시작하기
NetApp Data Classification

NetApp
February 06, 2026

목차

시작하기

NetApp Data Classification 에 대해 알아보세요

NetApp Data Classification 는 NetApp Console 위한 데이터 거버넌스 서비스로, 기업의 온프레미스 및 클라우드 데이터 소스를 스캔하여 데이터를 매핑하고 분류하며 개인 정보를 식별합니다. 이를 통해 보안 및 규정 준수 위험을 줄이고, 스토리지 비용을 절감하고, 데이터 마이그레이션 프로젝트를 지원할 수 있습니다.

버전 1.31부터 데이터 분류가 NetApp Console 의 핵심 기능으로 제공됩니다. 추가 비용은 없습니다. 분류 라이센스나 구독이 필요하지 않습니다. + 기존 버전 1.30 또는 이전 버전을 사용 중이신 경우, 구독이 만료될 때까지 해당 버전을 사용할 수 있습니다.

NetApp Console

데이터 분류는 NetApp Console 통해 접근할 수 있습니다.

NetApp Console 엔터프라이즈급 온프레미스 및 클라우드 환경 전반에서 NetApp 스토리지 및 데이터 서비스를 중앙에서 관리할 수 있는 기능을 제공합니다. NetApp 데이터 서비스에 액세스하고 사용하려면 콘솔이 필요합니다. 관리 인터페이스로서, 하나의 인터페이스에서 여러 스토리지 리소스를 관리할 수 있습니다. 콘솔 관리자는 기업 내 모든 시스템의 저장소와 서비스에 대한 액세스를 제어할 수 있습니다.

NetApp Console 사용하려면 라이선스나 구독이 필요하지 않으며, 스토리지 시스템이나 NetApp 데이터 서비스에 대한 연결을 보장하기 위해 클라우드에 Console 에이전트를 배포해야 할 때만 요금이 부과됩니다. 그러나 콘솔에서 액세스할 수 있는 일부 NetApp 데이터 서비스는 라이선스 기반이거나 구독 기반입니다.

자세히 알아보세요"NetApp Console" .

특징

데이터 분류는 인공지능(AI), 자연어 처리(NLP), 머신 러닝(ML)을 사용하여 스캔한 콘텐츠를 이해하고 엔터티를 추출하고 그에 따라 콘텐츠를 분류합니다. 이를 통해 데이터 분류는 다음과 같은 기능 영역을 제공할 수 있습니다.

"데이터 분류 사용 사례에 대해 알아보세요" .

규정 준수 유지

데이터 분류는 규정 준수 노력에 도움이 되는 다양한 도구를 제공합니다. 데이터 분류를 사용하여 다음을 수행할 수 있습니다.

- 개인 식별 정보(PII)를 식별합니다.
- GDPR, CCPA, PCI 및 HIPAA 개인정보 보호 규정에서 요구하는 대로 광범위한 민감한 개인 정보를 식별합니다.
- 이름이나 이메일 주소를 기반으로 데이터 주체 접근 요청(DSAR)에 응답합니다.

보안 강화

데이터 분류를 통해 범죄 목적으로 접근될 위험이 있는 데이터를 식별할 수 있습니다. 데이터 분류를 사용하여 다음을 수행할 수 있습니다.

- 전체 조직이나 대중에게 공개된, 공개 권한이 있는 모든 파일과 디렉토리(공유 및 폴더)를 식별합니다.

- 처음 지정된 위치 외부에 있는 민감한 데이터를 식별합니다.
- 데이터 보존 정책을 준수합니다.
- 정책을 사용하면 새로운 보안 문제를 자동으로 감지하여 보안 직원이 즉시 조치를 취할 수 있습니다.

저장 공간 사용량 최적화

데이터 분류는 스토리지 총 소유 비용(TCO)을 절감하는 데 도움이 되는 도구를 제공합니다. 데이터 분류를 사용하여 다음을 수행할 수 있습니다.

- 중복된 데이터나 업무와 관련 없는 데이터를 식별하여 저장 효율성을 높입니다.
- 비활성 데이터를 식별하여 비용이 덜 드는 개체 스토리지로 계층화하여 스토리지 비용을 절감하세요. "Cloud Volumes ONTAP 시스템의 계층화에 대해 자세히 알아보세요." . "온프레미스 ONTAP 시스템의 계층화에 대해 자세히 알아보세요." .

지원되는 시스템 및 데이터 소스

데이터 분류는 다음 유형의 시스템 및 데이터 소스에서 구조화된 데이터와 구조화되지 않은 데이터를 스캔하고 분석할 수 있습니다.

시스템

- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- Cloud Volumes ONTAP (AWS, Azure 또는 GCP에 배포됨)
- 온프레미스 ONTAP 클러스터
- StorageGRID
- Google Cloud NetApp Volumes

데이터 출처

- NetApp 파일 공유
- 데이터베이스:
  - Amazon 관계형 데이터베이스 서비스(Amazon RDS)
  - 몽고디비
  - MySQL
  - 신탁
  - 포스트그레스큐엘
  - SAP 하나
  - SQL 서버(MSSQL)

데이터 분류는 NFS 버전 3.x, 4.0, 4.1과 CIFS 버전 1.x, 2.0, 2.1, 3.0을 지원합니다.

비용

데이터 분류는 무료로 사용할 수 있습니다. 분류 라이센스나 유료 구독이 필요하지 않습니다.

인프라 비용

- 클라우드에 데이터 분류를 설치하려면 클라우드 인스턴스를 배포해야 하며, 배포된 클라우드 제공업체에서 요금이 부과됩니다. 보다 각 클라우드 공급자에 배포되는 인스턴스 유형 . 온프레미스 시스템에 데이터 분류를 설치하는 경우 비용이 발생하지 않습니다.
- 데이터 분류를 위해서는 콘솔 에이전트를 배포해야 합니다. 많은 경우 콘솔에서 다른 저장소와 서비스를 사용하고 있기 때문에 이미 콘솔 에이전트가 있는 것입니다. 콘솔 에이전트 인스턴스는 배포된 클라우드 공급자로부터 요금이 부과됩니다. 를 참조하십시오 "각 클라우드 공급자에 배포되는 인스턴스 유형" . 온프레미스 시스템에 콘솔 에이전트를 설치하는 경우 비용이 발생하지 않습니다.

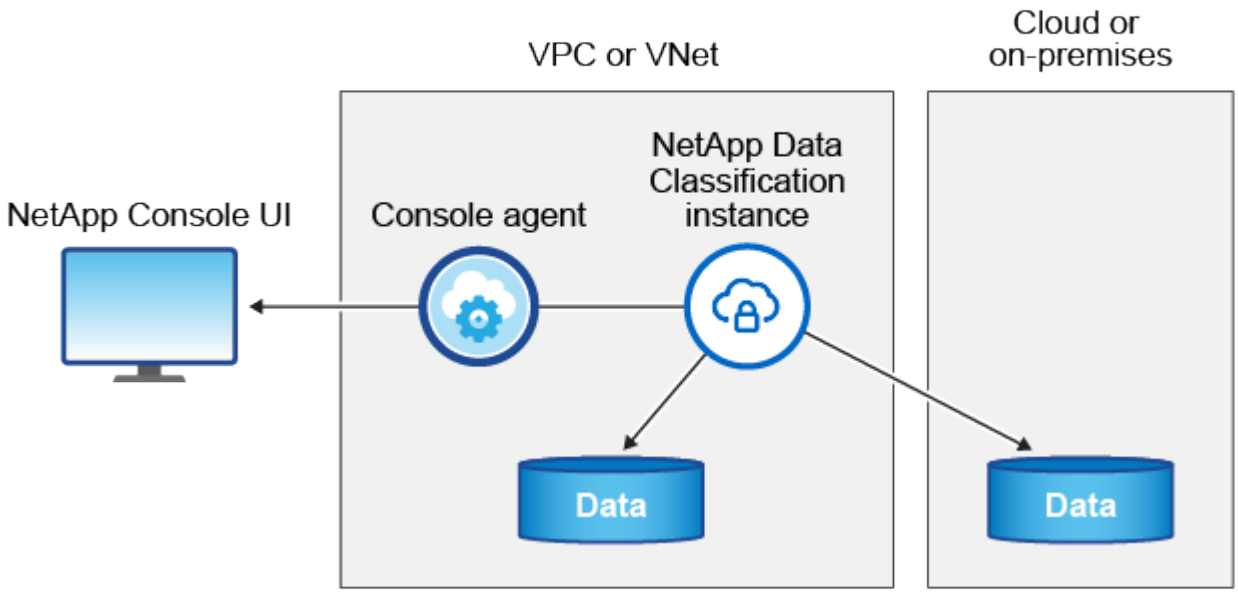
데이터 전송 비용

데이터 전송 비용은 설정에 따라 달라집니다. 데이터 분류 인스턴스와 데이터 소스가 동일한 가용성 영역 및 지역에 있는 경우 데이터 전송 비용이 발생하지 않습니다. 하지만 Cloud Volumes ONTAP 시스템과 같은 데이터 소스가 다른 가용성 영역이나 지역에 있는 경우 클라우드 공급자가 데이터 전송 비용을 청구합니다. 자세한 내용은 다음 링크를 참조하세요.

- "AWS: Amazon Elastic Compute Cloud(Amazon EC2) 가격"
- "Microsoft Azure: 대역폭 가격 세부 정보"
- "Google Cloud: Storage Transfer Service 가격 책정"

데이터 분류 인스턴스

클라우드에 데이터 분류를 배포하면 콘솔은 콘솔 에이전트와 동일한 서브넷에 인스턴스를 배포합니다. "콘솔 에이전트에 대해 자세히 알아보세요."



기본 인스턴스에 대해 다음 사항을 참고하세요.

- AWS에서는 데이터 분류가 실행됩니다. "m6i.4xlarge 인스턴스" 500GiB GP2 디스크 포함. 운영체제 이미지는 Amazon Linux 2입니다. AWS에 배포하는 경우 소량의 데이터를 스캔하는 경우 더 작은 인스턴스 크기를 선택할 수 있습니다.
- Azure에서 데이터 분류는 다음에서 실행됩니다."Standard_D16s_v3 VM" 500GiB 디스크 포함. 운영체제 이미지는 Ubuntu 22.04입니다.

- GCP에서 데이터 분류는 다음에서 실행됩니다."n2-standard-16 VM" 500GiB 표준 영구 디스크를 사용합니다. 운영체제 이미지는 Ubuntu 22.04입니다.
- 기본 인스턴스를 사용할 수 없는 지역에서는 데이터 분류가 대체 인스턴스에서 실행됩니다. "대체 인스턴스 유형을 확인하세요" .
- 인스턴스 이름은 *CloudCompliance_*이고, 생성된 해시*(UUID)*가 여기에 연결됩니다. 예*: _CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- 콘솔 에이전트당 하나의 데이터 분류 인스턴스만 배포됩니다.

사내 Linux 호스트나 선호하는 클라우드 공급업체의 호스트에 데이터 분류를 배포할 수도 있습니다. 어떤 설치 방법을 선택하든 소프트웨어의 기능은 정확히 동일합니다. 인스턴스에 인터넷 접속이 가능한 한 데이터 분류 소프트웨어 업그레이드는 자동화됩니다.

데이터 분류는 지속적으로 데이터를 스캔하므로 인스턴스는 항상 실행 상태를 유지해야 합니다.

다양한 인스턴스 유형에 배포

인스턴스 유형에 대한 다음 사양을 검토하세요.

시스템 크기	명세서	제한 사항
특대	32개 CPU, 128GB RAM, 1TiB SSD	최대 5억 개의 파일을 검색할 수 있습니다.
대형(기본값)	CPU 16개, 64GB RAM, 500GiB SSD	최대 2억 5천만 개의 파일을 스캔할 수 있습니다.

Azure 또는 GCP에서 데이터 분류를 배포할 때 더 작은 인스턴스 유형을 사용하려면 ng-contact-data-sense@netapp.com으로 이메일을 보내 지원을 요청하세요.

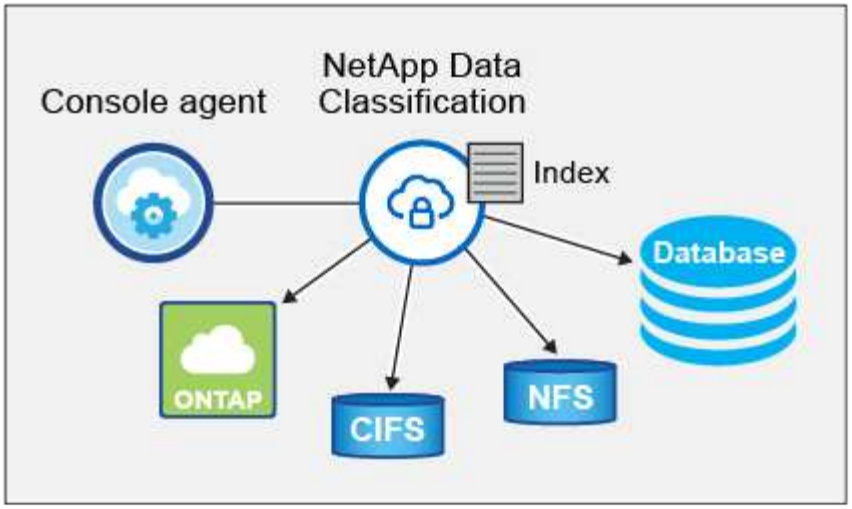
데이터 분류 스캐닝 작동 방식

높은 수준에서 데이터 분류 스캐닝은 다음과 같이 작동합니다.

1. 콘솔에서 데이터 분류 인스턴스를 배포합니다.
2. 하나 이상의 데이터 소스에서 고수준 매핑(매핑 전용 스캔이라고 함) 또는 심층 수준 스캐닝(맵 및 분류 스캔이라고 함)을 활성화합니다.
3. 데이터 분류는 AI 학습 프로세스를 사용하여 데이터를 스캔합니다.
4. 제공된 대시보드와 보고 도구를 사용하면 규정 준수 및 거버넌스 활동에 도움이 됩니다.

데이터 분류를 활성화하고 스캔하려는 저장소(볼륨, 데이터베이스 스키마 또는 기타 사용자 데이터)를 선택하면 즉시 데이터 스캔을 시작하여 개인 및 민감한 데이터를 식별합니다. 대부분의 경우 백업, 미러 또는 DR 사이트 대신 라이브 프로덕션 데이터 스캔에 집중해야 합니다. 그런 다음 데이터 분류는 조직 데이터를 매핑하고, 각 파일을 분류하고, 데이터에서 엔터티와 사전 정의된 패턴을 식별하여 추출합니다. 검사 결과는 개인 정보, 민감한 개인 정보, 데이터 범주 및 파일 유형의 인덱스입니다.

데이터 분류는 NFS 및 CIFS 볼륨을 마운트하여 다른 클라이언트와 마찬가지로 데이터에 연결합니다. NFS 볼륨은 자동으로 읽기 전용으로 액세스되는 반면, CIFS 볼륨을 스캔하려면 Active Directory 자격 증명을 제공해야 합니다.



초기 스캔 이후, 데이터 분류는 라운드 로빈 방식으로 데이터를 지속적으로 스캔하여 증분적 변경 사항을 감지합니다. 인스턴스를 계속 실행하는 것이 중요한 이유가 여기에 있습니다.

볼륨 수준이나 데이터베이스 스키마 수준에서 검사를 활성화하거나 비활성화할 수 있습니다.

데이터 분류는 스캔할 수 있는 데이터 양에 제한을 두지 않습니다. 각 콘솔 에이전트는 500TiB의 데이터를 스캔하고 표시하는 것을 지원합니다. 500TiB 이상의 데이터를 스캔하려면"다른 콘솔 에이전트를 설치하세요" 그 다음에"다른 데이터 분류 인스턴스 배포" . + 콘솔 UI는 단일 커넥터의 데이터를 표시합니다. 여러 콘솔 에이전트의 데이터를 보는 방법에 대한 팁은 다음을 참조하세요."여러 콘솔 에이전트와 함께 작업" .

매핑 스캔과 분류 스캔의 차이점은 무엇입니까?

데이터 분류에서는 두 가지 유형의 스캔을 수행할 수 있습니다.

- 매핑 전용 스캔은 데이터에 대한 개략적인 개요만 제공하며 선택된 데이터 소스에서 수행됩니다. 매핑 전용 스캔은 파일에 액세스하여 내부 데이터를 확인하지 않으므로 매핑 및 분류 스캔보다 시간이 덜 걸립니다. 연구할 분야를 파악하기 위해 먼저 이 작업을 수행한 다음 해당 분야에 대한 지도 및 분류 검사를 수행하는 것이 좋습니다.
- **Map & Classify** 스캔은 데이터에 대한 심층적인 스캔을 제공합니다.

매핑 스캔과 분류 스캔의 차이점에 대한 자세한 내용은 다음을 참조하세요."매핑 스캔과 분류 스캔의 차이점은 무엇인가요?" .

데이터 분류가 분류하는 정보

데이터 분류는 다음 데이터를 수집, 색인화하고 범주를 지정합니다.

- 파일에 대한 표준 메타데이터: 파일 유형, 크기, 생성 및 수정 날짜 등.
- 개인 데이터: 이메일 주소, 신분증 번호 또는 신용 카드 번호와 같은 개인 식별 정보(PII)로, 데이터 분류는 파일에서 특정 단어, 문자열 및 패턴을 사용하여 이를 식별합니다. "개인 데이터에 대해 자세히 알아보세요" .
- 민감한 개인 정보: 건강 데이터, 민족적 기원 또는 정치적 의견과 같은 특수 유형의 민감한 개인 정보(SPII)로, 일반 데이터 보호 규정(GDPR) 및 기타 개인정보 보호 규정에 정의되어 있습니다. "민감한 개인 데이터에 대해 자세히 알아보세요" .
- 범주: 데이터 분류는 스캔한 데이터를 여러 유형의 범주로 분류합니다. 카테고리는 각 파일의 콘텐츠와

메타데이터에 대한 AI 분석을 기반으로 한 주제입니다. "카테고리에 대해 자세히 알아보세요".

- 이름 엔터티 인식: 데이터 분류는 AI를 사용하여 문서에서 사람들의 실제 이름을 추출합니다. "데이터 주체 접근 요청에 응답하는 방법에 대해 알아보세요" .

네트워킹 개요

데이터 분류는 클라우드나 온프레미스 등 원하는 곳에 단일 서버 또는 클러스터를 배포합니다. 서버는 표준 프로토콜을 통해 데이터 소스에 연결하고, 동일한 서버에 배포된 Elasticsearch 클러스터에서 검색 결과를 인덱싱합니다. 이를 통해 멀티 클라우드, 크로스 클라우드, 프라이빗 클라우드 및 온프레미스 환경을 지원할 수 있습니다.

콘솔은 콘솔 에이전트에서 인바운드 HTTP 연결을 활성화하는 보안 그룹과 함께 데이터 분류 인스턴스를 배포합니다.

SaaS 모드에서 콘솔을 사용하는 경우 콘솔 연결은 HTTPS를 통해 제공되고 브라우저와 데이터 분류 인스턴스 간에 전송되는 개인 데이터는 TLS 1.2를 사용하여 종단 간 암호화로 보호되므로 NetApp 과 타사가 해당 데이터를 읽을 수 없습니다.

아웃바운드 규칙은 완전히 공개되어 있습니다. 데이터 분류 소프트웨어를 설치하고 업그레이드하고 사용 지표를 전송하려면 인터넷 접속이 필요합니다.

엄격한 네트워킹 요구 사항이 있는 경우"데이터 분류가 접촉하는 엔드포인트에 대해 알아보세요" .

NetApp Data Classification 액세스

NetApp Console 통해 NetApp Data Classification 액세스할 수 있습니다.

콘솔에 로그인하려면 NetApp 지원 사이트 자격 증명을 사용하거나 이메일과 비밀번호를 사용하여 NetApp Console 로그인에 가입할 수 있습니다. "콘솔에 로그인하는 방법에 대해 자세히 알아보세요" .

특정 작업에는 특정 콘솔 사용자 역할이 필요합니다. "모든 서비스에 대한 콘솔 액세스 역할에 대해 알아보세요." .

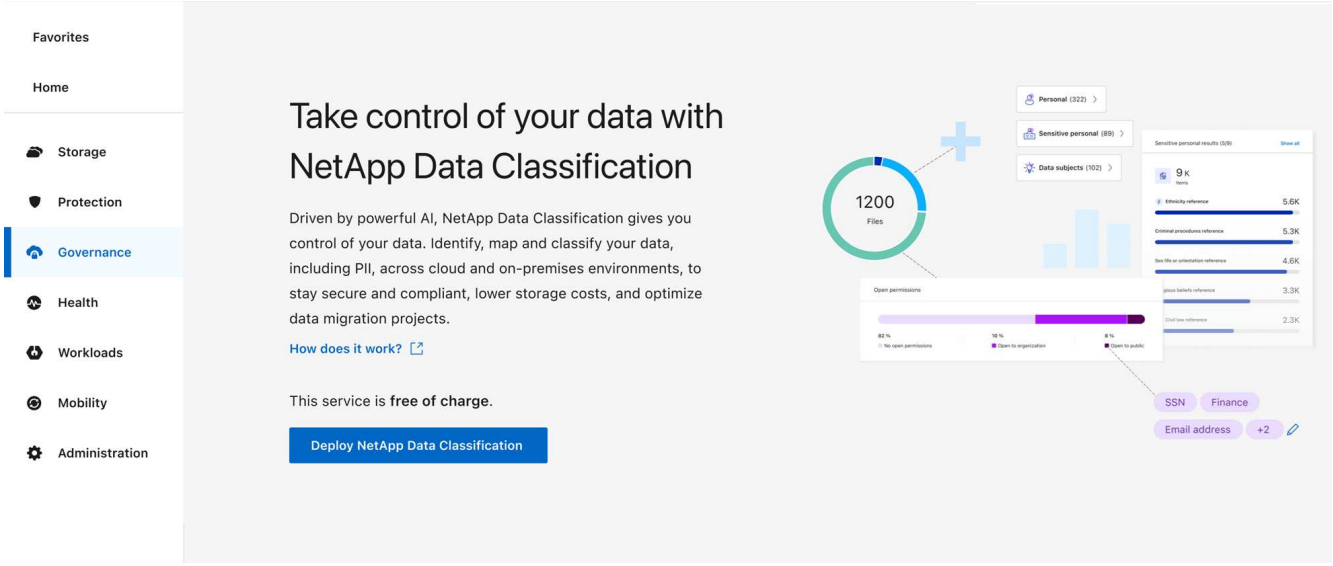
시작하기 전에

- "콘솔 에이전트를 추가해야 합니다."
- "귀하의 작업 부하에 적합한 데이터 분류 배포 스타일을 파악하세요."

단계

1. 웹 브라우저에서 다음으로 이동합니다."콘솔" .
2. 콘솔에 로그인합니다.
3. NetApp Console 의 메인 페이지에서 거버넌스 > *데이터 분류*를 선택합니다.
4. 처음으로 데이터 분류에 접속하는 경우 랜딩 페이지가 나타납니다.

   분류 인스턴스 배포를 시작하려면 *온프레미스 또는 클라우드에 분류 배포*를 선택하세요. 자세한 내용은 다음을 참조하세요."어떤 데이터 분류 배포를 사용해야 합니까?"



그렇지 않으면 데이터 분류 대시보드가 나타납니다.

데이터 분류 배포

어떤 **NetApp Data Classification** 배포를 사용해야 합니까**?**

NetApp Data Classification 다양한 방법으로 배포할 수 있습니다. 어떤 방법이 귀하의 필요에 맞는지 알아보세요.

데이터 분류는 다음과 같은 방법으로 배포될 수 있습니다.

- "콘솔을 사용하여 클라우드에 배포" . 콘솔은 콘솔 에이전트와 동일한 클라우드 공급자 네트워크에 데이터 분류 인스턴스를 배포합니다.
- "인터넷 접속이 가능한 Linux 호스트에 설치" . 인터넷 접속이 가능한 네트워크의 Linux 호스트나 클라우드의 Linux 호스트에 데이터 분류를 설치합니다. 온프레미스 ONTAP 시스템을 온프레미스에 있는 데이터 분류 인스턴스를 사용하여 스캔하려는 경우 이러한 유형의 설치가 좋은 옵션일 수 있지만, 이는 필수 사항은 아닙니다.
- "인터넷 접속이 불가능한 온프레미스 사이트의 Linux 호스트에 설치"_비공개 모드_라고도 합니다. 설치 스크립트를 사용하는 이 유형의 설치는 콘솔 SaaS 계층에 연결할 수 없습니다.

BlueXP 개인 모드(레거시 BlueXP 인터페이스)는 일반적으로 인터넷 연결이 없고 AWS Secret Cloud, AWS Top Secret Cloud, Azure IL6를 포함하는 보안 클라우드 지역이 있는 온프레미스 환경에서 사용됩니다. NetApp 기존 BlueXP 인터페이스를 통해 이러한 환경을 계속 지원합니다. 레거시 BlueXP 인터페이스의 개인 모드 설명서는 다음을 참조하세요."BlueXP 개인 모드에 대한 PDF 문서" .

인터넷 접속이 가능한 Linux 호스트에 설치하는 경우와 인터넷 접속이 불가능한 Linux 호스트에 온프레미스로 설치하는 경우 모두 설치 스크립트를 사용합니다. 스크립트는 시스템과 환경이 전제 조건을 충족하는지 확인하는 것으로 시작합니다. 필수 구성 요소가 충족되면 설치가 시작됩니다. 데이터 분류 설치를 실행하지 않고도 전제 조건을 독립적으로 확인하려면 전제 조건만 테스트하는 별도의 소프트웨어 패키지를 다운로드할 수 있습니다.

"Linux 호스트가 데이터 분류를 설치할 준비가 되었는지 확인하세요." .

NetApp Console 사용하여 클라우드에 NetApp Data Classification 배포합니다.

NetApp Console 사용하여 클라우드에 NetApp Data Classification 배포할 수 있습니다. 콘솔은 콘솔 에이전트와 동일한 클라우드 공급자 네트워크에 데이터 분류 인스턴스를 배포합니다.

또한 다음을 수행할 수도 있습니다."인터넷 접속이 가능한 Linux 호스트에 데이터 분류 설치" . 온프레미스 ONTAP 시스템을 온프레미스에 있는 데이터 분류 인스턴스를 사용하여 스캔하려는 경우 이러한 유형의 설치가 좋은 옵션일 수 있지만, 이는 필수 사항은 아닙니다. 어떤 설치 방법을 선택하든 소프트웨어의 기능은 정확히 동일합니다.

빠른 시작

다음 단계에 따라 빠르게 시작하거나, 나머지 섹션으로 스크롤하여 자세한 내용을 확인하세요.

1 콘솔 에이전트 만들기

아직 콘솔 에이전트가 없으면 하나 만드세요. 보다"AWS에서 콘솔 에이전트 만들기" ,"Azure에서 콘솔 에이전트 만들기" , 또는"GCP에서 콘솔 에이전트 만들기" .

당신도 할 수 있습니다"온프레미스에 콘솔 에이전트 설치" 네트워크의 Linux 호스트나 클라우드의 Linux 호스트에서.

2 필수 조건

사용 환경이 필수 요구 사항을 충족하는지 확인하십시오. 여기에는 인스턴스의 아웃바운드 인터넷 액세스, Console 에이전트와 Data Classification 간의 포트 443을 통한 연결 등이 포함됩니다. 전체 목록을 확인하세요.

3 데이터 분류 배포

설치 마법사를 실행하여 클라우드에 데이터 분류 인스턴스를 배포합니다.

콘솔 에이전트 만들기

아직 콘솔 에이전트가 없다면 클라우드 공급자에서 콘솔 에이전트를 만드세요. 보다 "AWS에서 콘솔 에이전트 만들기" 또는 "Azure에서 콘솔 에이전트 만들기" , 또는 "GCP에서 콘솔 에이전트 만들기" . 대부분의 경우 데이터 분류를 활성화하기 전에 콘솔 에이전트를 설정했을 가능성이 높습니다. "콘솔 기능에는 콘솔 에이전트가 필요합니다." 하지만 지금 당장 설정해야 하는 경우도 있습니다.

특정 클라우드 공급자에 배포된 콘솔 에이전트를 사용해야 하는 몇 가지 시나리오가 있습니다.

- AWS의 Cloud Volumes ONTAP 또는 Amazon FSx for ONTAP 버킷에서 데이터를 스캔할 때 AWS의 콘솔 에이전트를 사용합니다.
- Azure의 Cloud Volumes ONTAP 또는 Azure NetApp Files 에서 데이터를 스캔할 때 Azure의 콘솔 에이전트를 사용합니다.
  - Azure NetApp Files 의 경우, 스캔하려는 볼륨과 동일한 지역에 배포해야 합니다.
- GCP에서 Cloud Volumes ONTAP 의 데이터를 스캔할 때 GCP의 콘솔 에이전트를 사용합니다.

이러한 클라우드 콘솔 에이전트를 사용하면 온프레미스 ONTAP 시스템, NetApp 파일 공유 및 데이터베이스를 검사할 수 있습니다.

또한 다음을 수행할 수도 있습니다. "온프레미스에 콘솔 에이전트 설치" 네트워크나 클라우드 내의 Linux 호스트에서. 온프레미스에 데이터 분류를 설치하려는 일부 사용자는 온프레미스에 콘솔 에이전트를 설치하기로 선택할 수도 있습니다.

사용해야 하는 상황이 있을 수 있습니다. "여러 콘솔 에이전트" .

데이터 분류는 스캔할 수 있는 데이터 양에 제한을 두지 않습니다. 각 콘솔 에이전트는 500TiB의 데이터를 스캔하고 표시하는 것을 지원합니다. 500TiB 이상의 데이터를 스캔하려면"다른 콘솔 에이전트를 설치하세요" 그 다음에"다른 데이터 분류 인스턴스 배포" . + 콘솔 UI는 단일 커넥터의 데이터를 표시합니다. 여러 콘솔 에이전트의 데이터를 보는 방법에 대한 팁은 다음을 참조하세요."여러 콘솔 에이전트와 함께 작업" .

정부 지역 지원

콘솔 에이전트가 정부 지역(AWS GovCloud, Azure Gov 또는 Azure DoD)에 배포된 경우 데이터 분류가 지원됩니다. 이런 방식으로 배포할 경우 데이터 분류에는 다음과 같은 제한이 있습니다.

"정부 지역에 콘솔 에이전트를 배포하는 방법에 대해 알아보세요.".

필수 조건

클라우드에 데이터 분류를 배포하기 전에 지원되는 구성이 있는지 확인하려면 다음 필수 구성 요소를 검토하세요. 클라우드에 데이터 분류를 배포하면 콘솔 에이전트와 동일한 서브넷에 위치하게 됩니다.

데이터 분류에서 아웃바운드 인터넷 액세스 활성화

데이터 분류에는 아웃바운드 인터넷 액세스가 필요합니다. 가상 또는 물리적 네트워크에서 인터넷 접속을 위해 프록시 서버를 사용하는 경우, 데이터 분류 인스턴스에 다음 엔드포인트에 연결할 수 있는 아웃바운드 인터넷 접속 권한이 있는지 확인하세요. 프록시는 투명하지 않아야 합니다. 투명 프록시는 현재 지원되지 않습니다.

AWS, Azure 또는 GCP에서 데이터 분류를 배포하는지에 따라 아래 해당 표를 검토하세요.

AWS에 필요한 엔드포인트

엔드포인트	목적
\ https://api.console.netapp.com	NetApp 계정을 포함하는 콘솔 서비스와의 통신입니다.
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	중앙화된 사용자 인증을 위해 콘솔 웹사이트와 통신합니다.
\ https://cloud-compliance-support-netapp.s3.us-west-2.amazonaws.com \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	소프트웨어 이미지, 매니페스트, 템플릿에 대한 액세스를 제공합니다.
\ https://kinesis.us-east-1.amazonaws.com	NetApp 감사 기록에서 데이터를 스트리밍할 수 있도록 합니다.
\ https://cognito-idp.us-east-1.amazonaws.com \ https://cognito-identity.us-east-1.amazonaws.com \ https://user-feedback-store-prod.s3.us-west-2.amazonaws.com \ https://customer-data-production.s3.us-west-2.amazonaws.com	데이터 분류를 통해 매니페스트와 템플릿에 액세스하고 다운로드하며, 로그와 메트릭을 전송할 수 있습니다.

Azure에 필요한 엔드포인트

엔드포인트	목적
\ https://api.console.netapp.com	NetApp 계정을 포함하는 콘솔 서비스와의 통신입니다.
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	중앙화된 사용자 인증을 위해 콘솔 웹사이트와 통신합니다.
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	소프트웨어 이미지, 매니페스트, 템플릿에 대한 액세스를 제공하고 로그와 메트릭을 전송합니다.
\ https://support.compliance.api.console.netapp.com/	NetApp 감사 기록에서 데이터를 스트리밍할 수 있도록 합니다.

GCP에 필요한 엔드포인트

엔드포인트	목적
\ https://api.console.netapp.com	NetApp 계정을 포함하는 콘솔 서비스와의 통신입니다.
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	중앙화된 사용자 인증을 위해 콘솔 웹사이트와 통신합니다.

엔드포인트	목적
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	소프트웨어 이미지, 매니페스트, 템플릿에 대한 액세스를 제공하고 로그와 메트릭을 전송합니다.
\ https://support.compliance.api.console.netapp.com/	NetApp 감사 기록에서 데이터를 스트리밍할 수 있도록 합니다.

데이터 분류에 필요한 권한이 있는지 확인하세요.

데이터 분류에 리소스를 배포하고 데이터 분류 인스턴스에 대한 보안 그룹을 생성할 수 있는 권한이 있는지 확인하세요.

- "Google Cloud 권한"
- "AWS 권한"
- "Azure 권한"

콘솔 에이전트가 데이터 분류에 액세스할 수 있는지 확인하세요.

콘솔 에이전트와 데이터 분류 인스턴스 간의 연결을 보장합니다. 콘솔 에이전트의 보안 그룹은 포트 443을 통해 데이터 분류 인스턴스와의 인바운드 및 아웃바운드 트래픽을 허용해야 합니다. 이 연결을 통해 데이터 분류 인스턴스를 배포하고 규정 준수 및 거버넌스 탭에서 정보를 볼 수 있습니다. 데이터 분류는 AWS와 Azure의 정부 지역에서 지원됩니다.

AWS 및 AWS GovCloud 배포에는 추가적인 인바운드 및 아웃바운드 보안 그룹 규칙이 필요합니다. 보다 "AWS의 콘솔 에이전트에 대한 규칙" 자세한 내용은.

Azure 및 Azure Government 배포에는 추가적인 인바운드 및 아웃바운드 보안 그룹 규칙이 필요합니다. 보다 "Azure의 콘솔 에이전트에 대한 규칙" 자세한 내용은.

데이터 분류를 계속 실행할 수 있는지 확인하세요.

데이터 분류 인스턴스는 지속적으로 데이터를 스캔하기 위해 켜져 있어야 합니다.

데이터 분류에 대한 웹 브라우저 연결을 보장합니다.

데이터 분류가 활성화된 후, 사용자가 데이터 분류 인스턴스에 연결된 호스트에서 콘솔 인터페이스에 액세스하는지 확인하세요.

데이터 분류 인스턴스는 개인 IP 주소를 사용하여 인덱싱된 데이터가 인터넷에서 접근되지 않도록 보장합니다. 따라서 콘솔에 접속하는 데 사용하는 웹 브라우저는 해당 개인 IP 주소에 연결되어 있어야 합니다. 해당 연결은 클라우드 공급자(예: VPN)에 대한 직접 연결을 통해 이루어질 수도 있고, 데이터 분류 인스턴스와 동일한 네트워크 내부에 있는 호스트를 통해 이루어질 수도 있습니다.

vCPU 제한을 확인하세요

클라우드 제공업체의 vCPU 한도가 필요한 수의 코어를 갖춘 인스턴스를 배포할 수 있는지 확인하세요. 콘솔이 실행되는 지역에서 해당 인스턴스 패밀리에 대한 vCPU 제한을 확인해야 합니다. "필요한 인스턴스 유형을 확인하세요" .

vCPU 제한에 대한 자세한 내용은 다음 링크를 참조하세요.

- "AWS 설명서: Amazon EC2 서비스 할당량"
- "Azure 설명서: 가상 머신 vCPU 할당량"
- "Google Cloud 문서: 리소스 할당량"

클라우드에 데이터 분류 배포

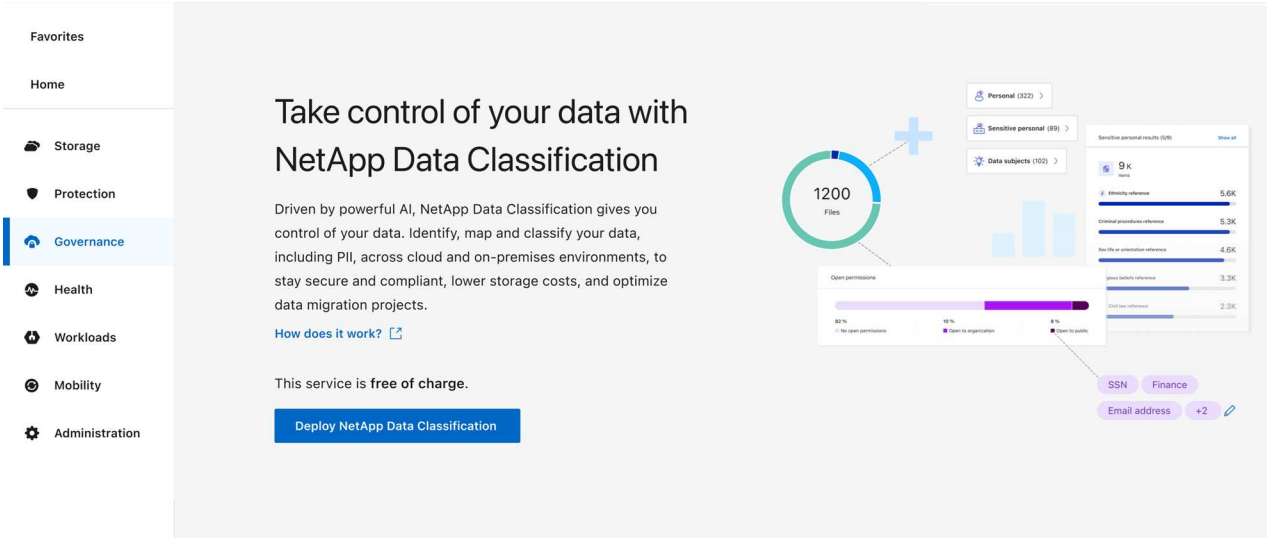
클라우드에 데이터 분류 인스턴스를 배포하려면 다음 단계를 따르세요. 콘솔 에이전트는 클라우드에 인스턴스를 배포한 다음 해당 인스턴스에 데이터 분류 소프트웨어를 설치합니다.

기본 인스턴스 유형을 사용할 수 없는 지역에서는 데이터 분류가 실행됩니다."대체 인스턴스 유형" .

AWS에 배포

단계

1. 데이터 분류의 메인 페이지에서 *온프레미스 또는 클라우드에 분류 배포*를 선택합니다.

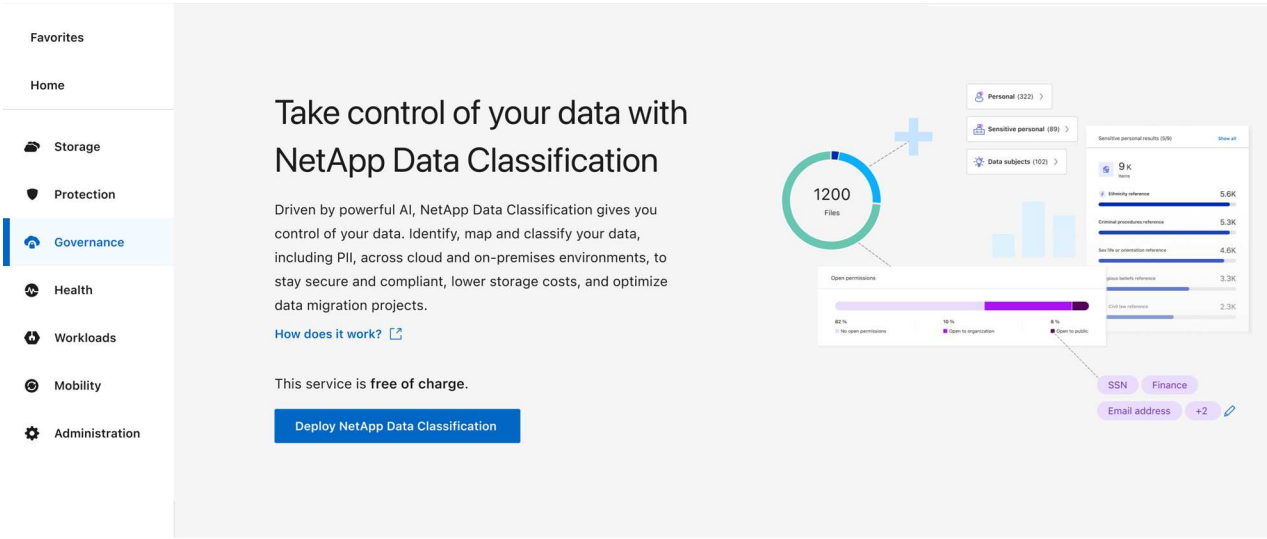


2. 설치 페이지에서 *배포 > 배포*를 선택하여 "대형" 인스턴스 크기를 사용하고 클라우드 배포 마법사를 시작합니다.
3. 마법사는 배포 단계를 진행하면서 진행 상황을 표시합니다. 입력이 필요하거나 문제가 발생하면 메시지가 표시됩니다.
4. 인스턴스가 배포되고 데이터 분류가 설치되면 *구성 계속*을 선택하여 구성 페이지로 이동합니다.

Azure에 배포

단계

1. 데이터 분류의 메인 페이지에서 *온프레미스 또는 클라우드에 분류 배포*를 선택합니다.



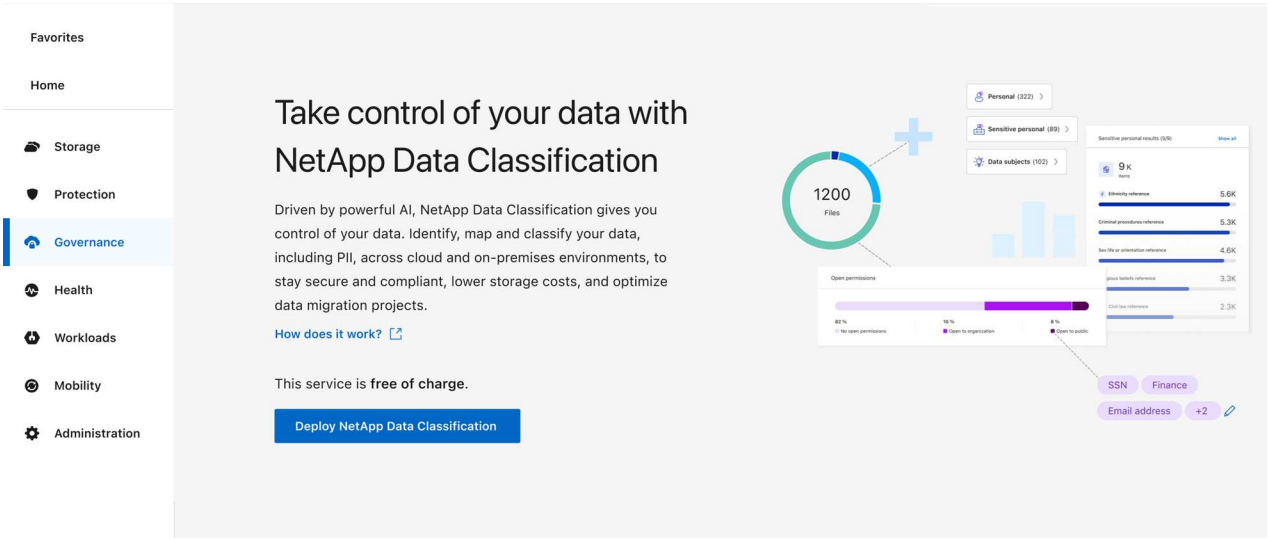
2. 클라우드 배포 마법사를 시작하려면 *배포*를 선택하세요.
3. 마법사는 배포 단계를 진행하면서 진행 상황을 표시합니다. 문제가 발생하면 멈추고 입력을 요청합니다.

4. 인스턴스가 배포되고 데이터 분류가 설치되면 *구성 계속*을 선택하여 구성 페이지로 이동합니다.

Google Cloud에 배포

단계

1. 데이터 분류의 메인 페이지에서 *거버넌스 > 분류*를 선택합니다.
2. *온프레미스 또는 클라우드에 분류 배포*를 선택합니다.



3. 클라우드 배포 마법사를 시작하려면 *배포*를 선택하세요.
4. 마법사는 배포 단계를 진행하면서 진행 상황을 표시합니다. 문제가 발생하면 멈추고 입력을 요청합니다.
5. 인스턴스가 배포되고 데이터 분류가 설치되면 *구성 계속*을 선택하여 구성 페이지로 이동합니다.

결과

콘솔은 클라우드 공급자에 데이터 분류 인스턴스를 배포합니다.

인스턴스가 인터넷에 연결되어 있는 한 콘솔 에이전트와 데이터 분류 소프트웨어의 업그레이드는 자동화됩니다.

다음은 무엇인가

구성 페이지에서 스캔할 데이터 소스를 선택할 수 있습니다.

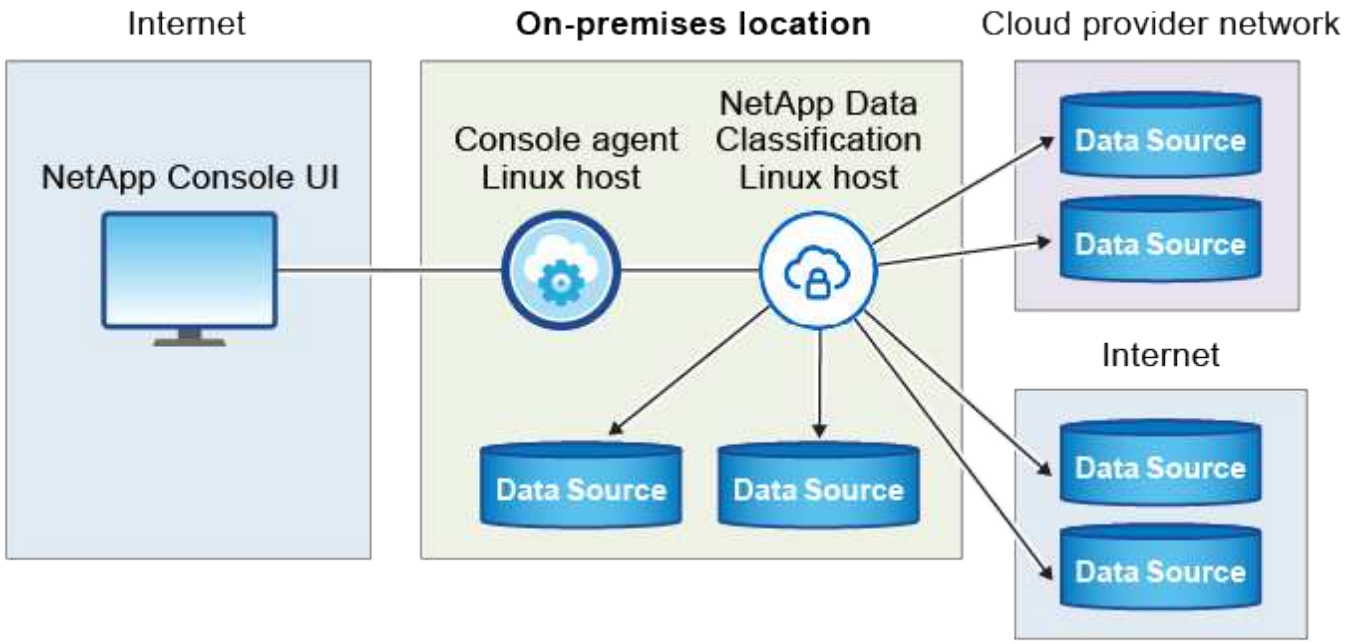
인터넷 접속이 가능한 호스트에 **NetApp Data Classification** 설치

네트워크의 Linux 호스트나 인터넷 접속이 가능한 클라우드의 Linux 호스트에 NetApp Data Classification 배포하려면 네트워크나 클라우드에 Linux 호스트를 수동으로 배포해야 합니다.

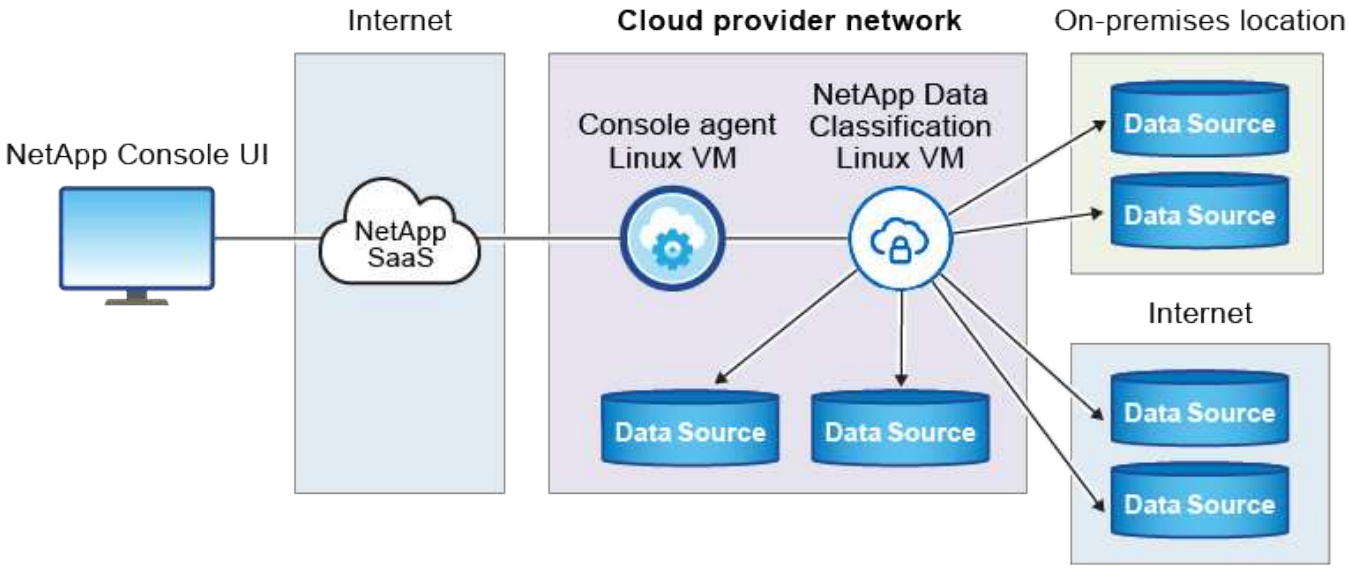
온프레미스 설치는 온프레미스에 있는 데이터 분류 인스턴스를 사용하여 온프레미스 ONTAP 시스템을 스캔하는 것을 선호하는 경우에 좋은 옵션입니다. 이것은 필수사항이 아닙니다. 어떤 설치 방법을 선택하든 소프트웨어의 기능은 동일합니다.

데이터 분류 설치 스크립트는 시스템과 환경이 필수 전제 조건을 충족하는지 확인하는 것으로 시작됩니다. 모든 전제 조건이 충족되면 설치가 시작됩니다. 데이터 분류 설치를 실행하지 않고도 전제 조건을 독립적으로 확인하려면 전제 조건만 테스트하는 별도의 소프트웨어 패키지를 다운로드할 수 있습니다. "Linux 호스트가 데이터 분류를 설치할 준비가 되었는지 확인하는 방법을 알아보세요." .

귀사 구내의 Linux 호스트에 일반적으로 설치되는 구성 요소와 연결은 다음과 같습니다.



클라우드에 있는 Linux 호스트에 일반적으로 설치되는 구성 요소와 연결은 다음과 같습니다.



빠른 시작

다음 단계에 따라 빠르게 시작하거나, 나머지 섹션으로 스크롤하여 자세한 내용을 확인하세요.

1 콘솔 에이전트 만들기

아직 콘솔 에이전트가 없는 경우 "온프레미스에 콘솔 에이전트 배포" 네트워크의 Linux 호스트나 클라우드의 Linux 호스트에서.

클라우드 공급자를 사용하여 콘솔 에이전트를 생성할 수도 있습니다. 보다 "AWS에서 콘솔 에이전트 만들기" , "Azure에서 콘솔 에이전트 만들기" , 또는 "GCP에서 콘솔 에이전트 만들기" .

2 필수 조건 검토

귀하의 환경이 전제 조건을 충족하는지 확인하세요. 여기에는 인스턴스에 대한 아웃바운드 인터넷 액세스, 포트 443을 통한 콘솔 에이전트와 데이터 분류 간의 연결 등이 포함됩니다. 전체 목록을 확인하세요 .

또한 다음을 충족하는 Linux 시스템이 필요합니다.다음 요구 사항 .

3 데이터 분류 다운로드 및 배포

NetApp 지원 사이트에서 클라우드 데이터 분류 소프트웨어를 다운로드하고 사용하려는 Linux 호스트에 설치 프로그램 파일을 복사합니다. 그런 다음 설치 마법사를 실행하고 메시지에 따라 데이터 분류 인스턴스를 배포합니다.

콘솔 에이전트 만들기

데이터 분류를 설치하고 사용하려면 콘솔 에이전트가 필요합니다. 대부분의 경우 데이터 분류를 활성화하기 전에 콘솔 에이전트를 설정했을 가능성이 높습니다. "콘솔 기능에는 콘솔 에이전트가 필요합니다." 하지만 지금 당장 설정해야 하는 경우도 있습니다.

클라우드 공급자 환경에서 하나를 생성하려면 다음을 참조하세요. "AWS에서 콘솔 에이전트 만들기" , "Azure에서 콘솔 에이전트 만들기" , 또는 "GCP에서 콘솔 에이전트 만들기" .

특정 클라우드 공급자에 배포된 콘솔 에이전트를 사용해야 하는 몇 가지 시나리오가 있습니다.

- AWS의 Cloud Volumes ONTAP 또는 Amazon FSx for ONTAP 에서 데이터를 스캔할 때 AWS의 콘솔 에이전트를 사용합니다.
- Azure의 Cloud Volumes ONTAP 또는 Azure NetApp Files 에서 데이터를 스캔할 때 Azure의 콘솔 에이전트를 사용합니다.

  Azure NetApp Files 의 경우, 스캔하려는 볼륨과 동일한 지역에 배포해야 합니다.

- GCP에서 Cloud Volumes ONTAP 의 데이터를 스캔할 때 GCP의 콘솔 에이전트를 사용합니다.

온프레미스 ONTAP 시스템, NetApp 파일 공유 및 데이터베이스 계정은 이러한 클라우드 콘솔 에이전트를 사용하여 스캔할 수 있습니다.

또한 다음을 수행할 수도 있습니다. "온프레미스에 콘솔 에이전트 배포" 네트워크의 Linux 호스트나 클라우드의 Linux 호스트에서. 온프레미스에 데이터 분류를 설치하려는 일부 사용자는 콘솔 에이전트도 온프레미스에 설치하기로 선택할 수도 있습니다.

데이터 분류를 설치할 때 콘솔 에이전트 시스템의 IP 주소나 호스트 이름이 필요합니다. 사내에 콘솔 에이전트를 설치한 경우 이 정보를 얻을 수 있습니다. 콘솔 에이전트가 클라우드에 배포된 경우 콘솔에서 다음 정보를 찾을 수 있습니다. 도움말 아이콘을 선택한 다음 *지원*을 선택하고 콘솔 에이전트를 선택합니다.

Linux 호스트 시스템 준비

데이터 분류 소프트웨어는 특정 운영 체제 요구 사항, RAM 요구 사항, 소프트웨어 요구 사항 등을 충족하는 호스트에서 실행되어야 합니다. Linux 호스트는 네트워크에 있을 수도 있고 클라우드에 있을 수도 있습니다.

데이터 분류를 계속 실행할 수 있는지 확인하세요. 데이터 분류 머신은 지속적으로 데이터를 스캔하기 위해 계속 켜져 있어야 합니다.

- 데이터 분류는 전용 호스트에서 수행되어야 합니다. 호스트는 다른 애플리케이션이나 바이러스 백신과 같은 타사 소프트웨어와 공유할 수 없습니다.
- 데이터 분류를 통해 스캔할 데이터 세트에 맞는 크기를 선택하십시오.

시스템 크기	**CPU**	**RAM(**스왑 메모리를 비활성화해야 함**)**	디스크
특대	32개의 CPU	128GB 램	• /에 1TiB SSD, 또는 /opt에 100GiB 사용 가능 • /var/lib/docker에서 895GiB 사용 가능 • /tmp에 5GiB • **Podman**의 경우 **/var/tmp**에 **30GB**
크기가 큰	16개의 CPU	64GB 램	• /에 500GiB SSD, 또는 /opt에 100GiB 사용 가능 • /var/lib/docker 또는 Podman /var/lib/containers에서 400GiB 사용 가능 • /tmp에 5GiB • **Podman**의 경우 **/var/tmp**에 **30GB**

- 데이터 분류 설치를 위해 클라우드에 컴퓨팅 인스턴스를 배포할 때 위의 "대규모" 시스템 요구 사항을 충족하는 시스템을 사용하는 것이 좋습니다.
  - **Amazon Elastic Compute Cloud(Amazon EC2)** 인스턴스 유형: "m6i.4xlarge". "추가 AWS 인스턴스 유형 보기" .
  - **Azure VM** 크기: "Standard_D16s_v3". "추가 Azure 인스턴스 유형 보기" .
  - **GCP** 머신 유형: "n2-standard-16". "추가 GCP 인스턴스 유형을 참조하세요." .
- **UNIX** 폴더 권한: 다음과 같은 최소 UNIX 권한이 필요합니다.

접는 사람	최소 권한
/임시	`rwxrwxrwt`
/고르다	`rwxr-xr-x`
/var/lib/도커	`rwx------`
/usr/lib/systemd/시스템	`rwxr-xr-x`

- 운영체제:
  - 다음 운영 체제에서는 Docker 컨테이너 엔진을 사용해야 합니다.
    - Red Hat Enterprise Linux 버전 7.8 및 7.9

        - Ubuntu 22.04(데이터 분류 버전 1.23 이상 필요)
        - Ubuntu 24.04(데이터 분류 버전 1.23 이상 필요)
    - 다음 운영 체제에서는 Podman 컨테이너 엔진을 사용해야 하며, Data Classification 버전 1.30 이상이 필요합니다.
        - Red Hat Enterprise Linux 버전 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 및 9.6.
    - 호스트 시스템에서 고급 벡터 확장(AVX2)을 활성화해야 합니다.
- **Red Hat Subscription Management**: 호스트는 Red Hat Subscription Management에 등록되어야 합니다. 등록되지 않은 경우, 설치 중에 시스템은 저장소에 접근하여 필요한 타사 소프트웨어를 업데이트할 수 없습니다.
- 추가 소프트웨어: 데이터 분류를 설치하기 전에 호스트에 다음 소프트웨어를 설치해야 합니다.
    - 사용 중인 운영체제에 따라 컨테이너 엔진 중 하나를 설치해야 합니다.
        - Docker Engine 버전 19.3.1 이상. "설치 지침 보기" .
        - Podman 버전 4 이상. Podman을 설치하려면 다음을 입력하세요.(`sudo yum install podman netavark -y`).
- Python 버전 3.6 이상. "설치 지침 보기" .
    - **NTP** 고려 사항: NetApp 데이터 분류 시스템을 구성하여 NTP(네트워크 시간 프로토콜) 서비스를 사용할 것을 권장합니다. 데이터 분류 시스템과 콘솔 에이전트 시스템 간의 시간은 동기화되어야 합니다.
- 방화벽 고려 사항: 방화벽을 사용하려는 경우 `firewalld` 데이터 분류를 설치하기 전에 해당 기능을 활성화하는 것이 좋습니다. 다음 명령을 실행하여 구성하세요. `firewalld` 데이터 분류와 호환되도록:

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

추가 데이터 분류 호스트를 스캐너 노드로 사용할 계획이라면 지금 바로 기본 시스템에 다음 규칙을 추가하세요.

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

Docker 또는 Podman을 활성화하거나 업데이트할 때마다 다시 시작해야 합니다. `firewalld` 설정.

데이터 분류 호스트 시스템의 IP 주소는 설치 후 변경할 수 없습니다.

데이터 분류에서 아웃바운드 인터넷 액세스 활성화

데이터 분류에는 아웃바운드 인터넷 액세스가 필요합니다. 가상 또는 물리적 네트워크에서 인터넷 접속을 위해 프록시

서버를 사용하는 경우, 데이터 분류 인스턴스에 다음 엔드포인트에 연결할 수 있는 아웃바운드 인터넷 접속 권한이 있는지 확인하세요.

엔드포인트	목적
\ https://api.console.netapp.com	NetApp 계정을 포함한 콘솔과의 통신.
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	중앙화된 사용자 인증을 위해 콘솔 웹사이트와 통신합니다.
\ https://support.compliance.api.bluexp.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	소프트웨어 이미지, 매니페스트, 템플릿에 대한 액세스를 제공하고 로그와 메트릭을 전송합니다.
\ https://support.compliance.api.bluexp.netapp.com/	NetApp 감사 기록에서 데이터를 스트리밍할 수 있도록 합니다.
\ https://github.com/docker \ https://download.docker.com	Docker 설치를 위한 필수 패키지를 제공합니다.
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Ubuntu 설치를 위한 필수 패키지를 제공합니다.

모든 필수 포트가 활성화되어 있는지 확인하세요

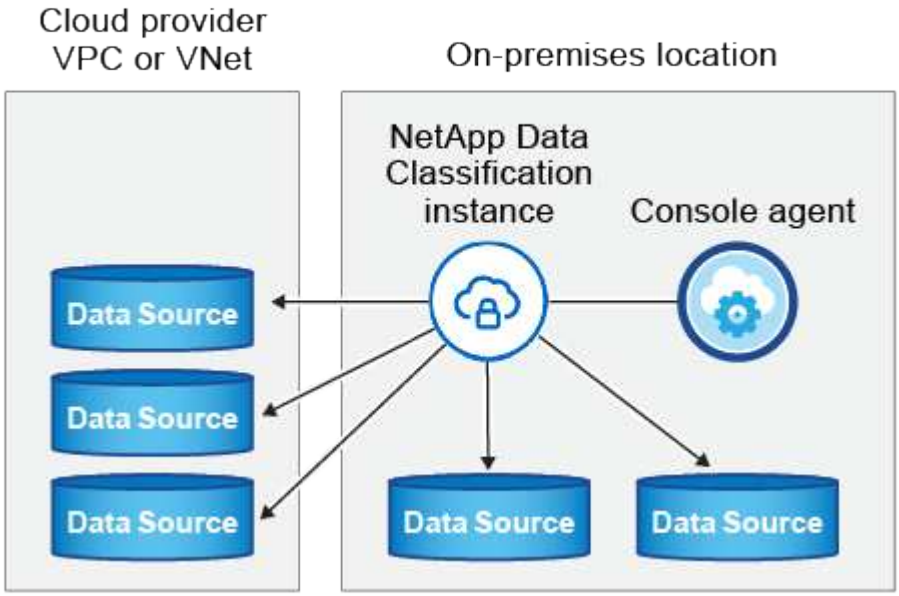
콘솔 에이전트, 데이터 분류, Active Directory 및 데이터 소스 간 통신에 필요한 모든 포트가 열려 있는지 확인해야 합니다.

연결 유형	포트	설명
콘솔 에이전트 <> 데이터 분류	8080(TCP), 443(TCP), 80. 9000	콘솔 에이전트의 방화벽 또는 라우팅 규칙은 포트 443을 통해 데이터 분류 인스턴스로의 인바운드 및 아웃바운드 트래픽을 허용해야 합니다. 콘솔에서 설치 진행 상황을 볼 수 있도록 포트 8080이 열려 있는지 확인하세요. Linux 호스트에서 방화벽을 사용하는 경우 Ubuntu 서버 내의 내부 프로세스에는 포트 9000이 필요합니다.

연결 유형	포트	설명
콘솔 에이전트 <> ONTAP 클러스터(NAS)	443(TCP)	콘솔은 HTTPS를 사용하여 ONTAP 클러스터를 검색합니다. 사용자 지정 방화벽 정책을 사용하는 경우 다음 요구 사항을 충족해야 합니다. • 콘솔 에이전트 호스트는 포트 443을 통해 아웃바운드 HTTPS 액세스를 허용해야 합니다. 콘솔 에이전트가 클라우드에 있는 경우 모든 아웃바운드 통신은 사전 정의된 방화벽이나 라우팅 규칙에 따라 허용됩니다. • ONTAP 클러스터는 포트 443을 통해 인바운드 HTTPS 액세스를 허용해야 합니다. 기본 "mgmt" 방화벽 정책은 모든 IP 주소에서 인바운드 HTTPS 액세스를 허용합니다. 이 기본 정책을 수정했거나 사용자 고유의 방화벽 정책을 만든 경우 HTTPS 프로토콜을 해당 정책과 연결하고 콘솔 에이전트 호스트에서 액세스를 활성화해야 합니다.
데이터 분류 <> ONTAP 클러스터	• NFS의 경우 - 111(TCP\UDP) 및 2049(TCP\UDP) • CIFS의 경우 - 139(TCP\UDP) 및 445(TCP\UDP)	데이터 분류에는 각 Cloud Volumes ONTAP 서브넷이나 온프레미스 ONTAP 시스템에 대한 네트워크 연결이 필요합니다. Cloud Volumes ONTAP 의 방화벽이나 라우팅 규칙은 데이터 분류 인스턴스에서 인바운드 연결을 허용해야 합니다. 다음 포트가 데이터 분류 인스턴스에 열려 있는지 확인하세요. • NFS - 111 및 2049의 경우 • CIFS - 139 및 445의 경우 NFS 볼륨 내보내기 정책은 데이터 분류 인스턴스에서의 액세스를 허용해야 합니다.
데이터 분류 <> Active Directory	389(TCP 및 UDP), 636(TCP), 3268(TCP), 3269(TCP)	회사 사용자를 위해 Active Directory가 이미 설정되어 있어야 합니다. 또한, 데이터 분류에는 CIFS 볼륨을 스캔하기 위한 Active Directory 자격 증명이 필요합니다. Active Directory에 대한 정보가 있어야 합니다. • DNS 서버 IP 주소 또는 여러 IP 주소 • 서버의 사용자 이름 및 비밀번호 • 도메인 이름(Active Directory 이름) • 보안 LDAP(LDAPS)를 사용하든 사용하지 않든 • LDAP 서버 포트(일반적으로 LDAP의 경우 389, 보안 LDAP의 경우 636)

Linux 호스트에 데이터 분류 설치

일반적인 구성의 경우 단일 호스트 시스템에 소프트웨어를 설치합니다. 여기에서 해당 단계를 확인하세요 .



보다Linux 호스트 시스템 준비 그리고필수 조건 검토 데이터 분류를 배포하기 전에 필요한 전체 요구 사항 목록을 확인하세요.

인스턴스에 인터넷 연결이 있는 한 데이터 분류 소프트웨어 업그레이드는 자동으로 수행됩니다.

현재 데이터 분류 기능은 온프레미스에 소프트웨어가 설치된 경우 S3 버킷, Azure NetApp Files 또는 FSx for ONTAP 검색할 수 없습니다. 이러한 경우 클라우드에 별도의 콘솔 에이전트와 데이터 분류 인스턴스를 배포해야 합니다. "커넥터 간 전환" 다양한 데이터 소스에 대해.

일반적인 구성을 위한 단일 호스트 설치

단일 온프레미스 호스트에 데이터 분류 소프트웨어를 설치할 때 요구 사항을 검토하고 다음 단계를 따르세요.

"이 영상을 시청하세요"데이터 분류를 설치하는 방법을 알아보세요.

데이터 분류를 설치할 때 모든 설치 활동이 기록됩니다. 설치 중에 문제가 발생하면 설치 감사 로그의 내용을 볼 수 있습니다. 에 쓰여있다 `/opt/netapp/install_logs/` .

시작하기 전에

- Linux 시스템이 다음 사항을 충족하는지 확인하십시오.호스트 요구 사항 .
- 시스템에 두 가지 필수 소프트웨어 패키지(Docker Engine 또는 Podman, Python 3)가 설치되어 있는지 확인합니다.
- Linux 시스템에서 루트 권한이 있는지 확인하세요.
- 인터넷에 접속하기 위해 프록시를 사용하는 경우:
  - 프록시 서버 정보(IP 주소 또는 호스트 이름, 연결 포트, 연결 방식: https 또는 http, 사용자 이름 및 비밀번호)가 필요합니다.
  - 프록시가 TLS 가로채기를 수행하는 경우 TLS CA 인증서가 저장된 Data Classification Linux 시스템의

경로를 알아야 합니다.

- 프록시는 투명하지 않아야 합니다. 데이터 분류는 현재 투명 프록시를 지원하지 않습니다.
- 사용자는 로컬 사용자여야 합니다. 도메인 사용자는 지원되지 않습니다.

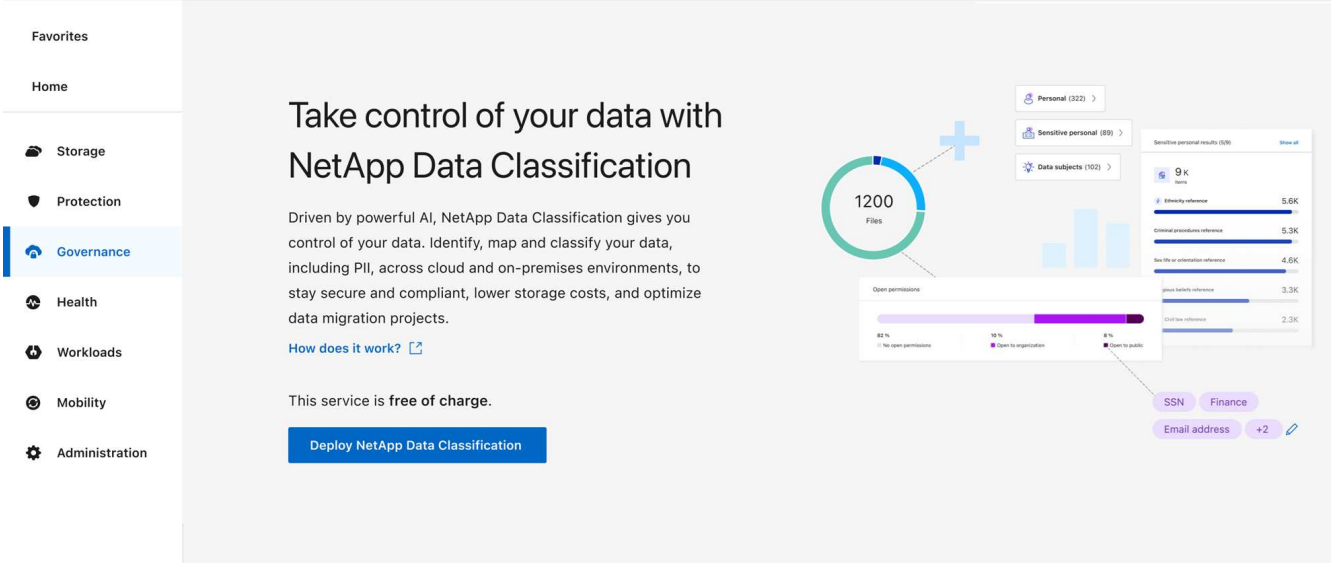
- 오프라인 환경이 요구 사항을 충족하는지 확인하세요.권한 및 연결 .

단계

1. 데이터 분류 소프트웨어를 다운로드하세요. "NetApp 지원 사이트" . 선택해야 하는 파일의 이름은 *DATASENSE-INSTALLER-<버전>.tar.gz*입니다.
2. 사용하려는 Linux 호스트에 설치 프로그램 파일을 복사합니다(사용 `scp` 또는 다른 방법).
3. 호스트 컴퓨터에서 설치 프로그램 파일의 압축을 풉니다. 예:

```
tar -xzf DATASENSE-INSTALLER-V1.25.0.tar.gz
```

4. 콘솔에서 *거버넌스 > 분류*를 선택합니다.
5. *온프레미스 또는 클라우드에 분류 배포*를 선택합니다.



6. 클라우드에서 준비한 인스턴스에 데이터 분류를 설치하는지, 아니면 사내에서 준비한 인스턴스에 데이터 분류를 설치하는지에 따라 적절한 배포 옵션을 선택하여 데이터 분류 설치를 시작합니다.
7. 온프레미스에 데이터 분류 배포 대화 상자가 표시됩니다. 제공된 명령을 복사합니다(예: `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`)을 텍스트 파일에 붙여넣어 나중에 사용할 수 있습니다. 그런 다음 *닫기*를 선택하여 대화 상자를 닫습니다.
8. 호스트 머신에서 복사한 명령을 입력한 다음 일련의 프롬프트를 따르거나 모든 필수 매개변수를 포함한 전체 명령을 명령줄 인수로 제공할 수 있습니다.

   설치 프로그램은 성공적인 설치를 위해 시스템 및 네트워킹 요구 사항이 충족되는지 사전 점검을 수행합니다. "이 영상을 시청하세요" 사전 확인 메시지와 그 의미를 이해합니다.

프롬프트에 따라 매개변수를 입력하세요.	전체 명령을 입력하세요:
a. 7단계에서 복사한 명령을 붙여넣습니다. `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` 클라우드 인스턴스(사내가 아닌)에 설치하는 경우 다음을 추가하세요. `--manual-cloud-install <cloud_provider>` . b. 콘솔 에이전트 시스템에서 액세스할 수 있도록 데이터 분류 호스트 머신의 IP 주소 또는 호스트 이름을 입력하세요. c. 데이터 분류 시스템에서 액세스할 수 있도록 콘솔 에이전트 호스트 머신의 IP 주소 또는 호스트 이름을 입력하세요. d. 지시에 따라 프록시 세부 정보를 입력하세요. 콘솔 에이전트가 이미 프록시를 사용하는 경우 데이터 분류가 자동으로 콘솔 에이전트에서 사용하는 프록시를 사용하므로 여기에 다시 정보를 입력할 필요가 없습니다.	또는 필요한 호스트 및 프록시 매개변수를 제공하여 전체 명령을 미리 만들 수 있습니다. `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token> --host <ds_host> --manager-host <cm_host> --manual-cloud-install <cloud_provider> --proxy-host <proxy_host> --proxy-port <proxy_port> --proxy-scheme <proxy_scheme> --proxy -user <proxy_user> --proxy-password <proxy_password> --cacert-folder-path <ca_cert_dir>`

변수 값:

- *account_id* = NetApp 계정 ID
- *client_id* = 콘솔 에이전트 클라이언트 ID(클라이언트 ID에 접미사 "clients"가 없으면 추가)
- *user_token* = JWT 사용자 액세스 토큰
- *ds_host* = 데이터 분류 Linux 시스템의 IP 주소 또는 호스트 이름입니다.
- *cm_host* = 콘솔 에이전트 시스템의 IP 주소 또는 호스트 이름입니다.
- *cloud_provider* = 클라우드 인스턴스에 설치하는 경우 클라우드 공급자에 따라 "AWS", "Azure" 또는 "Gcp"를 입력하세요.
- *proxy_host* = 호스트가 프록시 서버 뒤에 있는 경우 프록시 서버의 IP 또는 호스트 이름입니다.
- *proxy_port* = 프록시 서버에 연결할 포트(기본값 80).
- *proxy_scheme* = 연결 방식: https 또는 http(기본값은 http).
- *proxy_user* = 기본 인증이 필요한 경우 프록시 서버에 연결하는 인증된 사용자입니다. 사용자는 로컬 사용자여야 합니다. 도메인 사용자는 지원되지 않습니다.
- *proxy_password* = 지정한 사용자 이름에 대한 비밀번호입니다.
- *ca_cert_dir* = 추가 TLS CA 인증서 번들이 포함된 Data Classification Linux 시스템의 경로입니다. 프록시가 TLS 가로채기를 수행하는 경우에만 필요합니다.

결과

데이터 분류 설치 프로그램은 패키지를 설치하고, 설치를 등록하고, 데이터 분류를 설치합니다. 설치하는 데 10~20분이 걸릴 수 있습니다.

호스트 머신과 콘솔 에이전트 인스턴스 사이에 포트 8080을 통해 연결이 있는 경우 콘솔의 데이터 분류 탭에서 설치

진행률을 볼 수 있습니다.

다음은 무엇인가

구성 페이지에서 스캔할 데이터 소스를 선택할 수 있습니다.

인터넷 접속이 없는 **Linux** 호스트에 **NetApp Data Classification** 설치

인터넷 접속이 불가능한 온프레미스 사이트의 Linux 호스트에 NetApp Data Classification 설치하는 것을 _개인 모드_라고 합니다. 설치 스크립트를 사용하는 이 유형의 설치는 NetApp Console SaaS 계층에 연결되지 않습니다.

BlueXP 개인 모드(레거시 BlueXP 인터페이스)는 일반적으로 인터넷 연결이 없고 AWS Secret Cloud, AWS Top Secret Cloud, Azure IL6를 포함하는 보안 클라우드 지역이 있는 온프레미스 환경에서 사용됩니다. NetApp 기존 BlueXP 인터페이스를 통해 이러한 환경을 계속 지원합니다. 레거시 BlueXP 인터페이스의 개인 모드 설명서는 다음을 참조하세요."BlueXP 개인 모드에 대한 PDF 문서" .

Linux 호스트가 **NetApp Data Classification** 설치할 준비가 되었는지 확인하세요**.**

Linux 호스트에 NetApp Data Classification 수동으로 설치하기 전에 호스트에서 스크립트를 실행하여 Data Classification을 설치하는 데 필요한 모든 전제 조건이 충족되었는지 확인합니다. 이 스크립트는 네트워크의 Linux 호스트나 클라우드의 Linux 호스트에서 실행할 수 있습니다. 호스트는 인터넷에 연결되어 있을 수도 있고, 인터넷에 접속할 수 없는 사이트(다크 사이트)에 있을 수도 있습니다.

데이터 분류 설치 스크립트에는 환경이 요구 사항을 충족하는지 확인하는 테스트 스크립트가 포함되어 있습니다. 설치 스크립트를 실행하기 전에 Linux 호스트의 준비 상태를 확인하기 위해 이 스크립트를 별도로 실행할 수 있습니다.

시작하기

다음 작업을 수행하게 됩니다.

- 선택적으로, 콘솔 에이전트가 설치되어 있지 않으면 설치하세요. 콘솔 에이전트를 설치하지 않고도 테스트 스크립트를 실행할 수 있지만, 스크립트는 콘솔 에이전트와 데이터 분류 호스트 머신 간의 연결을 확인합니다. 따라서 콘솔 에이전트를 설치하는 것이 좋습니다.
- 호스트 머신을 준비하고 모든 요구 사항을 충족하는지 확인하세요.
- 데이터 분류 호스트 머신에서 아웃바운드 인터넷 액세스를 활성화합니다.
- 모든 시스템에서 필요한 포트가 모두 활성화되어 있는지 확인하세요.
- 필수 테스트 스크립트를 다운로드하여 실행하세요.

콘솔 에이전트 만들기

데이터 분류를 설치하고 사용하려면 콘솔 에이전트가 필요합니다. 하지만 콘솔 에이전트 없이도 필수 구성 요소 스크립트를 실행할 수 있습니다.

당신은 할 수 있습니다 "온프레미스에 콘솔 에이전트 설치" 네트워크 내의 Linux 호스트 또는 클라우드의 Linux 호스트에서 실행할 수 있습니다. 콘솔 에이전트가 온프레미스에 설치되어 있는 경우 데이터 분류 기능도 온프레미스에 설치할 수 있습니다.

클라우드 공급자 환경에서 콘솔 에이전트를 생성하려면 다음을 참조하세요.

- "AWS에서 콘솔 에이전트 만들기"
- "Azure에서 콘솔 에이전트 만들기"
- "GCP에서 콘솔 에이전트 만들기"

필수 조건 스크립트를 실행하려면 콘솔 에이전트 시스템의 IP 주소 또는 호스트 이름이 필요합니다. 콘솔 에이전트를 사내에 설치한 경우 이 정보를 확인할 수 있습니다. 콘솔 에이전트가 클라우드에 배포된 경우 콘솔에서 이 정보를 확인할 수 있습니다. 도움말 아이콘을 선택한 다음 *지원*을 선택하고, 에이전트 및 감사 섹션에서 *에이전트로 이동*을 선택하세요.

호스트 요구 사항 확인

데이터 분류 소프트웨어는 특정 운영 체제 요구 사항, RAM 요구 사항 및 소프트웨어 요구 사항을 충족하는 호스트에서 실행되어야 합니다.

- 데이터 분류는 전용 호스트에서 수행되어야 합니다. 호스트는 다른 애플리케이션이나 바이러스 백신과 같은 타사 소프트웨어와 공유할 수 없습니다.
- 데이터 분류를 통해 스캔할 데이터 세트에 맞는 크기를 선택하십시오.

시스템 크기	**CPU**	**RAM(**스왑 메모리를 비활성화해야 함**)**	디스크
특대	32개의 CPU	128GB 램	• /에 1TiB SSD, 또는 /opt에 100GiB 사용 가능 • /var/lib/docker에서 895GiB 사용 가능 • /tmp에 5GiB • **Podman**의 경우 **/var/tmp**에 **30GB**
크기가 큰	16개의 CPU	64GB 램	• /에 500GiB SSD, 또는 /opt에 100GiB 사용 가능 • /var/lib/docker 또는 Podman /var/lib/containers에서 400GiB 사용 가능 • /tmp에 5GiB • **Podman**의 경우 **/var/tmp**에 **30GB**

- 데이터 분류 설치를 위해 클라우드에 컴퓨팅 인스턴스를 배포할 때 위의 "대규모" 시스템 요구 사항을 충족하는 시스템을 사용하는 것이 좋습니다.
  - **Amazon Elastic Compute Cloud(Amazon EC2)** 인스턴스 유형: "m6i.4xlarge". "추가 AWS 인스턴스 유형 보기" .
  - **Azure VM** 크기: "Standard_D16s_v3". "추가 Azure 인스턴스 유형 보기" .

- **GCP** 머신 유형: "n2-standard-16". "추가 GCP 인스턴스 유형을 참조하세요." .

- **UNIX** 폴더 권한: 다음과 같은 최소 UNIX 권한이 필요합니다.

접는 사람	최소 권한
/임시	`rwxrwxrwt`
/고르다	`rwxr-xr-x`
/var/lib/도커	`rwx------`
/usr/lib/systemd/시스템	`rwxr-xr-x`

- 운영체제:
  - 다음 운영 체제에서는 Docker 컨테이너 엔진을 사용해야 합니다.
    - Red Hat Enterprise Linux 버전 7.8 및 7.9
    - Ubuntu 22.04(데이터 분류 버전 1.23 이상 필요)
    - Ubuntu 24.04(데이터 분류 버전 1.23 이상 필요)
  - 다음 운영 체제에서는 Podman 컨테이너 엔진을 사용해야 하며, Data Classification 버전 1.30 이상이 필요합니다.
    - Red Hat Enterprise Linux 버전 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 및 9.6.
  - 호스트 시스템에서 고급 벡터 확장(AVX2)을 활성화해야 합니다.
- **Red Hat Subscription Management**: 호스트는 Red Hat Subscription Management에 등록되어야 합니다. 등록되지 않은 경우, 설치 중에 시스템은 저장소에 접근하여 필요한 타사 소프트웨어를 업데이트할 수 없습니다.
- 추가 소프트웨어: 데이터 분류를 설치하기 전에 호스트에 다음 소프트웨어를 설치해야 합니다.
  - 사용 중인 운영체제에 따라 컨테이너 엔진 중 하나를 설치해야 합니다.
    - Docker Engine 버전 19.3.1 이상. "설치 지침 보기" .
    - Podman 버전 4 이상. Podman을 설치하려면 다음을 입력하세요.(`sudo yum install podman netavark -y`).
- Python 버전 3.6 이상. "설치 지침 보기" .
  - **NTP** 고려 사항: NetApp 데이터 분류 시스템을 구성하여 NTP(네트워크 시간 프로토콜) 서비스를 사용할 것을 권장합니다. 데이터 분류 시스템과 콘솔 에이전트 시스템 간의 시간은 동기화되어야 합니다.
- 방화벽 고려 사항: 방화벽을 사용하려는 경우 `firewalld` 데이터 분류를 설치하기 전에 해당 기능을 활성화하는 것이 좋습니다. 다음 명령을 실행하여 구성하세요. `firewalld` 데이터 분류와 호환되도록:

```
firewall-cmd --permanent --add-service=http
firewall-cmd --permanent --add-service=https
firewall-cmd --permanent --add-port=80/tcp
firewall-cmd --permanent --add-port=8080/tcp
firewall-cmd --permanent --add-port=443/tcp
firewall-cmd --reload
```

추가 데이터 분류 호스트를 스캐너 노드로 사용하려는 경우(분산 모델에서), 이때 다음 규칙을 기본 시스템에

추가하세요.

```
firewall-cmd --permanent --add-port=2377/tcp
firewall-cmd --permanent --add-port=7946/udp
firewall-cmd --permanent --add-port=7946/tcp
firewall-cmd --permanent --add-port=4789/udp
```

Docker 또는 Podman을 활성화하거나 업데이트할 때마다 다시 시작해야 합니다. `firewalld` 설정.

데이터 분류에서 아웃바운드 인터넷 액세스 활성화

데이터 분류에는 아웃바운드 인터넷 액세스가 필요합니다. 가상 또는 물리적 네트워크에서 인터넷 접속을 위해 프록시 서버를 사용하는 경우, 데이터 분류 인스턴스에 다음 엔드포인트에 연결할 수 있는 아웃바운드 인터넷 접속 권한이 있는지 확인하세요.

인터넷 연결이 없는 사이트에 설치된 호스트 시스템에는 이 섹션이 필요하지 않습니다.

엔드포인트	목적
\ https://api.console.netapp.com	NetApp 계정을 포함하는 콘솔 서비스와의 통신입니다.
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	중앙화된 사용자 인증을 위해 콘솔 웹사이트와 통신합니다.
\ https://support.compliance.api.console.netapp.com/ \ https://hub.docker.com \ https://auth.docker.io \ https://registry-1.docker.io \ https://index.docker.io/ \ https://dseasb33srnrn.cloudfront.net/ \ https://production.cloudflare.docker.com/	소프트웨어 이미지, 매니페스트, 템플릿에 대한 액세스를 제공하고 로그와 메트릭을 전송합니다.
\ https://support.compliance.api.console.netapp.com/	NetApp 감사 기록에서 데이터를 스트리밍할 수 있도록 합니다.
\ https://github.com/docker \ https://download.docker.com	Docker 설치를 위한 필수 패키지를 제공합니다.
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Ubuntu 설치를 위한 필수 패키지를 제공합니다.

모든 필수 포트가 활성화되어 있는지 확인하세요

콘솔 에이전트, 데이터 분류, Active Directory 및 데이터 소스 간 통신에 필요한 모든 포트가 열려 있는지 확인해야 합니다.

연결 유형	포트	설명
콘솔 에이전트 <> 데이터 분류	8080(TCP), 443(TCP), 80. 9000	콘솔 에이전트의 방화벽 또는 라우팅 규칙은 포트 443을 통해 데이터 분류 인스턴스로의 인바운드 및 아웃바운드 트래픽을 허용해야 합니다. 콘솔에서 설치 진행 상황을 볼 수 있도록 포트 8080이 열려 있는지 확인하세요. Linux 호스트에서 방화벽을 사용하는 경우 Ubuntu 서버 내의 내부 프로세스에는 포트 9000이 필요합니다.
콘솔 에이전트 <> ONTAP 클러스터(NAS)	443(TCP)	콘솔은 HTTPS를 사용하여 ONTAP 클러스터를 검색합니다. 사용자 지정 방화벽 정책을 사용하는 경우 콘솔 에이전트 호스트는 포트 443을 통한 아웃바운드 HTTPS 액세스를 허용해야 합니다. 콘솔 에이전트가 클라우드에 있는 경우 모든 아웃바운드 통신은 사전 정의된 방화벽이나 라우팅 규칙에 따라 허용됩니다.

데이터 분류 필수 조건 스크립트 실행

데이터 분류 필수 조건 스크립트를 실행하려면 다음 단계를 따르세요.

"이 영상을 시청하세요"필수 구성 요소 스크립트를 실행하고 결과를 해석하는 방법을 알아보세요.

시작하기 전에

- Linux 시스템이 다음 사항을 충족하는지 확인하십시오.호스트 요구 사항 .
- 시스템에 두 가지 필수 소프트웨어 패키지(Docker Engine 또는 Podman, Python 3)가 설치되어 있는지 확인합니다.
- Linux 시스템에서 루트 권한이 있는지 확인하세요.

단계

1. 데이터 분류 전제 조건 스크립트를 다운로드하세요. "NetApp 지원 사이트" . 선택해야 하는 파일의 이름은 *standalone-pre-requisite-tester-<version>*입니다.
2. 사용하려는 Linux 호스트에 파일을 복사합니다(사용 `scp` 또는 다른 방법).
3. 스크립트를 실행할 수 있는 권한을 할당합니다.

```
chmod +x standalone-pre-requisite-tester-v1.25.0
```

4. 다음 명령을 사용하여 스크립트를 실행하세요.

```
./standalone-pre-requisite-tester-v1.25.0 <--darksite>
```

   인터넷 접속이 불가능한 호스트에서 스크립트를 실행하는 경우에만 "--darksite" 옵션을 추가하세요. 호스트가 인터넷에 연결되어 있지 않으면 특정 필수 테스트가 건너뜁니다.

5. 스크립트는 데이터 분류 호스트 머신의 IP 주소를 입력하라는 메시지를 표시합니다.
   - IP 주소나 호스트 이름을 입력하세요.

6. 스크립트는 콘솔 에이전트가 설치되어 있는지 여부를 묻습니다.
   - 콘솔 에이전트가 설치되어 있지 않으면 *N*을 입력하세요.
   - 콘솔 에이전트가 설치되어 있는 경우 *Y*를 입력하세요. 그런 다음 테스트 스크립트가 이 연결성을 테스트할 수 있도록 콘솔 에이전트의 IP 주소나 호스트 이름을 입력합니다.
7. 스크립트는 시스템에서 다양한 테스트를 실행하고 진행 상황에 따라 결과를 표시합니다. 완료되면 세션 로그를 다음 이름의 파일에 기록합니다. `prerequisites-test-<timestamp>.log` 디렉토리에서 `/opt/netapp/install_logs` .

결과

모든 필수 테스트가 성공적으로 실행되었다면 준비가 되면 호스트에 데이터 분류를 설치할 수 있습니다.

문제가 발견되면 "권장" 또는 "필수"로 분류하여 수정합니다. 권장되는 문제는 일반적으로 데이터 분류 스캐닝 및 분류 작업의 실행 속도를 느리게 만드는 항목입니다. 이러한 항목은 수정할 필요가 없지만 해결하는 것이 좋습니다.

"필수" 문제가 있는 경우 문제를 해결하고 필수 구성 요소 테스트 스크립트를 다시 실행해야 합니다.

데이터 소스에서 스캐닝을 활성화하세요

NetApp Data Classification 사용하여 데이터 소스 스캔

NetApp Data Classification 사용자가 선택한 저장소(볼륨, 데이터베이스 스키마 또는 기타 사용자 데이터)의 데이터를 스캔하여 개인 데이터와 민감한 데이터를 식별합니다. 그런 다음 데이터 분류는 조직 데이터를 매핑하고, 각 파일을 분류하고, 데이터에서 미리 정의된 패턴을 식별합니다. 검사 결과는 개인 정보, 민감한 개인 정보, 데이터 범주 및 파일 유형의 인덱스입니다.

초기 스캔 이후, 데이터 분류는 라운드 로빈 방식으로 데이터를 지속적으로 스캔하여 증분적 변경 사항을 감지합니다. 인스턴스를 계속 실행하는 것이 중요한 이유가 여기에 있습니다.

볼륨 수준이나 데이터베이스 스키마 수준에서 검사를 활성화하거나 비활성화할 수 있습니다.

매핑 스캔과 분류 스캔의 차이점은 무엇입니까?

데이터 분류에서는 두 가지 유형의 스캔을 수행할 수 있습니다.

- 매핑 전용 스캔은 데이터에 대한 개략적인 개요만 제공하며 선택된 데이터 소스에서 수행됩니다. 매핑 전용 스캔은 파일에 액세스하여 내부 데이터를 확인하지 않기 때문에 매핑 및 분류 스캔보다 시간이 덜 걸립니다. 연구할 분야를 파악하기 위해 먼저 이 작업을 수행한 다음 해당 분야에 대한 지도 및 분류 검사를 수행하는 것이 좋습니다.
- **Map & Classify** 스캔은 데이터에 대한 심층적인 스캔을 제공합니다.

아래 표는 몇 가지 차이점을 보여줍니다.

특징	스캔 매핑 및 분류	매핑 전용 스캔
스캔 속도	느린	빠른
가격	무료	무료
용량	500TiB*로 제한됨	500TiB*로 제한됨

특징	스캔 매핑 및 분류	매핑 전용 스캔
파일 유형 및 사용 용량 목록	예	예
파일 개수 및 사용 용량	예	예
파일의 나이와 크기	예	예
실행할 수 있는 능력"데이터 매핑 보고서"	예	예
파일 세부 정보를 보려면 데이터 조사 페이지로 이동하세요.	예	아니요
파일 내에서 이름 검색	예	아니요
만들다"저장된 쿼리" 사용자 정의 검색 결과를 제공하는	예	아니요
다른 보고서를 실행하는 기능	예	아니요
파일의 메타데이터를 볼 수 있는 기능**	아니요	예

{별표} 데이터 분류는 스캔할 수 있는 데이터 양에 제한을 두지 않습니다. 각 콘솔 에이전트는 500TiB의 데이터를 스캔하고 표시하는 것을 지원합니다. 500TiB 이상의 데이터를 스캔하려면"다른 콘솔 에이전트를 설치하세요" 그 다음에"다른 데이터 분류 인스턴스 배포" . + 콘솔 UI는 단일 커넥터의 데이터를 표시합니다. 여러 콘솔 에이전트의 데이터를 보는 방법에 대한 팁은 다음을 참조하세요."여러 콘솔 에이전트와 함께 작업" .

{별표}{별표} 매핑 스캔 중에 파일에서 다음 메타데이터가 추출됩니다.

- 체계
- 시스템 유형
- 저장 저장소
- 파일 유형
- 사용된 용량
- 파일 수
- 파일 크기
- 파일 생성
- 파일 마지막 접근
- 파일이 마지막으로 수정되었습니다
- 파일 발견 시간
- 권한 추출

거버넌스 대시보드 차이점:

특징	지도 및 분류	지도
오래된 데이터	예	예
비업무용 데이터	예	예
중복된 파일	예	예
미리 정의된 저장된 쿼리	예	아니요
기본 저장된 쿼리	예	예
DDA 보고서	예	예
매핑 보고서	예	예
감도 수준 감지	예	아니요
광범위한 권한이 있는 민감한 데이터	예	아니요
공개 권한	예	예
데이터의 시대	예	예
데이터 크기	예	예
카테고리	예	아니요
파일 유형	예	예

규정 준수 대시보드 차이점:

특징	지도 및 분류	지도
개인정보	예	아니요
민감한 개인 정보	예	아니요
개인정보 위험 평가 보고서	예	아니요
HIPAA 보고서	예	아니요
PCI DSS 보고서	예	아니요

조사 필터의 차이점은 다음과 같습니다.

특징	지도 및 분류	지도
저장된 쿼리	예	예
시스템 유형	예	예
체계	예	예
저장 저장소	예	예
파일 유형	예	예
파일 크기	예	예
생성 시간	예	예
발견된 시간	예	예
마지막 수정	예	예
마지막 접근	예	예
공개 권한	예	예
파일 디렉토리 경로	예	예
범주	예	아니요
민감도 수준	예	아니요
식별자의 수	예	아니요
개인정보	예	아니요
민감한 개인 데이터	예	아니요
데이터 주체	예	아니요
중복	예	예
분류 상태	예	상태는 항상 "제한된 통찰력"입니다.
스캔 분석 이벤트	예	예
파일 해시	예	예
접근 권한이 있는 사용자 수	예	예
사용자/그룹 권한	예	예
파일 소유자	예	예
디렉토리 유형	예	예

NetApp Data Classification 사용하여 Amazon FSx 에서 ONTAP 볼륨 스캔

NetApp Data Classification 사용하여 Amazon FSx for ONTAP 볼륨을 스캔하려면 몇 가지 단계를 완료하세요.

시작하기 전에

- 데이터 분류를 배포하고 관리하려면 AWS에서 활성 콘솔 에이전트가 필요합니다.
- 시스템을 생성할 때 선택한 보안 그룹은 데이터 분류 인스턴스의 트래픽을 허용해야 합니다. FSx for ONTAP 파일 시스템에 연결된 ENI를 사용하여 연관된 보안 그룹을 찾고 AWS Management Console을 사용하여 편집할 수 있습니다.

  "Linux 인스턴스용 AWS 보안 그룹"

  "Windows 인스턴스용 AWS 보안 그룹"

  "AWS 탄력적 네트워크 인터페이스(ENI)"

- 다음 포트가 데이터 분류 인스턴스에 열려 있는지 확인하세요.
  - NFS의 경우 포트 111과 2049.
  - CIFS의 경우 포트 139 및 445.

데이터 분류 인스턴스 배포

"데이터 분류 배포"아직 배포된 인스턴스가 없는 경우.

AWS용 콘솔 에이전트와 스캔하려는 FSx 볼륨과 동일한 AWS 네트워크에 데이터 분류를 배포해야 합니다.

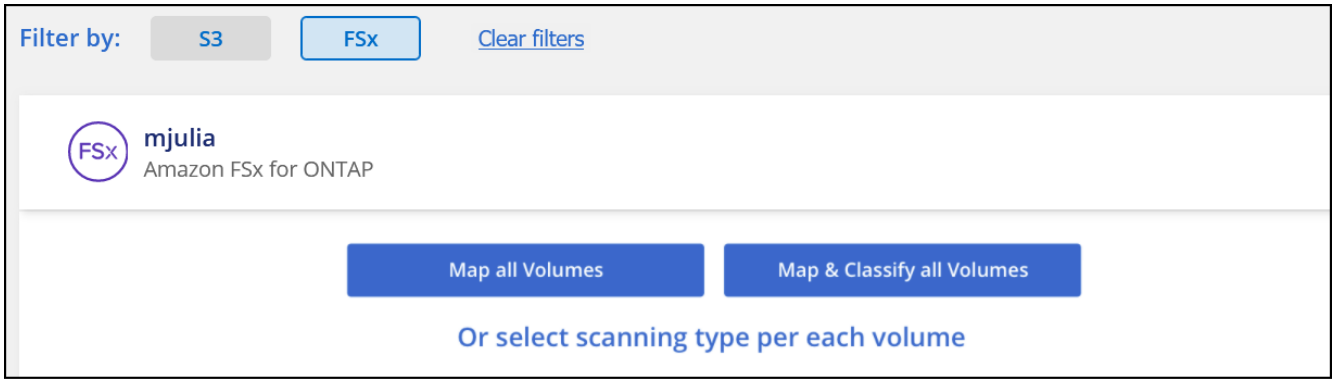
참고: FSx 볼륨을 스캔할 때 온프레미스 위치에 데이터 분류를 배포하는 것은 현재 지원되지 않습니다.

인스턴스에 인터넷 연결이 있는 한 데이터 분류 소프트웨어 업그레이드는 자동으로 수행됩니다.

시스템에서 데이터 분류를 활성화하세요

FSx for ONTAP 볼륨에 대한 데이터 분류를 활성화할 수 있습니다.

1. NetApp Console 에서 *거버넌스 > 분류*를 선택합니다.
2. 데이터 분류 메뉴에서 *구성*을 선택합니다.



3. 각 시스템의 볼륨을 스캔하는 방법을 선택합니다. "매핑 및 분류 스캔에 대해 알아보세요":
   - 모든 볼륨을 매핑하려면 *모든 볼륨 매핑*을 선택하세요.
   - 모든 볼륨을 매핑하고 분류하려면 *모든 볼륨 매핑 및 분류*를 선택하세요.

◦ 각 볼륨에 대한 스캐닝을 사용자 지정하려면 *또는 각 볼륨에 대한 스캐닝 유형 선택*을 선택한 다음 매핑 및 /또는 분류하려는 볼륨을 선택합니다.

4. 확인 대화 상자에서 *승인*을 선택하면 데이터 분류가 볼륨 검사를 시작합니다.

결과

데이터 분류는 시스템에서 선택한 볼륨에 대한 스캔을 시작합니다. 데이터 분류가 초기 스캔을 완료하는 즉시 규정 준수 대시보드에서 결과를 확인할 수 있습니다. 걸리는 시간은 데이터 양에 따라 다릅니다. 몇 분이나 몇 시간이 걸릴 수 있습니다. 구성 메뉴로 이동한 다음 시스템 구성을 선택하면 초기 검사의 진행 상황을 추적할 수 있습니다. 진행률 표시줄에서 각 검사의 진행 상황을 추적하세요. 진행률 표시줄 위에 마우스를 올리면 볼륨의 전체 파일 대비 검사된 파일 수를 확인할 수 있습니다.

- 기본적으로 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우 시스템은 볼륨의 파일을 검사하지 않습니다. 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문입니다. 마지막 액세스 시간이 재설정되어도 상관없다면 *또는 각 볼륨에 대한 스캐닝 유형을 선택하세요*를 선택하세요. 결과 페이지에는 데이터 분류가 권한에 관계없이 볼륨을 검사하도록 설정할 수 있는 설정이 있습니다.
- 데이터 분류는 볼륨 아래의 파일 공유를 하나만 스캔합니다. 볼륨에 여러 개의 주식이 있는 경우 다른 주식을 주식 그룹으로 별도로 스캔해야 합니다. "이 데이터 분류 제한에 대한 자세한 내용을 확인하세요." .

데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요.

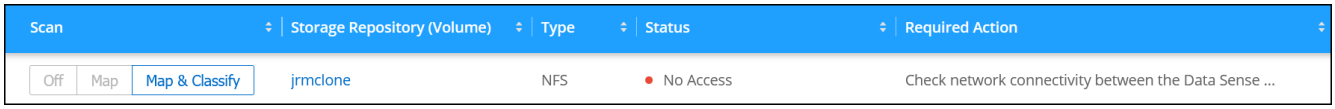
네트워킹, 보안 그룹 및 내보내기 정책을 확인하여 데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요.

CIFS 볼륨에 액세스할 수 있도록 데이터 분류에 CIFS 자격 증명을 제공해야 합니다.

단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 구성 페이지에서 *세부 정보 보기*를 선택하여 상태를 검토하고 오류를 수정하세요.

   예를 들어, 다음 이미지는 데이터 분류 인스턴스와 볼륨 간의 네트워크 연결 문제로 인해 볼륨 데이터 분류가 스캔할 수 없는 상황을 보여줍니다.



3. FSx for ONTAP 볼륨이 포함된 각 네트워크와 데이터 분류 인스턴스 사이에 네트워크 연결이 있는지 확인하세요.

   FSx for ONTAP 의 경우 데이터 분류는 콘솔과 동일한 지역의 볼륨만 스캔할 수 있습니다.

4. NFS 볼륨 내보내기 정책에 데이터 분류 인스턴스의 IP 주소가 포함되어 있는지 확인하여 각 볼륨의 데이터에 액세스할 수 있도록 합니다.
5. CIFS를 사용하는 경우 CIFS 볼륨을 스캔할 수 있도록 Active Directory 자격 증명을 사용하여 데이터 분류를 제공합니다.

   a. 데이터 분류 메뉴에서 *구성*을 선택합니다.

   b. 각 시스템에 대해 *CIFS 자격 증명 편집*을 선택하고 데이터 분류가 시스템의 CIFS 볼륨에 액세스하는 데

필요한 사용자 이름과 비밀번호를 입력합니다.

자격 증명은 읽기 전용일 수 있지만, 관리자 자격 증명을 제공하면 데이터 분류에서 높은 권한이 필요한 모든 데이터를 읽을 수 있습니다. 자격 증명은 데이터 분류 인스턴스에 저장됩니다.

데이터 분류 검사를 통해 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 CIFS에서 쓰기 속성 권한이나 NFS에서 쓰기 권한이 있는 것이 좋습니다. 가능하다면 Active Directory 사용자를 모든 파일에 대한 권한이 있는 조직의 상위 그룹의 일부로 구성하세요.

자격 증명을 입력한 후 모든 CIFS 볼륨이 성공적으로 인증되었다는 메시지가 표시됩니다.

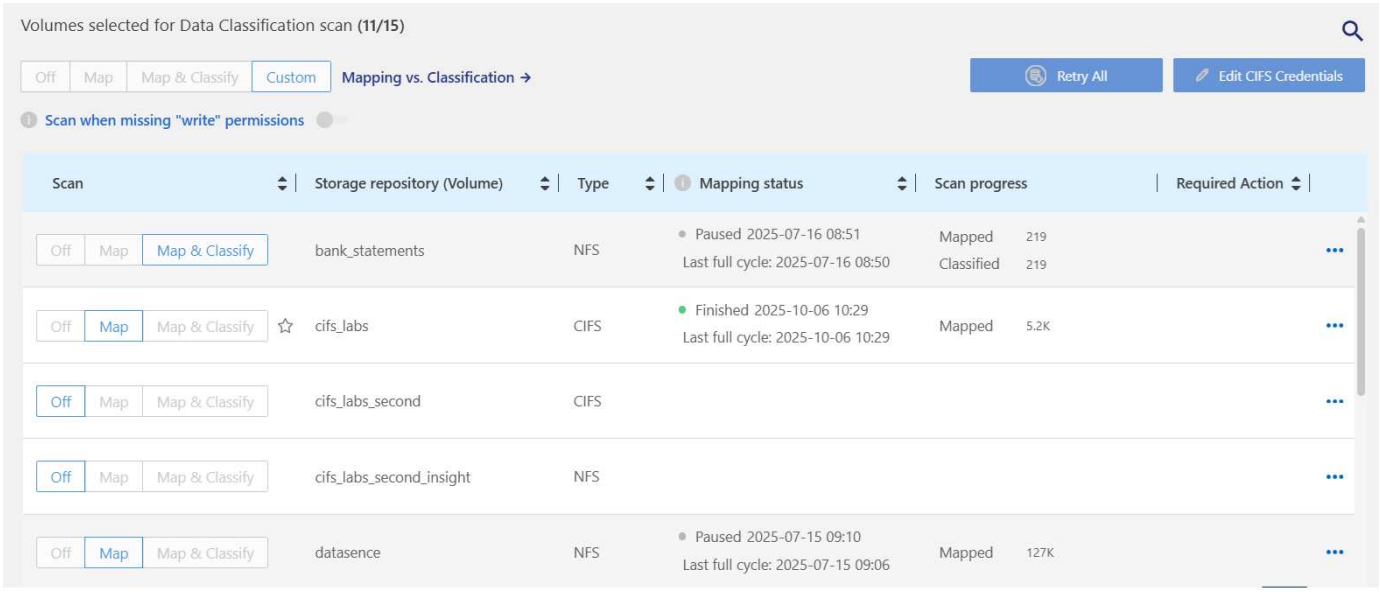
볼륨에서 스캔 활성화 및 비활성화

구성 페이지에서 언제든지 모든 시스템의 검사를 시작하거나 중지할 수 있습니다. 매핑 전용 스캔에서 매핑 및 분류 스캔으로 전환할 수도 있고, 그 반대로도 가능합니다. 시스템의 모든 볼륨을 스캔하는 것이 좋습니다.

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 선택한 경우에만 자동으로 스캔됩니다. 제목 영역에서 사용자 지정 또는 *끄기*로 설정하면 시스템에 추가하는 각 새 볼륨에 대해 매핑 및/또는 전체 스캐닝을 활성화해야 합니다.

기본적으로 페이지 상단의 쓰기 권한이 없는 경우 검사 스위치는 비활성화되어 있습니다. 즉, 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우, 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문에 시스템은 파일을 검사하지 않습니다. 마지막 접근 시간이 재설정되어도 상관없다면 스위치를 켜면 권한에 관계없이 모든 파일이 검사됩니다. "자세히 알아보기".

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 지정한 경우에만 자동으로 스캔됩니다. 모든 볼륨에 대한 설정이 사용자 지정 또는 *끄기*인 경우, 새로 추가하는 각 볼륨에 대해 수동으로 스캐닝을 활성화해야 합니다.



단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 시스템을 선택한 다음 *구성*을 선택하세요.

3. 모든 볼륨에 대한 검사를 활성화하거나 비활성화하려면 모든 볼륨 위의 제목에서 맵, 맵 및 분류 또는 끄기를 선택합니다.

   개별 볼륨에 대한 검사를 활성화하거나 비활성화하려면 목록에서 볼륨을 찾은 다음 볼륨 이름 옆에 있는 매핑, 매핑 및 분류 또는 끄기를 선택합니다.

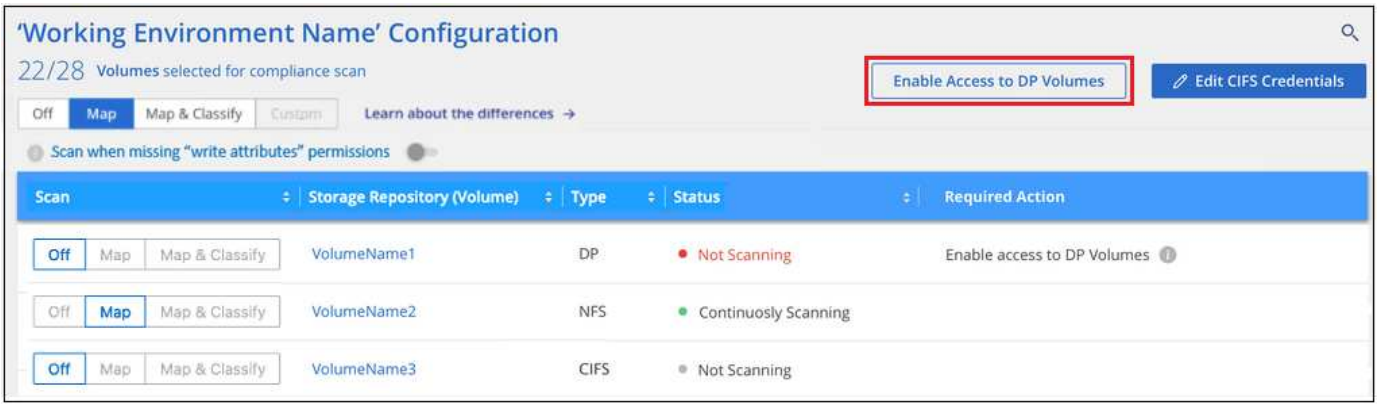
결과

스캐닝을 활성화하면 데이터 분류가 시스템에서 선택한 볼륨을 스캐닝하기 시작합니다. 데이터 분류가 스캔을 시작하자마자 규정 준수 대시보드에 결과가 나타나기 시작합니다. 검사 완료 시간은 데이터 양에 따라 달라지며, 몇 분에서 몇 시간까지 걸릴 수 있습니다.

데이터 보호 볼륨 스캔

기본적으로 데이터 보호(DP) 볼륨은 외부에 노출되지 않고 데이터 분류에서 액세스할 수 없으므로 스캔되지 않습니다. 이는 FSx for ONTAP 파일 시스템의 SnapMirror 작업을 위한 대상 볼륨입니다.

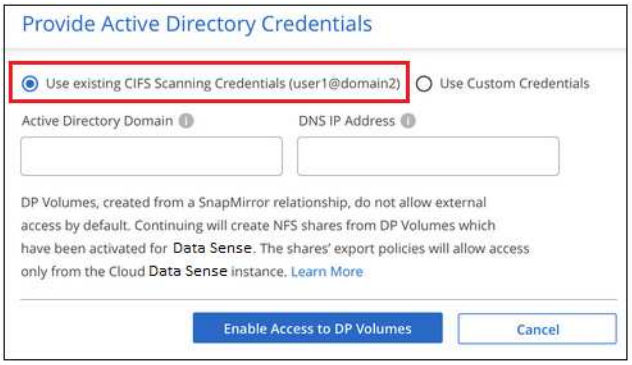
처음에 볼륨 목록은 이러한 볼륨을 유형 **DP**, 상태 스캔 안 함 및 필요한 작업 *DP 볼륨에 대한 액세스 활성화*로 식별합니다.

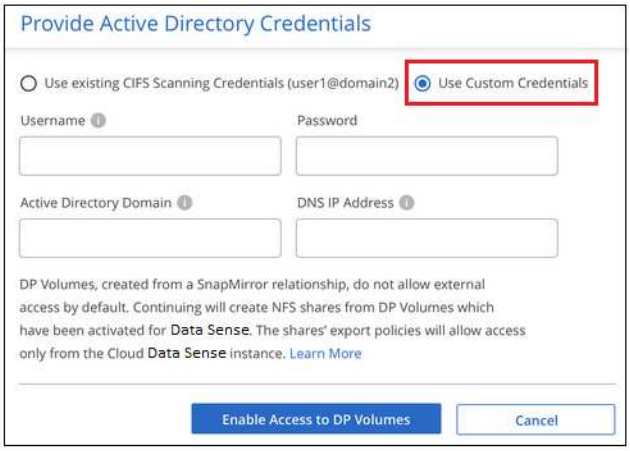


단계

다음 데이터 보호 볼륨을 스캔하려면 다음을 수행하세요.

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 페이지 상단에서 *DP 볼륨에 대한 액세스 활성화*를 선택합니다.
3. 확인 메시지를 검토하고 *DP 볼륨에 대한 액세스 활성화*를 다시 선택합니다.
   - ONTAP 파일 시스템용 소스 FSx에서 원래 NFS 볼륨으로 생성된 볼륨이 활성화됩니다.
   - ONTAP 파일 시스템용 소스 FSx에서 CIFS 볼륨으로 처음 생성된 볼륨의 경우 해당 DP 볼륨을 스캔하려면 CIFS 자격 증명을 입력해야 합니다. 데이터 분류가 CIFS 볼륨을 검색할 수 있도록 이미 Active Directory 자격 증명을 입력한 경우 해당 자격 증명을 사용하거나 다른 관리자 자격 증명 집합을 지정할 수 있습니다.





4. 스캔하려는 각 DP 볼륨을 활성화합니다.

결과

데이터 분류가 활성화되면 스캐닝을 위해 활성화된 각 DP 볼륨에서 NFS 공유가 생성됩니다. 공유 내보내기 정책은 데이터 분류 인스턴스에서만 액세스를 허용합니다.

처음에 DP 볼륨에 대한 액세스를 활성화했을 때 CIFS 데이터 보호 볼륨이 없었고 나중에 볼륨을 추가한 경우, 구성 페이지 상단에 **CIFS DP**에 대한 액세스 활성화 버튼이 나타납니다. 이 버튼을 선택하고 CIFS 자격 증명을 추가하여 이러한 CIFS DP 볼륨에 대한 액세스를 활성화합니다.

Active Directory 자격 증명은 첫 번째 CIFS DP 볼륨의 스토리지 VM에만 등록되므로 해당 SVM의 모든 DP 볼륨이 검사됩니다. 다른 SVM에 있는 볼륨에는 Active Directory 자격 증명이 등록되지 않으므로 해당 DP 볼륨은 검사되지 않습니다.

NetApp Data Classification 사용하여 Azure NetApp Files 볼륨 스캔

Azure NetApp Files 에 대한 NetApp Data Classification 시작하려면 몇 가지 단계를 완료하세요.

검사하려는 **Azure NetApp Files** 시스템을 검색하세요**.**

검사하려는 Azure NetApp Files 시스템이 시스템으로 NetApp Console 에 아직 없는 경우"시스템 페이지에 추가하세요" .

데이터 분류 인스턴스 배포

"데이터 분류 배포"아직 배포된 인스턴스가 없는 경우.

Azure NetApp Files 볼륨을 스캔할 때는 데이터 분류를 클라우드에 배포해야 하며, 스캔하려는 볼륨과 동일한 지역에 배포해야 합니다.

참고**:** Azure NetApp Files 볼륨을 스캔할 때 온-프레미스 위치에 데이터 분류를 배포하는 것은 현재 지원되지 않습니다.

시스템에서 데이터 분류를 활성화하세요

Azure NetApp Files 볼륨에서 데이터 분류를 활성화할 수 있습니다.

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.

2. 각 시스템의 볼륨을 스캔하는 방법을 선택합니다. "매핑 및 분류 스캔에 대해 알아보세요":
   - 모든 볼륨을 매핑하려면 *모든 볼륨 매핑*을 선택하세요.
   - 모든 볼륨을 매핑하고 분류하려면 *모든 볼륨 매핑 및 분류*를 선택하세요.
   - 각 볼륨에 대한 스캐닝을 사용자 지정하려면 *또는 각 볼륨에 대한 스캐닝 유형 선택*을 선택한 다음 매핑하거나 매핑하고 분류하려는 볼륨을 선택합니다.

     보다볼륨에서 스캔을 활성화하거나 비활성화합니다. 자세한 내용은.

3. 확인 대화 상자에서 *승인*을 선택합니다.

결과

데이터 분류는 시스템에서 선택한 볼륨에 대한 스캔을 시작합니다. 데이터 분류가 초기 스캔을 완료하면 규정 준수 대시보드에서 결과를 확인할 수 있습니다. 걸리는 시간은 데이터 양에 따라 다릅니다. 몇 분이나 몇 시간이 걸릴 수 있습니다. 구성 메뉴로 이동한 다음 시스템 구성을 선택하면 초기 검사의 진행 상황을 추적할 수 있습니다. 데이터 분류에서는 각 스캔에 대한 진행률 표시줄이 표시됩니다. 볼륨에 있는 전체 파일 수에 대한 검사된 파일 수를 확인하려면 진행률 표시줄 위에 마우스를 올려놓으세요.

- 기본적으로 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우 시스템은 볼륨의 파일을 검사하지 않습니다. 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문입니다. 마지막 액세스 시간이 재설정되어도 상관없다면 *또는 각 볼륨에 대한 스캐닝 유형을 선택하세요*를 선택하세요. 결과 페이지에는 데이터 분류가 권한에 관계없이 볼륨을 검사하도록 설정할 수 있는 설정이 있습니다.
- 데이터 분류는 볼륨 아래의 파일 공유를 하나만 스캔합니다. 볼륨에 여러 개의 주식이 있는 경우 다른 주식을 주식 그룹으로 별도로 스캔해야 합니다. "이 데이터 분류 제한 사항에 대해 알아보세요" .

데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요.

네트워킹, 보안 그룹 및 내보내기 정책을 확인하여 데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요. CIFS 볼륨에 액세스할 수 있도록 데이터 분류에 CIFS 자격 증명을 제공해야 합니다.

Azure NetApp Files 의 경우 데이터 분류는 콘솔과 동일한 지역의 볼륨만 검색할 수 있습니다.

체크리스트

- 데이터 분류 인스턴스와 Azure NetApp Files 볼륨이 포함된 각 네트워크 간에 네트워크 연결이 있는지 확인하세요.
- 다음 포트가 데이터 분류 인스턴스에 열려 있는지 확인하세요.
  - NFS의 경우 포트 111과 2049.

◦ CIFS의 경우 포트 139 및 445.

- NFS 볼륨 내보내기 정책에 데이터 분류 인스턴스의 IP 주소가 포함되어 있는지 확인하여 각 볼륨의 데이터에 액세스할 수 있도록 합니다.

단계

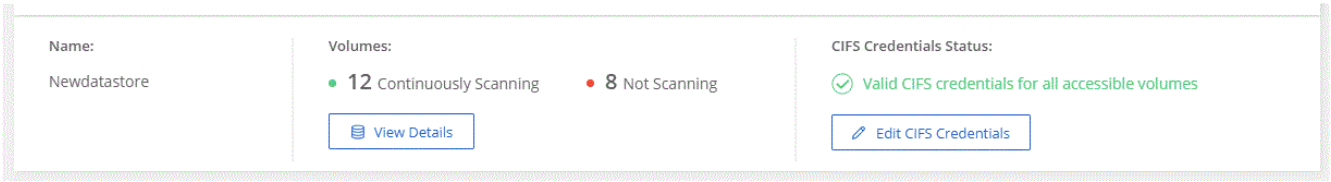
1. 데이터 분류 메뉴에서 *구성*을 선택합니다.

   a. CIFS(SMB)를 사용하는 경우 Active Directory 자격 증명이 올바른지 확인하세요. 각 시스템에 대해 *CIFS 자격 증명 편집*을 선택한 다음 데이터 분류가 시스템의 CIFS 볼륨에 액세스하는 데 필요한 사용자 이름과 비밀번호를 입력합니다.

      자격 증명은 읽기 전용일 수 있으며, 관리자 자격 증명을 제공하면 데이터 분류에서 높은 권한이 필요한 모든 데이터를 읽을 수 있습니다. 자격 증명은 데이터 분류 인스턴스에 저장됩니다.

      데이터 분류 검사를 통해 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 CIFS에서 쓰기 속성 권한이나 NFS에서 쓰기 권한이 있는 것이 좋습니다. 가능하다면 Active Directory 사용자를 모든 파일에 대한 권한이 있는 조직의 상위 그룹의 일부로 구성하세요.

      자격 증명을 입력한 후 모든 CIFS 볼륨이 성공적으로 인증되었다는 메시지가 표시됩니다.



2. 구성 페이지에서 *세부 정보 보기*를 선택하여 각 CIFS 및 NFS 볼륨의 상태를 검토합니다. 필요한 경우 네트워크 연결 문제 등의 오류를 수정하세요.

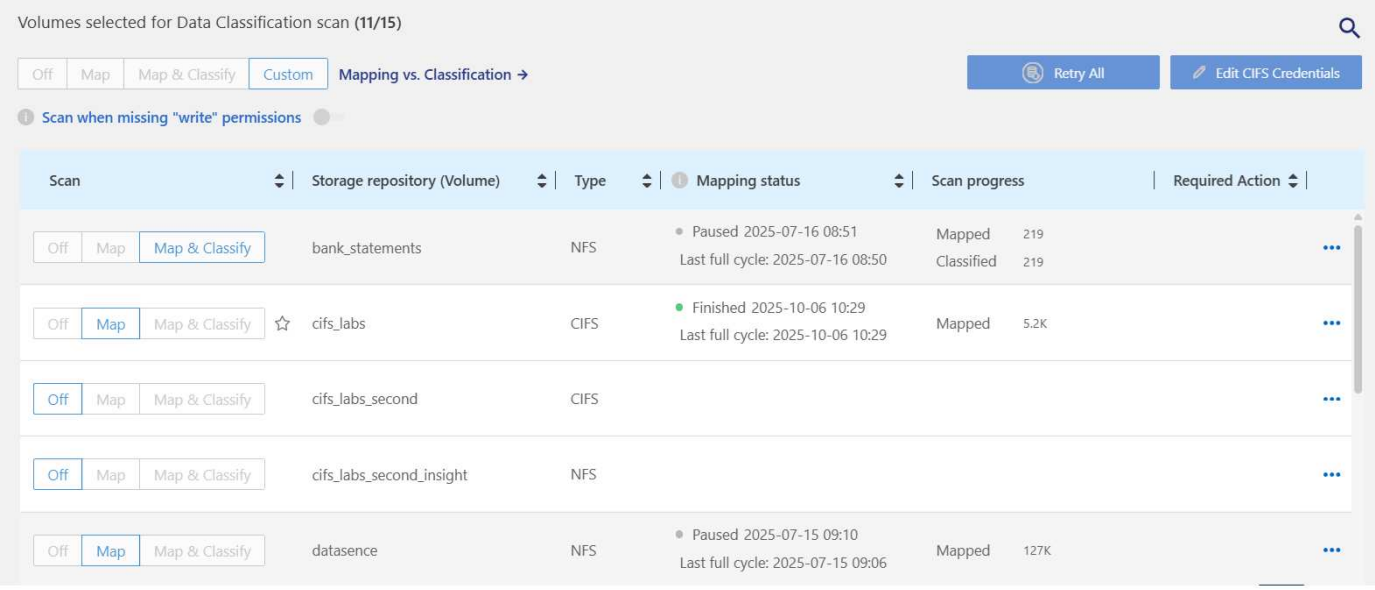
볼륨에서 스캔을 활성화하거나 비활성화합니다.

구성 페이지에서 언제든지 모든 시스템의 검사를 시작하거나 중지할 수 있습니다. 매핑 전용 스캔에서 매핑 및 분류 스캔으로 전환할 수도 있고, 그 반대로도 가능합니다. 시스템의 모든 볼륨을 스캔하는 것이 좋습니다.

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 선택한 경우에만 자동으로 스캔됩니다. 제목 영역에서 사용자 지정 또는 *끄기*로 설정하면 시스템에 추가하는 각 새 볼륨에 대해 매핑 및/또는 전체 스캐닝을 활성화해야 합니다.

기본적으로 페이지 상단의 쓰기 권한이 없는 경우 검사 스위치는 비활성화되어 있습니다. 즉, 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우, 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문에 시스템은 파일을 검사하지 않습니다. 마지막 접근 시간이 재설정되어도 상관없다면 스위치를 켜면 권한에 관계없이 모든 파일이 검사됩니다. "자세히 알아보기".

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 지정한 경우에만 자동으로 스캔됩니다. 모든 볼륨에 대한 설정이 사용자 지정 또는 *끄기*인 경우, 새로 추가하는 각 볼륨에 대해 수동으로 스캐닝을 활성화해야 합니다.



단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 시스템을 선택한 다음 *구성*을 선택하세요.
3. 모든 볼륨에 대한 검사를 활성화하거나 비활성화하려면 모든 볼륨 위의 제목에서 맵, 맵 및 분류 또는 끄기를 선택합니다.

   개별 볼륨에 대한 검사를 활성화하거나 비활성화하려면 목록에서 볼륨을 찾은 다음 볼륨 이름 옆에 있는 매핑, 매핑 및 분류 또는 끄기를 선택합니다.

결과

스캐닝을 활성화하면 데이터 분류가 시스템에서 선택한 볼륨을 스캐닝하기 시작합니다. 데이터 분류가 스캔을 시작하자마자 규정 준수 대시보드에 결과가 나타나기 시작합니다. 검사 완료 시간은 데이터 양에 따라 달라지며, 몇 분에서 몇 시간까지 걸릴 수 있습니다.

NetApp Data Classification 사용하여 Cloud Volumes ONTAP 및 온프레미스 ONTAP 볼륨 스캔

NetApp Data Classification 사용하여 Cloud Volumes ONTAP 및 온프레미스 ONTAP 볼륨 스캔을 시작하려면 몇 가지 단계를 완료하세요.

필수 조건

데이터 분류를 활성화하기 전에 지원되는 구성이 있는지 확인하세요.

- 인터넷을 통해 액세스할 수 있는 Cloud Volumes ONTAP 및 온프레미스 ONTAP 시스템을 스캔하는 경우 다음을 수행할 수 있습니다."클라우드에 데이터 분류 배포" 또는"인터넷 접속이 가능한 사내 위치에서" .
- 인터넷 접속이 불가능한 다크 사이트에 설치된 온프레미스 ONTAP 시스템을 스캔하는 경우 다음이 필요합니다. "인터넷 접속이 없는 동일한 온프레미스 위치에 데이터 분류를 배포합니다." . 이렇게 하려면 콘솔 에이전트를 동일한 온프레미스 위치에 배포해야 합니다.

데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요.

네트워킹, 보안 그룹 및 내보내기 정책을 확인하여 데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요. CIFS 볼륨에 액세스할 수 있도록 데이터 분류에 CIFS 자격 증명을 제공해야 합니다.

체크리스트

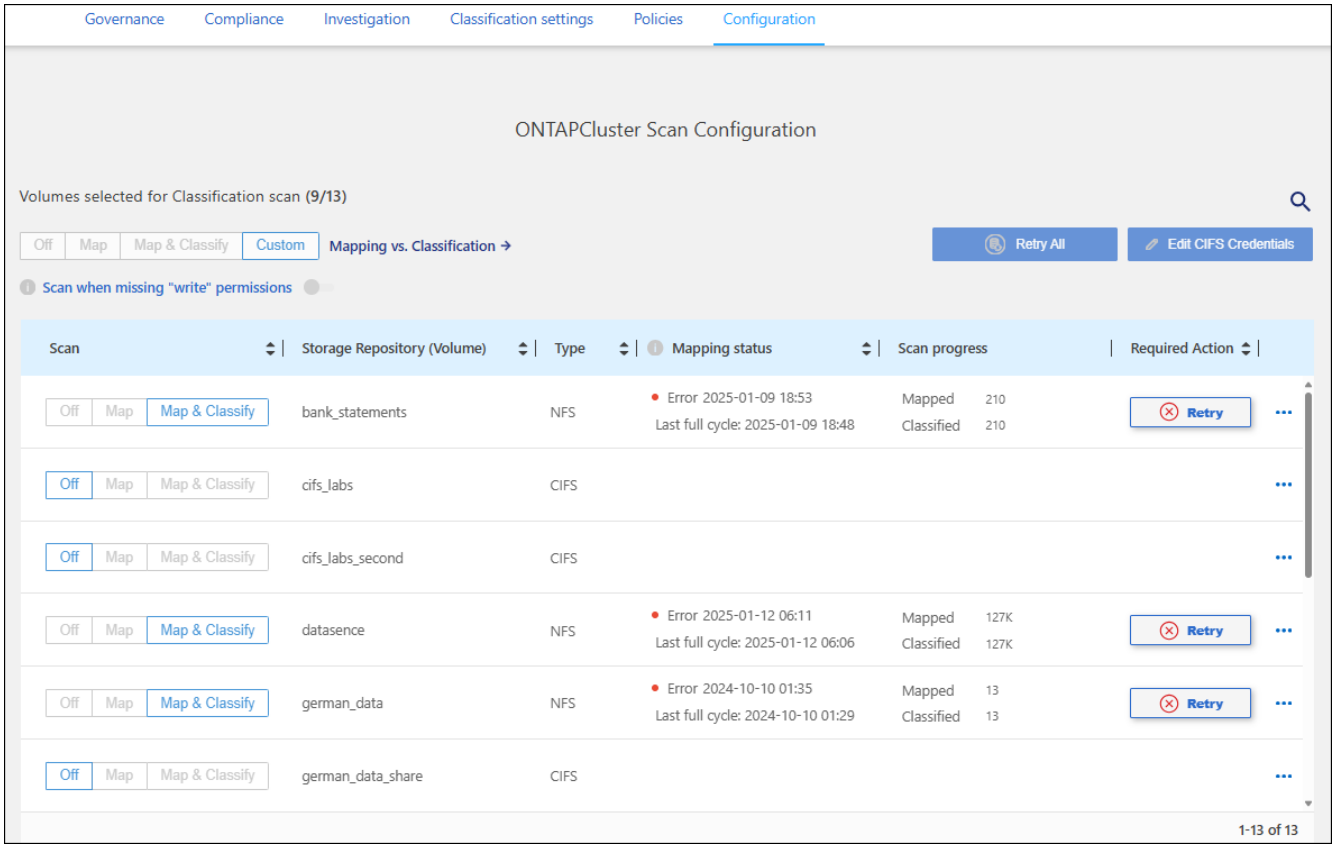
- 데이터 분류 인스턴스와 Cloud Volumes ONTAP 또는 온프레미스 ONTAP 클러스터의 볼륨이 포함된 각 네트워크 간에 네트워크 연결이 있는지 확인하세요.
- Cloud Volumes ONTAP 의 보안 그룹이 데이터 분류 인스턴스에서 들어오는 트래픽을 허용하는지 확인하세요.

  데이터 분류 인스턴스의 IP 주소에서 발생하는 트래픽에 대해 보안 그룹을 열거나, 가상 네트워크 내부의 모든 트래픽에 대해 보안 그룹을 열 수 있습니다.

- NFS 볼륨 내보내기 정책에 데이터 분류 인스턴스의 IP 주소가 포함되어 있는지 확인하여 각 볼륨의 데이터에 액세스할 수 있도록 합니다.

단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.



2. CIFS를 사용하는 경우 CIFS 볼륨을 스캔할 수 있도록 Active Directory 자격 증명을 사용하여 데이터 분류를 제공합니다. 각 시스템에 대해 *CIFS 자격 증명 편집*을 선택하고 데이터 분류가 시스템의 CIFS 볼륨에 액세스하는 데 필요한 사용자 이름과 비밀번호를 입력합니다.

   자격 증명은 읽기 전용일 수 있지만, 관리자 자격 증명을 제공하면 데이터 분류에서 높은 권한이 필요한 모든 데이터를 읽을 수 있습니다. 자격 증명은 데이터 분류 인스턴스에 저장됩니다.

   데이터 분류 검사를 통해 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 CIFS에서 쓰기

속성 권한이나 NFS에서 쓰기 권한이 있는 것이 좋습니다. 가능하다면 Active Directory 사용자를 모든 파일에 대한 권한이 있는 조직의 상위 그룹의 일부로 구성하세요.

자격 증명을 올바르게 입력한 경우 모든 CIFS 볼륨이 성공적으로 인증되었음을 확인하는 메시지가 표시됩니다.

3. 구성 페이지에서 *구성*을 선택하여 각 CIFS 및 NFS 볼륨의 상태를 검토하고 오류를 수정합니다.

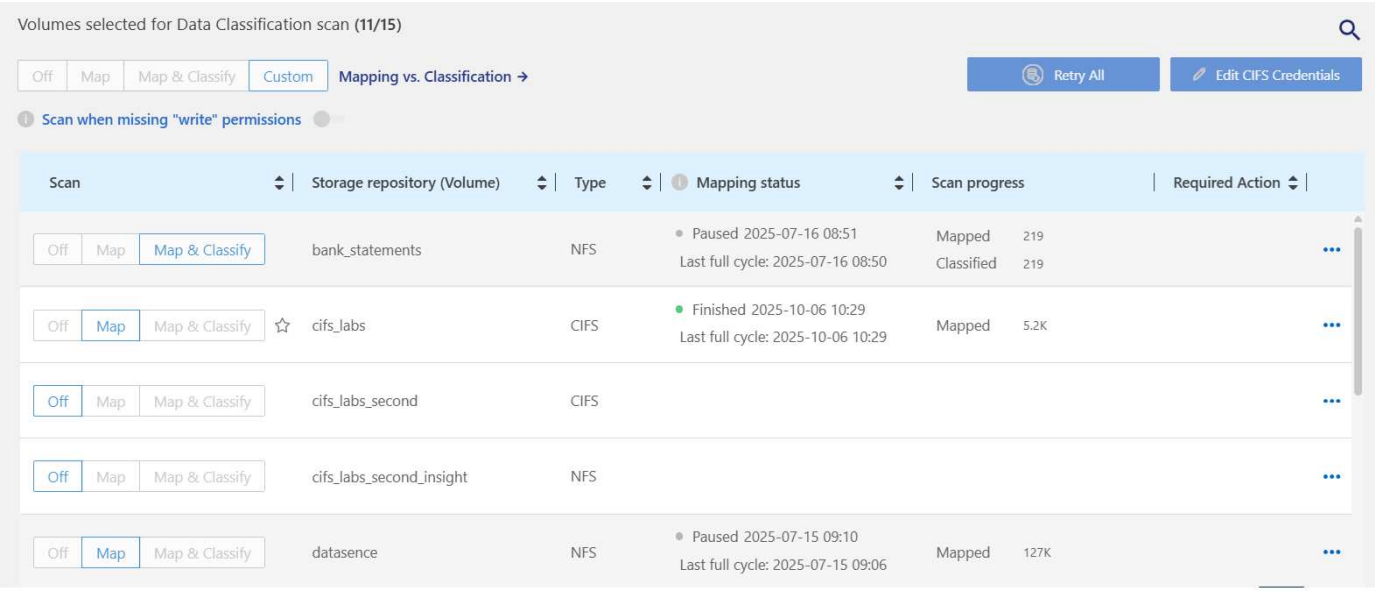
볼륨에서 스캔을 활성화하거나 비활성화합니다.

구성 페이지에서 언제든지 모든 시스템의 검사를 시작하거나 중지할 수 있습니다. 매핑 전용 스캔에서 매핑 및 분류 스캔으로 전환할 수도 있고, 그 반대로도 가능합니다. 시스템의 모든 볼륨을 스캔하는 것이 좋습니다.

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 선택한 경우에만 자동으로 스캔됩니다. 제목 영역에서 사용자 지정 또는 *끄기*로 설정하면 시스템에 추가하는 각 새 볼륨에 대해 매핑 및/또는 전체 스캐닝을 활성화해야 합니다.

기본적으로 페이지 상단의 쓰기 권한이 없는 경우 검사 스위치는 비활성화되어 있습니다. 즉, 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우, 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문에 시스템은 파일을 검사하지 않습니다. 마지막 접근 시간이 재설정되어도 상관없다면 스위치를 켜면 권한에 관계없이 모든 파일이 검사됩니다. "자세히 알아보기".

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 지정한 경우에만 자동으로 스캔됩니다. 모든 볼륨에 대한 설정이 사용자 지정 또는 *끄기*인 경우, 새로 추가하는 각 볼륨에 대해 수동으로 스캐닝을 활성화해야 합니다.



단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 시스템을 선택한 다음 *구성*을 선택하세요.
3. 모든 볼륨에 대한 검사를 활성화하거나 비활성화하려면 모든 볼륨 위의 제목에서 맵, 맵 및 분류 또는 끄기를 선택합니다.

   개별 볼륨에 대한 검사를 활성화하거나 비활성화하려면 목록에서 볼륨을 찾은 다음 볼륨 이름 옆에 있는 매핑, 매핑 및 분류 또는 끄기를 선택합니다.

결과

스캐닝을 활성화하면 데이터 분류가 시스템에서 선택한 볼륨을 스캐닝하기 시작합니다. 데이터 분류가 스캔을 시작하자마자 규정 준수 대시보드에 결과가 나타나기 시작합니다. 검사 완료 시간은 데이터 양에 따라 달라지며, 몇 분에서 몇 시간까지 걸릴 수 있습니다.

데이터 분류는 볼륨 아래의 파일 공유를 하나만 스캔합니다. 볼륨에 여러 개의 주식이 있는 경우 다른 주식을 주식 그룹으로 별도로 스캔해야 합니다. "이 데이터 분류 제한에 대한 자세한 내용을 확인하세요." .

NetApp Data Classification 사용하여 데이터베이스 스키마 스캔

NetApp Data Classification 사용하여 데이터베이스 스키마 스캔을 시작하려면 몇 가지 단계를 완료하세요.

필수 조건 검토

데이터 분류를 활성화하기 전에 지원되는 구성이 있는지 확인하려면 다음 필수 구성 요소를 검토하세요.

지원되는 데이터베이스

데이터 분류는 다음 데이터베이스에서 스키마를 스캔할 수 있습니다.

- Amazon 관계형 데이터베이스 서비스(Amazon RDS)
- 몽고디비
- MySQL
- 신탁
- 포스트그레스큐엘
- SAP 하나
- SQL 서버(MSSQL)

데이터베이스에서 통계 수집 기능을 *활성화*해야 합니다.

데이터베이스 요구 사항

데이터 분류 인스턴스에 연결된 모든 데이터베이스는 호스팅 위치에 관계없이 스캔할 수 있습니다. 데이터베이스에 연결하려면 다음 정보가 필요합니다.

- IP 주소 또는 호스트 이름
- 포트
- 서비스 이름(Oracle 데이터베이스에 액세스하는 경우에만 해당)
- 스키마에 대한 읽기 액세스를 허용하는 자격 증명

  사용자 이름과 비밀번호를 선택할 때는 스캔하려는 모든 스키마와 테이블에 대한 전체 읽기 권한이 있는 것을 선택하는 것이 중요합니다. 데이터 분류 시스템에 필요한 모든 권한을 갖춘 전담 사용자를 만드는 것이 좋습니다.

MongoDB의 경우 읽기 전용 관리자 역할이 필요합니다.

데이터 분류 인스턴스 배포

아직 배포된 인스턴스가 없으면 데이터 분류를 배포합니다.

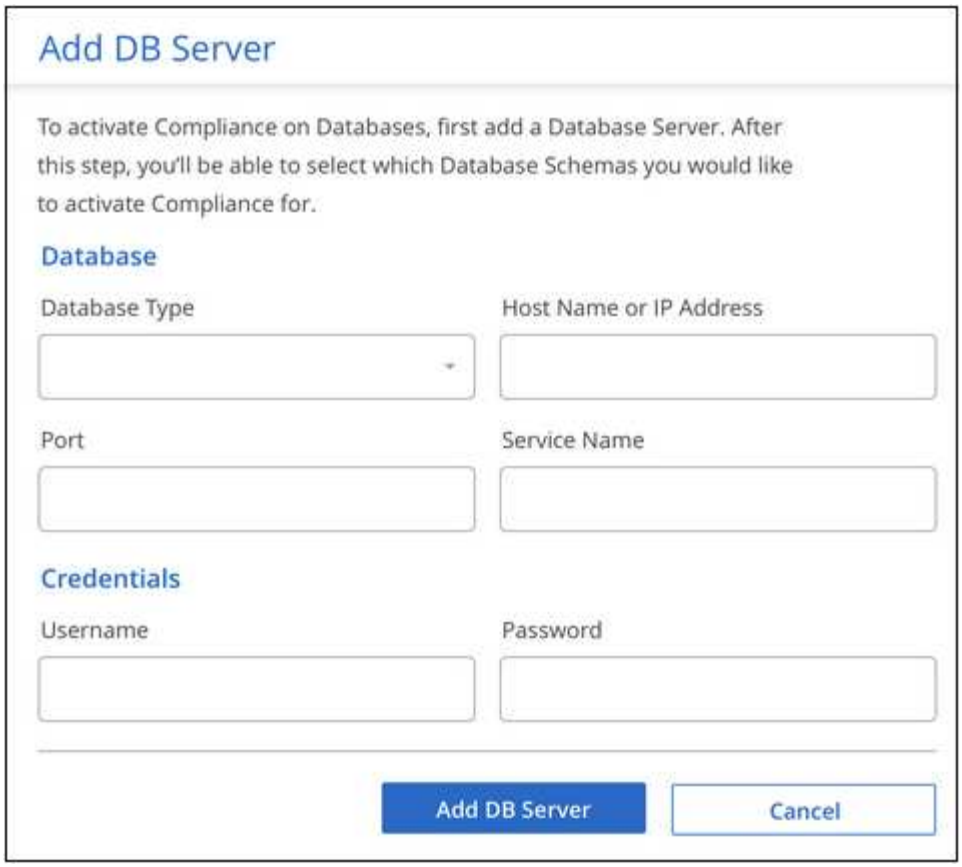
인터넷을 통해 접근 가능한 데이터베이스 스키마를 스캔하는 경우 다음을 수행할 수 있습니다."클라우드에 데이터 분류 배포" 또는"인터넷 접속이 가능한 온프레미스 위치에 데이터 분류를 배포합니다." .

인터넷 접속이 불가능한 다크 사이트에 설치된 데이터베이스 스키마를 스캔하는 경우 다음이 필요합니다."인터넷 접속이 없는 동일한 온프레미스 위치에 데이터 분류를 배포합니다." . 이를 위해서는 콘솔 에이전트가 동일한 온프레미스 위치에 배포되어야 합니다.

데이터베이스 서버 추가

스키마가 있는 데이터베이스 서버를 추가합니다.

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 구성 페이지에서 시스템 추가 > *데이터베이스 서버 추가*를 선택합니다.
3. 데이터베이스 서버를 식별하는 데 필요한 정보를 입력하세요.
    a. 데이터베이스 유형을 선택하세요.
    b. 데이터베이스에 연결하려면 포트와 호스트 이름 또는 IP 주소를 입력하세요.
    c. Oracle 데이터베이스의 경우 서비스 이름을 입력합니다.
    d. 데이터 분류가 서버에 액세스할 수 있도록 자격 증명을 입력하세요.
    e. *DB 서버 추가*를 선택합니다.



데이터베이스가 시스템 목록에 추가되었습니다.

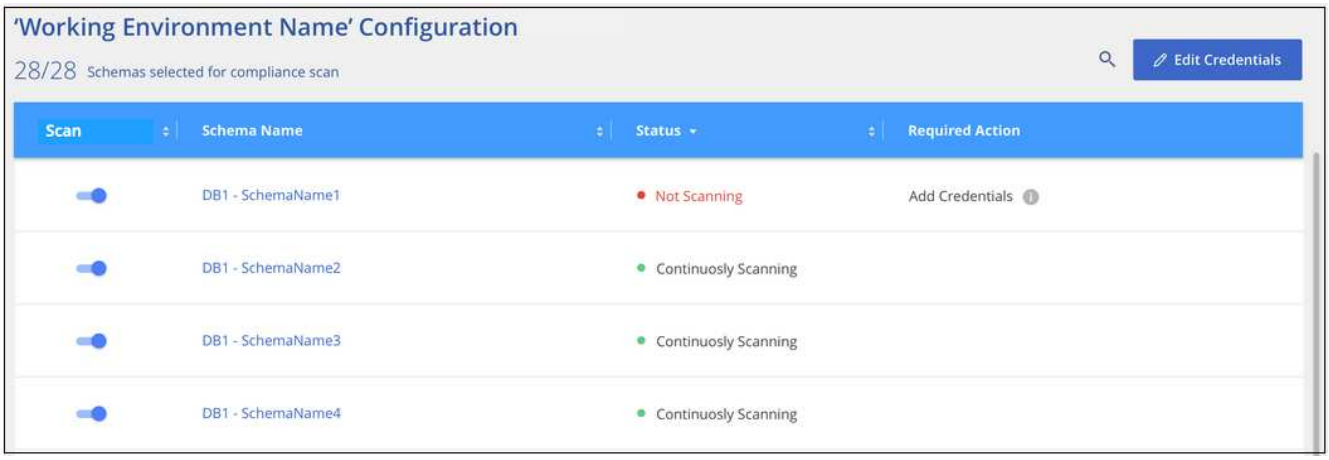
데이터베이스 스키마에 대한 스캔 활성화 및 비활성화

언제든지 스키마 전체 스캐닝을 중지하거나 시작할 수 있습니다.

데이터베이스 스키마에 대해 매핑 전용 스캔을 선택하는 옵션은 없습니다.

1. 구성 페이지에서 구성하려는 데이터베이스에 대한 구성 버튼을 선택합니다.

2. 슬라이더를 오른쪽으로 움직여 검사할 스키마를 선택합니다.



결과

데이터 분류는 활성화된 데이터베이스 스키마를 스캔하기 시작합니다. 구성 메뉴로 이동한 다음 시스템 구성을 선택하면 초기 검사의 진행 상황을 추적할 수 있습니다. 각 스캔의 진행 상황은 진행률 표시줄로 표시됩니다. 진행률 표시줄 위에 마우스를 올려 놓으면 볼륨에 있는 전체 파일 수에 대한 검사된 파일 수를 확인할 수 있습니다. 오류가 있는 경우 오류 수정에 필요한 작업과 함께 상태 열에 오류가 표시됩니다.

데이터 분류는 하루에 한 번씩 데이터베이스를 스캔합니다. 데이터베이스는 다른 데이터 소스처럼 지속적으로 스캔되지 않습니다.

NetApp Data Classification 사용하여 Google Cloud NetApp Volumes 스캔

NetApp Data Classification 시스템으로서 Google Cloud NetApp Volumes 지원합니다. Google Cloud NetApp Volumes 시스템을 스캔하는 방법을 알아보세요.

스캔하려는 **Google Cloud NetApp Volumes** 시스템을 검색하세요**.**

스캔하려는 Google Cloud NetApp Volumes 시스템이 NetApp Console 에 시스템으로 아직 없는 경우"시스템 페이지에 추가하세요" .

데이터 분류 인스턴스 배포

"데이터 분류 배포"아직 배포된 인스턴스가 없는 경우.

Google Cloud NetApp Volumes 스캔할 때는 데이터 분류를 클라우드에 배포해야 하며, 스캔하려는 볼륨과 동일한 지역에 배포해야 합니다.

참고**:** 현재 Google Cloud NetApp Volumes 스캔할 때 온프레미스 위치에 데이터 분류를 배포하는 것은 지원되지 않습니다.

시스템에서 데이터 분류를 활성화하세요

Google Cloud NetApp Volumes 시스템에서 데이터 분류를 활성화할 수 있습니다.

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 각 시스템의 볼륨을 스캔하는 방법을 선택합니다. "매핑 및 분류 스캔에 대해 알아보세요":
    - 모든 볼륨을 매핑하려면 *모든 볼륨 매핑*을 선택하세요.

- 모든 볼륨을 매핑하고 분류하려면 *모든 볼륨 매핑 및 분류*를 선택하세요.
- 각 볼륨에 대한 스캐닝을 사용자 지정하려면 *또는 각 볼륨에 대한 스캐닝 유형 선택*을 선택한 다음 매핑 및/또는 분류하려는 볼륨을 선택합니다.

  보다볼륨에서 스캔 활성화 및 비활성화 자세한 내용은.

3. 확인 대화 상자에서 *승인*을 선택합니다.

결과

데이터 분류는 시스템에서 선택한 볼륨에 대한 스캔을 시작합니다. 데이터 분류가 초기 스캔을 완료하면 규정 준수 대시보드에서 결과를 확인할 수 있습니다. 걸리는 시간은 데이터 양에 따라 다릅니다. 몇 분에서 몇 시간까지 걸립니다. 구성 메뉴의 시스템 구성 섹션에서 초기 검사의 진행 상황을 추적할 수 있습니다. 데이터 분류에서는 각 스캔에 대한 진행률 표시줄이 표시됩니다. 진행률 표시줄 위에 마우스를 올려 놓으면 볼륨의 전체 파일 대비 검사된 파일 수를 확인할 수 있습니다.

- 기본적으로 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우 시스템은 볼륨의 파일을 검사하지 않습니다. 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문입니다. 마지막 액세스 시간이 재설정되어도 상관없다면 *또는 각 볼륨에 대한 스캐닝 유형을 선택하세요*를 선택하세요. 결과 페이지에는 데이터 분류가 권한에 관계없이 볼륨을 검사하도록 설정할 수 있는 설정이 있습니다.
- 데이터 분류는 볼륨 아래의 파일 공유를 하나만 스캔합니다. 볼륨에 여러 개의 주식이 있는 경우 다른 주식을 주식 그룹으로 별도로 스캔해야 합니다. "이 데이터 분류 제한 사항에 대해 알아보세요".

데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요.

네트워킹, 보안 그룹 및 내보내기 정책을 확인하여 데이터 분류가 볼륨에 액세스할 수 있는지 확인하세요. CIFS 볼륨의 경우 CIFS 자격 증명을 사용하여 데이터 분류를 제공해야 합니다.

Google Cloud NetApp Volumes 의 경우 데이터 분류는 콘솔과 동일한 지역의 볼륨만 스캔할 수 있습니다.

체크리스트

- Google Cloud NetApp Volumes 대한 볼륨이 포함된 각 네트워크와 데이터 분류 인스턴스 사이에 네트워크 연결이 있는지 확인하세요.
- 다음 포트가 데이터 분류 인스턴스에 열려 있는지 확인하세요.
  - NFS의 경우 포트 111과 2049.
  - CIFS의 경우 포트 139 및 445.
- NFS 볼륨 내보내기 정책에 데이터 분류 인스턴스의 IP 주소가 포함되어 있는지 확인하여 각 볼륨의 데이터에 액세스할 수 있도록 합니다.

단계

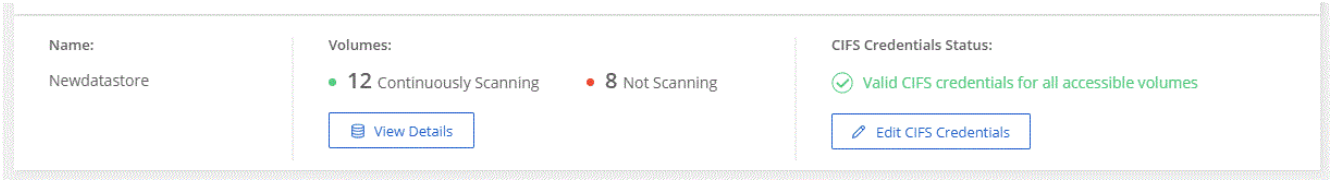
1. 데이터 분류 메뉴에서 *구성*을 선택합니다.

   a. CIFS(SMB)를 사용하는 경우 Active Directory 자격 증명이 올바른지 확인하세요. 각 시스템에 대해 *CIFS 자격 증명 편집*을 선택한 다음 데이터 분류가 시스템의 CIFS 볼륨에 액세스하는 데 필요한 사용자 이름과 비밀번호를 입력합니다.

      자격 증명은 읽기 전용일 수 있지만, 관리자 자격 증명을 제공하면 데이터 분류에서 높은 권한이 필요한 모든 데이터를 읽을 수 있습니다. 자격 증명은 데이터 분류 인스턴스에 저장됩니다.

데이터 분류 검사를 통해 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 CIFS에서 쓰기 속성 권한이나 NFS에서 쓰기 권한이 있는 것이 좋습니다. 가능하다면 Active Directory 사용자를 모든 파일에 대한 권한이 있는 조직의 상위 그룹의 일부로 구성하세요.

자격 증명을 입력한 후 모든 CIFS 볼륨이 성공적으로 인증되었다는 메시지가 표시됩니다.



2. 구성 페이지에서 *세부 정보 보기*를 선택하여 각 CIFS 및 NFS 볼륨의 상태를 검토하고 오류를 수정합니다.

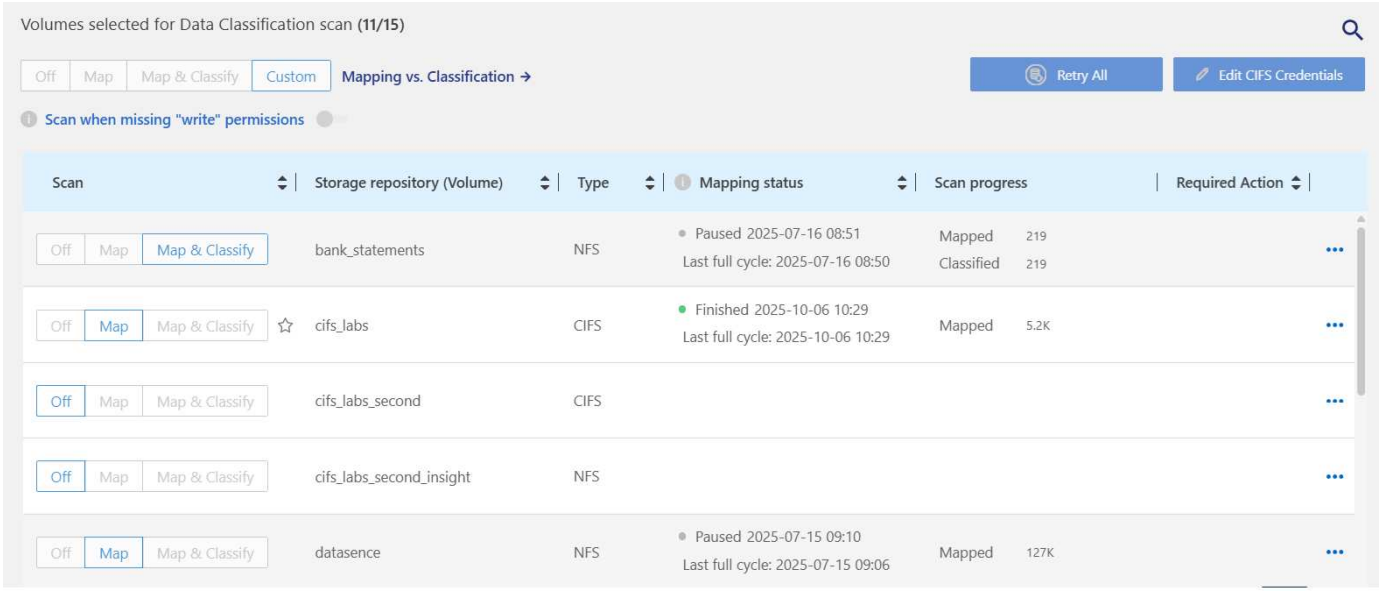
볼륨에서 스캔 활성화 및 비활성화

구성 페이지에서 언제든지 모든 시스템의 검사를 시작하거나 중지할 수 있습니다. 매핑 전용 스캔에서 매핑 및 분류 스캔으로 전환할 수도 있고, 그 반대로도 가능합니다. 시스템의 모든 볼륨을 스캔하는 것이 좋습니다.

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 선택한 경우에만 자동으로 스캔됩니다. 제목 영역에서 사용자 지정 또는 *끄기*로 설정하면 시스템에 추가하는 각 새 볼륨에 대해 매핑 및/또는 전체 스캐닝을 활성화해야 합니다.

기본적으로 페이지 상단의 쓰기 권한이 없는 경우 검사 스위치는 비활성화되어 있습니다. 즉, 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우, 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문에 시스템은 파일을 검사하지 않습니다. 마지막 접근 시간이 재설정되어도 상관없다면 스위치를 켜면 권한에 관계없이 모든 파일이 검사됩니다. "자세히 알아보기".

시스템에 추가된 새 볼륨은 제목 영역에서 지도 또는 지도 및 분류 설정을 지정한 경우에만 자동으로 스캔됩니다. 모든 볼륨에 대한 설정이 사용자 지정 또는 *끄기*인 경우, 새로 추가하는 각 볼륨에 대해 수동으로 스캐닝을 활성화해야 합니다.



단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 시스템을 선택한 다음 *구성*을 선택하세요.
3. 모든 볼륨에 대한 검사를 활성화하거나 비활성화하려면 모든 볼륨 위의 제목에서 맵, 맵 및 분류 또는 끄기를 선택합니다.

   개별 볼륨에 대한 검사를 활성화하거나 비활성화하려면 목록에서 볼륨을 찾은 다음 볼륨 이름 옆에 있는 매핑, 매핑 및 분류 또는 끄기를 선택합니다.

결과

스캐닝을 활성화하면 데이터 분류가 시스템에서 선택한 볼륨을 스캐닝하기 시작합니다. 데이터 분류가 스캔을 시작하자마자 규정 준수 대시보드에 결과가 나타나기 시작합니다. 검사 완료 시간은 데이터 양에 따라 달라지며, 몇 분에서 몇 시간까지 걸릴 수 있습니다.

NetApp Data Classification 사용하여 파일 공유 스캔

파일 공유를 스캔하려면 먼저 NetApp Data Classification 에서 파일 공유 그룹을 만들어야 합니다. 파일 공유 그룹은 온프레미스 또는 클라우드에서 호스팅되는 NFS 또는 CIFS(SMB) 공유를 위한 것입니다.

데이터 분류 핵심 버전에서는 NetApp 아닌 파일 공유에서 데이터를 스캔하는 기능이 지원되지 않습니다.

필수 조건

데이터 분류를 활성화하기 전에 지원되는 구성이 있는지 확인하려면 다음 필수 구성 요소를 검토하세요.

- 공유는 클라우드나 온프레미스 등 어디에서나 호스팅될 수 있습니다. 이전 NetApp 7-Mode 스토리지 시스템의 CIFS 공유는 파일 공유로 스캔될 수 있습니다.
  - 데이터 분류는 7-Mode 시스템에서 권한이나 "마지막 액세스 시간"을 추출할 수 없습니다.
  - 7-Mode 시스템에서 일부 Linux 버전과 CIFS 공유 간에 알려진 문제로 인해 NTLM 인증이 활성화된 SMBv1만 사용하도록 공유를 구성해야 합니다.
- 데이터 분류 인스턴스와 공유 간에 네트워크 연결이 필요합니다.
- DFS(분산 파일 시스템) 공유를 일반 CIFS 공유로 추가할 수 있습니다. 데이터 분류에서는 공유가 단일 CIFS 공유로 결합된 여러 서버/볼륨에 기반을 두고 있다는 사실을 인식하지 못하기 때문에 메시지가 실제로는 다른 서버/볼륨에 있는 폴더/공유 중 하나에만 적용되는 경우에도 공유에 대한 권한 또는 연결 오류가 발생할 수 있습니다.
- CIFS(SMB) 공유의 경우 공유에 대한 읽기 액세스 권한을 제공하는 Active Directory 자격 증명이 있는지 확인하세요. 데이터 분류에서 높은 권한이 필요한 데이터를 스캔해야 하는 경우 관리자 자격 증명이 선호됩니다.

  데이터 분류 검사를 통해 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 CIFS에서 쓰기 속성 권한이나 NFS에서 쓰기 권한이 있는 것이 좋습니다. 가능하다면 Active Directory 사용자를 모든 파일에 대한 권한이 있는 조직의 상위 그룹의 일부로 구성하세요.
- 그룹 내의 모든 CIFS 파일 공유는 동일한 Active Directory 자격 증명을 사용해야 합니다.
- NFS와 CIFS(Kerberos 또는 NTLM 사용) 공유를 혼합할 수 있습니다. 그룹에 주식을 별도로 추가해야 합니다. 즉, 프로토콜당 한 번씩, 총 두 번 프로세스를 완료해야 합니다.

◦ CIFS 인증 유형(Kerberos 및 NTLM)을 혼합하여 파일 공유 그룹을 만들 수 없습니다.

- Kerberos 인증을 사용하는 CIFS를 사용하는 경우 제공된 IP 주소가 데이터 분류에 액세스할 수 있는지 확인하세요. IP 주소에 접근할 수 없으면 파일 공유를 추가할 수 없습니다.

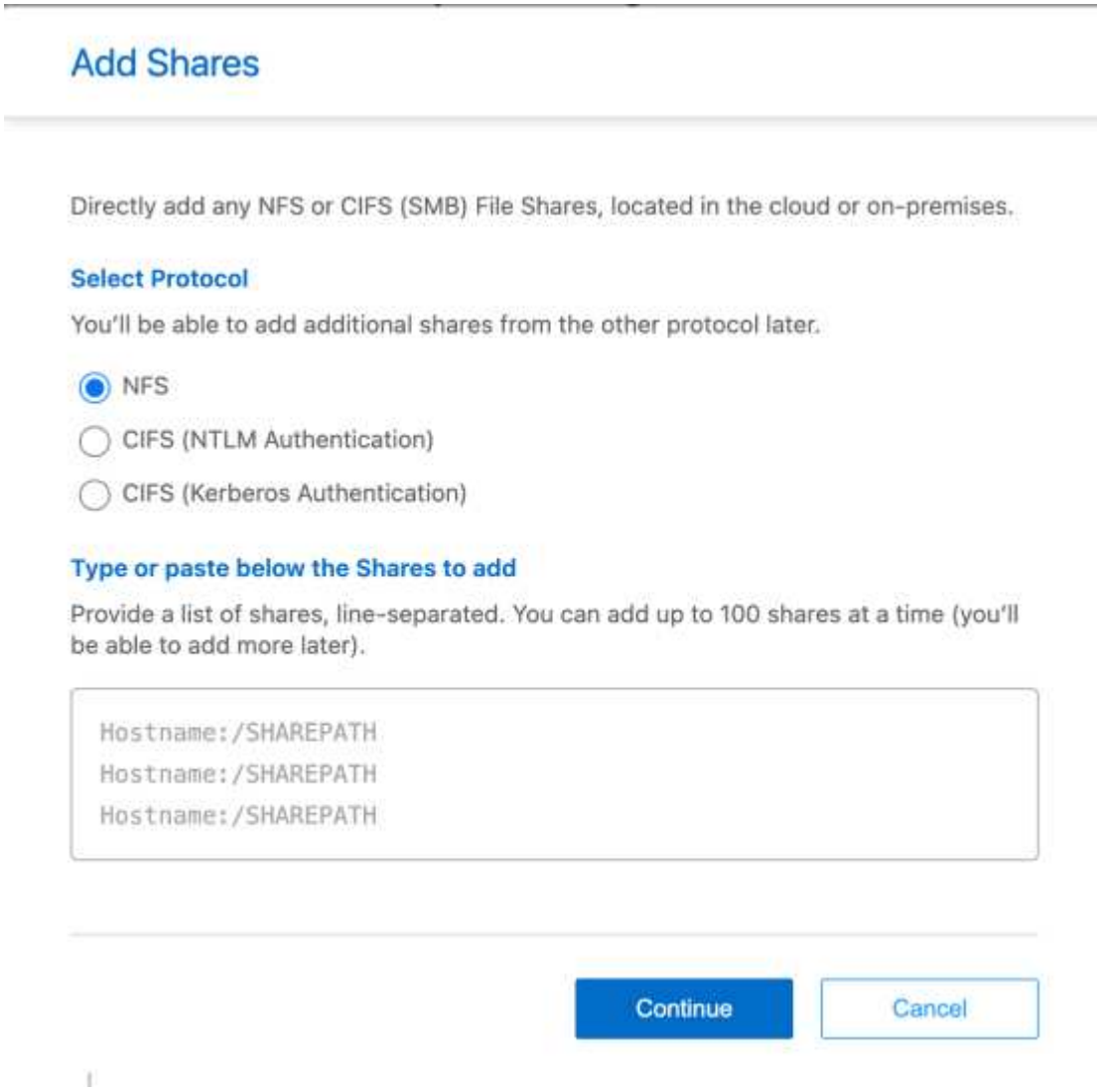
파일 공유 그룹 만들기

그룹에 파일 공유를 추가할 때는 다음 형식을 사용해야 합니다. `<host_name>:/<share_path>` .

파일 공유를 개별적으로 추가할 수도 있고, 검사하려는 파일 공유를 줄로 구분하여 나열하여 입력할 수도 있습니다. 한 번에 최대 100개의 주식을 추가할 수 있습니다.

단계

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 구성 페이지에서 시스템 추가 > *파일 공유 그룹 추가*를 선택합니다.
3. 파일 공유 그룹 추가 대화 상자에서 공유 그룹의 이름을 입력한 다음 *계속*을 선택합니다.
4. 추가할 파일 공유에 대한 프로토콜을 선택하세요.



a. NTLM 인증을 사용하여 CIFS 공유를 추가하는 경우 Active Directory 자격 증명을 입력하여 CIFS 볼륨에

액세스합니다. 읽기 전용 자격 증명도 지원되지만 관리자 자격 증명을 사용하여 전체 액세스 권한을 제공하는 것이 좋습니다. 저장을 선택하세요.

5. 검사하려는 파일 공유를 추가합니다(한 줄에 파일 공유 하나씩). 그런 다음 계속을 선택하세요.
6. 확인 대화 상자에는 추가된 주식 수가 표시됩니다.

   대화 상자에 추가할 수 없는 공유가 나열되면 이 정보를 캡처하여 문제를 해결하세요. 문제가 명명 규칙과 관련된 경우, 수정된 이름으로 공유를 다시 추가할 수 있습니다.

7. 볼륨에 대한 스캐닝을 구성합니다.
   - 파일 공유에서 매핑 전용 검사를 활성화하려면 *매핑*을 선택합니다.
   - 파일 공유에 대한 전체 검사를 활성화하려면 *매핑 및 분류*를 선택하세요.
   - 파일 공유에서 스캐닝을 비활성화하려면 *끄기*를 선택하세요.

기본적으로 페이지 상단의 **"쓰기 속성"** 권한이 없는 경우 검사 스위치는 비활성화되어 있습니다. 즉, 데이터 분류에 CIFS에서 쓰기 속성 권한이 없거나 NFS에서 쓰기 권한이 없는 경우, 데이터 분류는 "마지막 액세스 시간"을 원래 타임스탬프로 되돌릴 수 없기 때문에 시스템은 파일을 검사하지 않습니다. + "쓰기 속성" 권한이 없는 경우 검사*를 *켜기*로 전환하면 검사에서 마지막으로 액세스한 시간을 재설정하고 권한에 관계없이 모든 파일을 검사합니다. + 마지막으로 액세스한 타임스탬프에 대해 자세히 알아보려면 다음을 참조하세요."데이터 분류의 데이터 소스에서 수집된 메타데이터" .

결과

데이터 분류는 추가한 파일 공유에 있는 파일의 스캔을 시작합니다. 당신은 할 수 있습니다스캐닝 진행 상황을 추적하세요 대시보드에서 검사 결과를 확인하세요.

Kerberos 인증을 사용하는 CIFS 구성에 대한 스캔이 성공적으로 완료되지 않으면 구성 탭에서 오류를 확인하세요.

파일 공유 그룹 편집

파일 공유 그룹을 만든 후에는 CIFS 프로토콜을 편집하거나 파일 공유를 추가 및 제거할 수 있습니다.

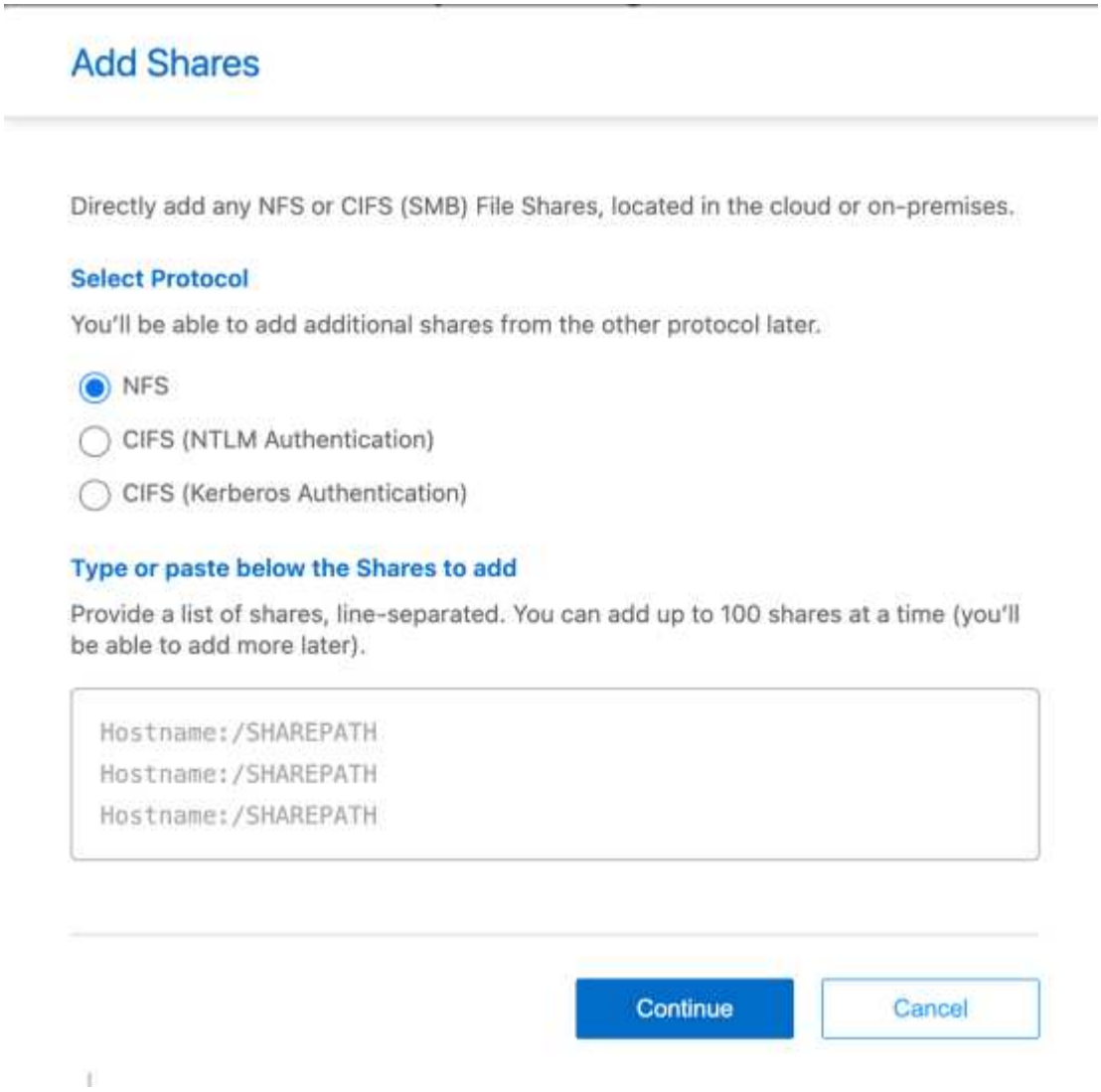
CIFS 프로토콜 구성 편집

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 구성 페이지에서 수정하려는 파일 공유 그룹을 선택합니다.
3. **CIFS** 자격 증명 편집을 선택합니다.

4. 인증 방법을 선택하세요: **NTLM** 또는 **Kerberos**.
5. Active Directory 사용자 이름과 암호를 입력합니다.
6. 저장을 선택하여 프로세스를 완료하세요.

스캔에 파일 공유 추가

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 구성 페이지에서 수정하려는 파일 공유 그룹을 선택합니다.
3. **+** 공유 추가를 선택하세요.
4. 추가할 파일 공유에 대한 프로토콜을 선택하세요.



이미 구성한 프로토콜에 파일 공유를 추가하는 경우 변경할 필요가 없습니다.

두 번째 프로토콜을 사용하여 파일 공유를 추가하는 경우 다음에서 자세히 설명한 대로 인증을 올바르게 구성했는지 확인하십시오."전제 조건" .

5. 형식을 사용하여 검사하려는 파일 공유를 추가합니다(줄당 파일 공유 하나). `<host_name>:/<share_path>` .
6. 계속을 선택하여 파일 공유 추가를 완료합니다.

스캔에서 파일 공유 제거

1. 데이터 분류 메뉴에서 *구성*을 선택합니다.
2. 파일 공유를 제거할 시스템을 선택하세요.
3. *구성*을 선택하세요.
4. 구성 페이지에서 작업을 선택하세요. ••• 제거하려는 파일 공유에 대해.
5. 작업 메뉴에서 *공유 제거*를 선택합니다.

스캐닝 진행 상황을 추적하세요

초기 스캔의 진행 상황을 추적할 수 있습니다.

1. 구성 메뉴를 선택하세요.
2. 시스템 구성을 선택하세요.
3. 저장소의 경우, 검사 진행률 열을 확인하여 상태를 확인하세요.

NetApp Data Classification 사용하여 StorageGRID 데이터 스캔

NetApp Data Classification 사용하여 StorageGRID 내에서 직접 데이터 스캔을 시작하려면 몇 가지 단계를 완료하세요.

StorageGRID 요구 사항 검토

데이터 분류를 활성화하기 전에 지원되는 구성이 있는지 확인하려면 다음 필수 구성 요소를 검토하세요.

- 개체 스토리지 서비스에 연결하려면 엔드포인트 URL이 필요합니다.
- 데이터 분류가 버킷에 액세스할 수 있도록 StorageGRID 에서 액세스 키와 비밀 키가 필요합니다.

데이터 분류 인스턴스 배포

아직 배포된 인스턴스가 없으면 데이터 분류를 배포합니다.

인터넷을 통해 액세스할 수 있는 StorageGRID 의 데이터를 스캔하는 경우 다음을 수행할 수 있습니다."클라우드에 데이터 분류 배포" 또는"인터넷 접속이 가능한 온프레미스 위치에 데이터 분류를 배포합니다." .

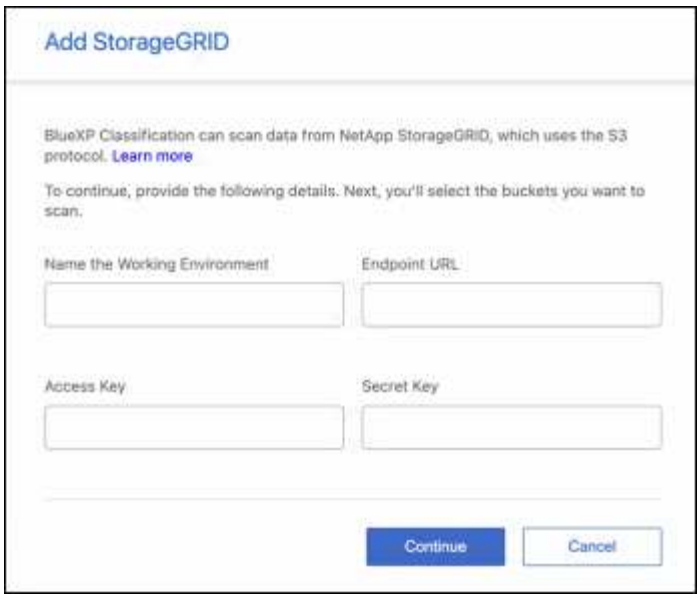
인터넷 접속이 불가능한 어두운 장소에 설치된 StorageGRID 에서 데이터를 스캔하는 경우 다음이 필요합니다."인터넷 접속이 없는 동일한 온프레미스 위치에 데이터 분류를 배포합니다." . 이를 위해서는 콘솔 에이전트가 동일한 온프레미스 위치에 배포되어야 합니다.

데이터 분류에 StorageGRID 서비스 추가

StorageGRID 서비스를 추가합니다.

단계

1. 데이터 분류 메뉴에서 구성 옵션을 선택합니다.
2. 구성 페이지에서 시스템 추가 > * StorageGRID 추가*를 선택합니다.
3. StorageGRID 서비스 추가 대화 상자에서 StorageGRID 서비스에 대한 세부 정보를 입력하고 *계속*을 선택합니다.
   a. 시스템에 사용할 이름을 입력하세요. 이 이름은 연결하려는 StorageGRID 서비스의 이름을 반영해야 합니다.
   b. 개체 스토리지 서비스에 액세스하려면 Endpoint URL을 입력하세요.
   c. Data Classification이 StorageGRID 의 버킷에 액세스할 수 있도록 액세스 키와 비밀 키를 입력하세요.



결과

StorageGRID 시스템 목록에 추가되었습니다.

StorageGRID 버킷에서 스캔 활성화 및 비활성화

StorageGRID 에서 데이터 분류를 활성화한 후 다음 단계는 스캔할 버킷을 구성하는 것입니다. 데이터 분류는 해당 버킷을 검색하여 사용자가 만든 시스템에 표시합니다.

단계

1. 구성 페이지에서 StorageGRID 시스템을 찾으세요.
2. StorageGRID 시스템 타일에서 *구성*을 선택합니다.
3. 다음 단계 중 하나를 완료하여 스캐닝을 활성화하거나 비활성화하세요.
   - 버킷에서 매핑 전용 스캔을 활성화하려면 *맵*을 선택합니다.
   - 버킷에 대한 전체 검사를 활성화하려면 *매핑 및 분류*를 선택합니다.
   - 버킷에서 스캐닝을 비활성화하려면 *끄기*를 선택하세요.

결과

데이터 분류는 활성화된 버킷을 스캔하기 시작합니다. 구성 메뉴로 이동한 다음 시스템 구성을 선택하면 초기 검사의 진행 상황을 추적할 수 있습니다. 각 스캔의 진행 상황은 진행률 표시줄로 표시됩니다. 진행률 표시줄 위에 마우스를 올려 놓으면 볼륨의 전체 파일 대비 검사된 파일 수를 확인할 수 있습니다. 오류가 있는 경우 오류는 상태 열에 표시되고 오류를 수정하는 데 필요한 작업도 함께 표시됩니다.

Active Directory를 NetApp Data Classification 와 통합하세요

NetApp Data Classification 와 글로벌 Active Directory를 통합하면 데이터 분류 보고서에서 파일 소유자와 파일에 액세스할 수 있는 사용자 및 그룹에 대한 결과를 더욱 향상시킬 수 있습니다.

특정 데이터 소스(아래 나열됨)를 설정하는 경우 데이터 분류가 CIFS 볼륨을 스캔할 수 있도록 Active Directory 자격

증명을 입력해야 합니다. 이 통합은 데이터 분류에 해당 데이터 소스에 있는 데이터의 파일 소유자 및 권한 세부 정보를 제공합니다. 해당 데이터 소스에 대해 입력한 Active Directory는 여기에 입력한 글로벌 Active Directory 자격 증명과 다를 수 있습니다. 데이터 분류는 모든 통합 Active Directory에서 사용자 및 권한 세부 정보를 찾습니다.

이 통합은 데이터 분류의 다음 위치에 추가 정보를 제공합니다.

- "파일 소유자"를 사용할 수 있습니다"필터" 조사 창에서 파일 메타데이터에서 결과를 확인하세요. SID(보안 식별자)를 포함하는 파일 소유자 대신 실제 사용자 이름이 채워집니다.

  파일 소유자에 대한 자세한 정보(계정 이름, 이메일 주소, SAM 계정 이름)를 보거나 해당 사용자가 소유한 항목을 볼 수도 있습니다.

- 당신은 볼 수 있습니다"전체 파일 권한" "모든 권한 보기" 버튼을 클릭하면 각 파일과 디렉토리에 대한 권한이 표시됩니다.
- 에서"거버넌스 대시보드" , 공개 권한 패널에는 데이터에 대한 더 자세한 정보가 표시됩니다.

로컬 사용자 SID와 알 수 없는 도메인의 SID는 실제 사용자 이름으로 변환되지 않습니다.

지원되는 데이터 소스

데이터 분류와 Active Directory를 통합하면 다음 데이터 소스에서 데이터를 식별할 수 있습니다.

- 온프레미스 ONTAP 시스템
- Cloud Volumes ONTAP
- Azure NetApp Files
- ONTAP 용 FSx

Active Directory 서버에 연결

데이터 분류를 배포하고 데이터 소스에 대한 스캐닝을 활성화한 후에는 데이터 분류를 Active Directory와 통합할 수 있습니다. Active Directory는 DNS 서버 IP 주소나 LDAP 서버 IP 주소를 사용하여 액세스할 수 있습니다.

Active Directory 자격 증명은 읽기 전용일 수 있지만, 관리자 자격 증명을 제공하면 데이터 분류에서 높은 권한이 필요한 모든 데이터를 읽을 수 있습니다. 자격 증명은 데이터 분류 인스턴스에 저장됩니다.

CIFS 볼륨/파일 공유의 경우 데이터 분류 검사에서 파일의 "마지막 액세스 시간"이 변경되지 않았는지 확인하려면 사용자에게 쓰기 속성 권한이 있어야 합니다. 가능하다면 Active Directory로 구성된 사용자를 모든 파일에 대한 권한이 있는 조직 내 상위 그룹에 포함하는 것이 좋습니다.

요구 사항

- 회사 사용자를 위해 Active Directory가 이미 설정되어 있어야 합니다.
- Active Directory에 대한 정보가 있어야 합니다.
  - DNS 서버 IP 주소 또는 여러 IP 주소

    또는

  LDAP 서버 IP 주소 또는 여러 IP 주소

- 서버에 접속하기 위한 사용자 이름과 비밀번호
- 도메인 이름(Active Directory 이름)
- 보안 LDAP(LDAPS)를 사용하든 사용하지 않든
- LDAP 서버 포트(일반적으로 LDAP의 경우 389, 보안 LDAP의 경우 636)

- 다음 포트는 데이터 분류 인스턴스의 아웃바운드 통신을 위해 열려 있어야 합니다.

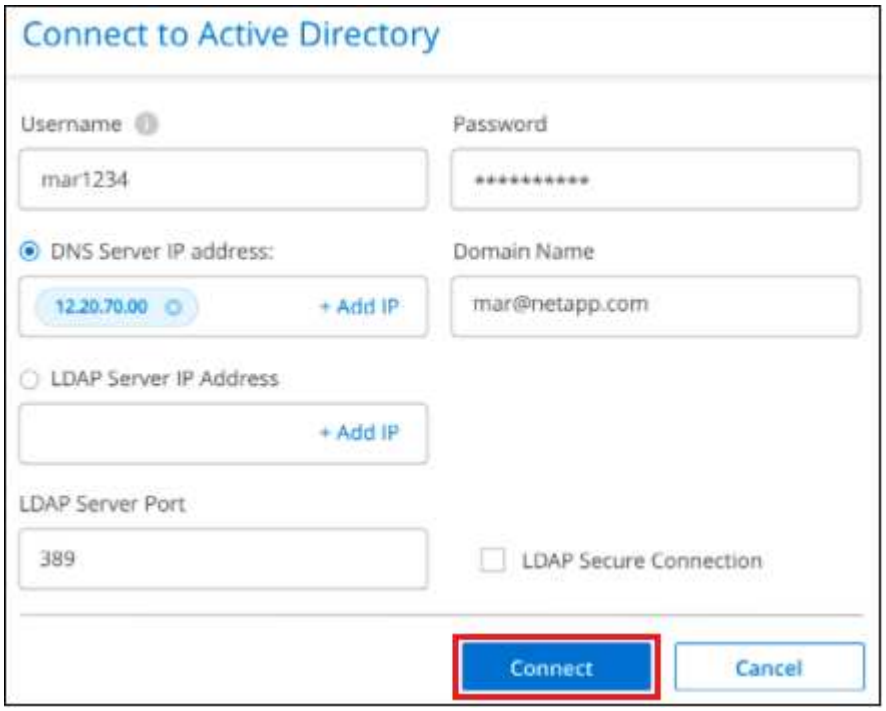
규약	포트	목적지	목적
TCP 및 UDP	389	액티브 디렉토리	LDAP
TCP	636	액티브 디렉토리	SSL을 통한 LDAP
TCP	3268	액티브 디렉토리	글로벌 카탈로그
TCP	3269	액티브 디렉토리	SSL을 통한 글로벌 카탈로그

단계

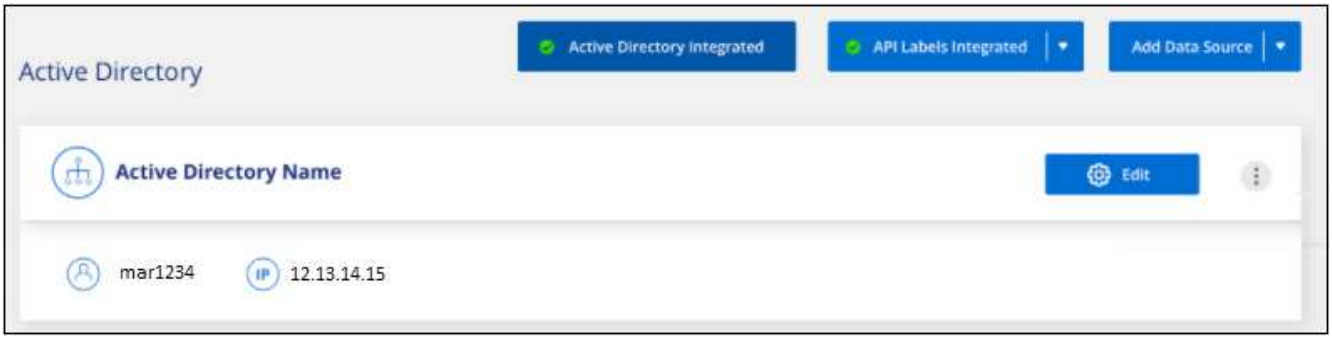
1. 데이터 분류 구성 페이지에서 *Active Directory 추가*를 클릭합니다.

2. Active Directory에 연결 대화 상자에서 Active Directory 세부 정보를 입력하고 *연결*을 클릭합니다.

   필요한 경우 *IP 추가*를 선택하여 여러 개의 IP 주소를 추가할 수 있습니다.



   데이터 분류가 Active Directory에 통합되었으며, 구성 페이지에 새로운 섹션이 추가되었습니다.



Active Directory 통합 관리

Active Directory 통합에서 값을 수정해야 하는 경우 편집 버튼을 클릭하고 변경합니다.

통합을 선택하여 삭제할 수도 있습니다. 버튼을 클릭한 다음 *Active Directory 제거*를 클릭합니다.

저작권 정보

상표 정보