ONTAP 환경과 상호 작용하는 방식에 엄격한 제한을 두므로 의도치 않은 변경 위험 없이 AI 기반 진단을 신뢰할 수 있습니다. Amazon Bedrock은 변경 명령을 차단하고 크기 및 실행 시간으로 출력을 제한하는 읽기 전용 AWS CLI 및 ONTAP CLI 진단 툴을 사용하며, 사용되는 모든 데이터는 일시적이며 NetApp Workload Factory 또는 AWS에 보관되지 않습니다.

툴 안전 제어

툴	기능	안전 제어 장치
AWS CLI (읽기 전용)	읽기 전용 AWS CLI 명령(describe, list, get)을 실행합니다.	생성, 삭제, 수정, 업데이트, 추가, 제거 등 모든 변경 동사를 차단합니다. 단계당 최대 10개의 명령만 허용됩니다. 출력 크기는 20KB에서 잘립니다.
ONTAP CLI (읽기 전용)	SSH를 사용하여 읽기 전용 ONTAP CLI 명령을 실행합니다.	변경 동사(delete, destroy, create, modify, move)를 모두 차단합니다. 출력 내용이 잘립니다.

에이전트는 리소스를 수정, 생성 또는 삭제할 수 없습니다. 에이전트는 관찰 및 진단만 가능합니다.

데이터 보존 및 모델 훈련 경계

AWS 정책에 따라 Amazon Bedrock은 기본 모델을 학습하거나 개선하기 위해 사용자의 입력 및 출력을 저장하거나 사용하지 않습니다. 온디맨드 추론(Workload Factory에서 사용)의 경우, AWS는 데이터를 일시적으로 처리하며 응답을 반환한 후에는 데이터를 보관하지 않습니다.

NetApp Workload Factory용 AI 모델 파라미터 및 청구

NetApp Workload Factory는 AI 거버넌스와 관련된 모델 구성, 사용량 및 청구 범위를 정의합니다. 구성된 모델은 결정론적 매개변수를 사용하고 일관된 분석 결과를 지원하기 위해 구조화된 JSON을 반환합니다. AWS는 Amazon Bedrock 추론에 대한 비용을 사용자 계정으로 청구하며, Workload Factory는 AI 사용량을 추적하고 계정별 월별 비용을 투명하게 제공합니다. 이 섹션에서는 모델 매개변수와 AI 사용량에 적용되는 제한 사항에 대해 설명합니다.

모델 및 매개변수

매개변수	가치
모델	Anthropic Claude(교차 지역 추론 프로필 사용)
온도	0 (결정론적, 무작위성 없음)
최대 출력 토큰	2,048
최대 추론 단계	분석당 15개
출력 형식	구조화된 JSON (요약, 심각도, 근본 원인, 시정 조치)

청구 및 사용 한도

- AWS는 Bedrock 추론에 대한 비용을 사용자 계정(추론 프로필, 자격 증명)으로 청구합니다.
- Workload Factory는 AI 사용량을 내부적으로 추적합니다.
- Workload Factory는 계정별 월별 비용 현황을 제공합니다.

  "NetApp Workload Factory에서 스토리지 비용을 추적합니다"

Ask Me AI 어시스턴트

NetApp Workload Factory용 Ask Me 보안 아키텍처

Ask Me는 NetApp Workload Factory 내에 내장된 생성형 AI 기반 어시스턴트입니다. Amazon FSx for NetApp ONTAP 및 Workload Factory 주제에 대한 실시간 대화형 지원을 제공합니다. 답변은 검증된 NetApp 문서를 기반으로 하며, Ask Me는 퍼블릭 인터넷에 액세스하거나 모델을 학습시키는 데 고객 데이터를 사용할 수 없습니다. 다중 보안 레이어는 악성 쿼리를 차단하고, 지원되는 도메인으로 답변을 제한하며, 사용자 입력이 시스템 지침을 재정의하지 못하도록 방지합니다. Ask Me가 툴을 사용할 때 권한 부여 경계, 읽기 전용 쿼리 적용, 임시 샌드박스 및

감사 로깅은 실행을 제한하고 모니터링하는 데 도움이 됩니다.

Workload Factory 관리 설정에서 언제든지 Ask Me를 옵트아웃할 수 있으며, 이렇게 하면 사용자 계정에 대한 어시스턴트가 비활성화됩니다. "GenAI 기능 관리에 대해 자세히 알아보십시오."

Ask Me의 **RAG** 아키텍처

Ask Me는 검색 증강 생성(RAG)을 사용합니다.

- 엄선된 지식 기반: 답변은 검증되고 게시된 NetApp 문서 모음을 기반으로 합니다.
- 검색 점수화 및 필터링: 충분히 관련성이 높은 콘텐츠만 모델 컨텍스트에 포함됩니다.
- 인터넷 접속 불가: 모델이 외부 콘텐츠를 가져오거나, 검색하거나, 스크래핑할 수 없습니다.
- 출처 표기: 답변에는 사용된 출처 문서에 대한 참조가 포함됩니다.

다중 계층 프롬프트 보안

- 1단계: 보안 분류기(LLM-as-a-judge)는 악의적인 쿼리(프롬프트 주입, 권한 상승, 데이터 유출, 소셜 엔지니어링, 정책 위반)를 차단합니다.

  시스템은 차단된 쿼리를 기록하지만, 생성 모델은 이를 전혀 인식하지 못합니다.

- 2단계: 시스템 프롬프트 강화는 사용자 입력이 시스템 명령을 덮어쓰는 것을 방지합니다.
- 레이어 3: 도메인 범위 지정으로 인해 FSx ONTAP 및 Workload Factory 주제에 대한 응답이 제한됩니다.
- 레이어 4: 툴 사용 거버넌스는 강화된 샌드박스에서 매개변수를 검증하고 쿼리를 실행합니다. 모델은 임의의 작업을 실행할 수 없습니다.

모델 데이터 격리 및 개인 정보 보호

- 고객 데이터를 사용한 모델 학습 없음
- 요청 수준 격리(테넌트 간 데이터 유출 방지)
- 제한된 기록을 가진 세션 범위의 대화 컨텍스트
- 영구적인 모델 메모리 없음
- 관리형 추론 인프라

툴 사용 및 에이전트 보안

- 권한 범위 기반 툴(사용자 권한 경계 적용)
- 엄격한 툴 실행 시간 제한
- 검색된 데이터를 위한 임시 세션 범위 샌드박스
- 허용 목록을 사용한 읽기 전용 쿼리 강제 적용
- 민감한 매개변수 스크러빙을 통한 툴 호출 감사 기능

NetApp Workload Factory용 Ask Me 액세스 제어

NetApp Workload Factory의 Ask Me 액세스 보안은 인증된 세션, 사용자 수준의 책임, 런타임 작업 중 보호된 비밀 키 처리에 중점을 둡니다. 인증 제어는 각 세션을 검증하고 상호 작용을 특정 사용자와 연결합니다. 민감한 자격 증명은 런타임에 관리되는 비밀 키 저장소에서 검색되고 로그에서 삭제됩니다. 또한 이 서비스는 루트 권한이 없는 컨테이너 실행 및 제한된 CORS 허용 목록을 포함한 인프라 제어를 적용합니다.

인증

- 대상, 발급자 및 RS256 등 알고리즘 검사를 포함한 암호화 검증을 통한 서명된 JWT 인증
- 서비스 경계에서 미들웨어를 이용한 인증
- 사용자 ID 범위를 지정하여 상호 작용이 특정 사용자에게 귀속되도록 합니다.

자격 증명 보호 및 기밀 처리

- 안전한 저장소: 서비스에서 사용하는 민감한 자격 증명은 관리형 비밀 저장소에 저장되며 런타임에 검색됩니다. 비밀 정보는 애플리케이션 코드, 구성 파일 또는 컨테이너 이미지에 저장되지 않습니다.
- 로그 정제: 민감한 값(자격 증명, 토큰, 암호, 액세스 키, 인증서)은 로그에 기록되기 전에 삭제됩니다.

인프라 보안

- 루트 권한이 없는 컨테이너 실행
- CORS는 신뢰할 수 있는 NetApp 도메인의 명시적인 허용 목록으로 제한됩니다.

NetApp Workload Factory의 Ask Me 실행 모드

NetApp Workload Factory는 Ask Me에 대해 서로 다른 보안 및 개인 정보 보호 특성을 가진 여러 실행 모드를 제공합니다. Ask Me가 툴 통합을 사용하는 경우(예: 고객의 Workload Factory 리소스를 쿼리하는 경우), 툴 실행은 계층화된 보안 조치를 통해 제어됩니다. 툴을 사용한 실행은 인증된 사용자의 권한 범위 내에서 작동하며 각 작업의 지속 시간과 범위에 제한이 적용됩니다. 검색된 데이터는 외부 액세스가 차단되는 임시 세션 샌드박스에 저장되며 세션이 종료되면 삭제됩니다. 독립적인 프롬프트 및 코드 제어를 통해 읽기 전용 쿼리가 적용되며, 감사 기록은 민감한 값이 삭제된 상태로 툴 활동을 기록합니다.

툴 실행 보호 장치

- 권한 범위 툴은 인증된 사용자의 권한 내에서 작동합니다.
- 툴 실행 시간 초과는 장시간 실행되는 작업을 제한합니다.
- 임시 데이터 샌드박스는 툴로 검색된 데이터를 세션 범위 내에 저장하고, 외부 액세스, 파일 I/O, 네트워크 호출 및 엔진 수준에서의 확장 프로그램 로딩을 차단하며, 세션이 종료되면 소멸됩니다.
- 읽기 전용 쿼리 강제 적용은 엄격한 허용 목록을 통해 생성된 쿼리의 유효성을 검사합니다.
- 프롬프트 적용 보안과 코드 적용 보안은 독립적으로 적용됩니다.
- 툴 호출 감사 기능은 툴 이름, 매개변수, 타임스탬프 및 결과 상태를 로깅하며, 민감한 값은 스크러빙됩니다.

출력 제어

- 토큰 제한 시행
- 분류/보안 중요 작업에 대한 결정론적 출력(온도 0)
- 조기 종료 및 정리 기능을 갖춘 스트리밍

NetApp Workload Factory용 Ask Me 개인 정보 보호 및 가용성

NetApp Workload Factory의 Ask Me 개인정보 보호 제어는 데이터 노출을 제한하고, 세션 처리를 격리하며, 사용자 상호 작용에 대한 개인정보 보호 경계를 유지합니다. 사용자의 입력은 관리형 AI 서비스에서 처리되며, 이 서비스는 해당 입력을 기본 모델 학습이나 개선에 사용하지 않습니다. 검색된 운영 데이터는 세션 기간 동안 임시 메모리 저장소에 보관되며 영구 저장되거나 공유되지 않습니다. 여러 클라우드 지역에 걸쳐 있는 멀티 모델 폴백 체인은 기본 모델의 성능 저하 또는 사용 불가 시에도 가용성을 보장하며, 폴백 모델에도 동일한 보안 제어가 적용됩니다.

데이터 처리 및 개인정보 보호

- 프롬프트에 포함된 사용자 데이터: 사용자의 입력을 기본 모델 학습 또는 개선에 사용하지 않는 관리형 AI 서비스에서 처리됩니다.
- 기술 자료 콘텐츠: 엄선되어 게시된 NetApp 문서만 포함되며, 사용자 데이터는 포함되지 않습니다.
- 운영 데이터: 사용자가 자신의 리소스를 조회할 때만 검색되며, 세션 기간 동안 임시 메모리 저장소에 보관되고 영구 저장되거나 공유되지 않습니다.
- 감사 로깅: 쿼리 메타데이터(익명 처리된 사용자 ID, 타임스탬프, 지속 시간)가 기록되며, 민감한 필드는 삭제됩니다.
- 피드백 데이터: 별도로 저장되며 쿼리 ID에 연결되고, 익명화된 사용자 ID 외의 개인 식별 정보와는 연결되지 않습니다.

가용성 및 격리 제어

Ask Me는 여러 클라우드 지역에 걸쳐 다중 모델 대체 체인을 사용합니다.

기본 모델이 조절되거나 사용할 수 없는 경우, 시스템은 대체 모델로 페일오버할 수 있습니다.

모든 모델은 동일한 보안 제어, 시스템 프롬프트 강화 및 격리 보장이 적용됩니다.