始めましょう
NetApp Data Classification

NetApp
February 11, 2026

目次

始めましょう

NetApp Data Classificationについて学ぶ

NetApp Data Classification は、 NetApp Consoleのデータ ガバナンス サービスであり、企業のオンプレミスおよびクラウド データ ソースをスキャンしてデータをマッピングおよび分類し、個人情報を識別します。これにより、セキュリティとコンプライアンスのリスクが軽減され、ストレージ コストが削減され、データ移行プロジェクトが支援されます。

バージョン 1.31 以降、データ分類はNetApp Console内のコア機能として利用できます。追加料金はかかりません。分類ライセンスやサブスクリプションは必要ありません。 + 旧バージョン 1.30 以前を使用している場合は、サブスクリプションの有効期限が切れるまでそのバージョンを利用できます。

NetApp Console

データ分類には、NetApp Consoleからアクセスできます。

NetApp Consoleは、オンプレミスとクラウド環境全体にわたるエンタープライズ グレードのNetAppストレージとデータ サービスの集中管理を提供します。NetAppデータ サービスにアクセスして使用するには、コンソールが必要です。管理インターフェースとして、1 つのインターフェースから多数のストレージ リソースを管理できます。コンソール管理者は、企業内のすべてのシステムのストレージとサービスへのアクセスを制御できます。

NetApp Consoleの使用を開始するためにライセンスやサブスクリプションは必要ありません。ストレージ システムまたはNetAppデータ サービスへの接続を確保するためにクラウドにコンソール エージェントを展開する必要がある場合にのみ料金が発生します。ただし、コンソールからアクセスできる一部のNetAppデータ サービスは、ライセンスまたはサブスクリプションベースです。

詳細はこちら"NetApp Console"。

機能

データ分類では、人工知能 (AI)、自然言語処理 (NLP)、機械学習 (ML) を使用してスキャンしたコンテンツを理解し、エンティティを抽出してそれに応じてコンテンツを分類します。これにより、データ分類では次の機能領域が提供できるようになります。

"データ分類のユースケースについて学ぶ" 。

コンプライアンスを維持する

データ分類では、コンプライアンスの取り組みに役立ついくつかのツールが提供されます。データ分類を使用すると次のことができます。

- 個人を特定できる情報 (PII) を特定します。
- GDPR、CCPA、PCI、HIPAA のプライバシー規制で要求される、広範囲にわたる機密個人情報を特定します。
- 名前または電子メール アドレスに基づいて、データ主体アクセス要求 (DSAR) に応答します。

セキュリティを強化する

データ分類により、犯罪目的でアクセスされる危険性があるデータを識別できます。データ分類を使用すると次のことができます。

- 組織全体または一般に公開されている、オープン権限を持つすべてのファイルとディレクトリ (共有とフォルダー) を識別します。
- 最初の専用場所の外部に存在する機密データを識別します。
- データ保持ポリシーに準拠します。
- *Policies* を使用すると、新しいセキュリティ問題が自動的に検出され、セキュリティ スタッフがすぐに対処できるようになります。

ストレージ使用量を最適化する

データ分類は、ストレージの総所有コスト (TCO) の削減に役立つツールを提供します。データ分類を使用すると次のことができます。

- 重複データやビジネスに関係のないデータを識別して、ストレージ効率を向上させます。
- 非アクティブなデータを特定し、より安価なオブジェクト ストレージに階層化することで、ストレージコストを節約します。 "Cloud Volumes ONTAPシステムの階層化について詳しくは" 。 "オンプレミスのONTAPシステムからの階層化の詳細" 。

サポートされているシステムとデータソース

データ分類では、次の種類のシステムおよびデータ ソースからの構造化データと非構造化データをスキャンして分析できます。

システム

- Amazon FSx for NetApp ONTAP管理
- Azure NetApp Files
- Cloud Volumes ONTAP （AWS、Azure、または GCP にデプロイ）
- オンプレミスのONTAPクラスタ
- StorageGRID
- Google Cloud NetApp Volumes

データソース

- NetAppファイル共有
- データベース:
  - Amazon リレーショナルデータベースサービス (Amazon RDS)
  - MongoDB
  - MySQL
  - Oracle
  - PostgreSQL
  - SAP HANA

- SQL サーバー (MSSQL)

データ分類では、NFS バージョン 3.x、4.0、4.1、および CIFS バージョン 1.x、2.0、2.1、3.0 がサポートされています。

料金

データ分類は無料でご利用いただけます。分類ライセンスや有料サブスクリプションは必要ありません。

インフラコスト

- クラウドにデータ分類をインストールするには、クラウド インスタンスを展開する必要があり、展開先のクラウド プロバイダーから料金が発生します。見る各クラウドプロバイダーに展開されるインスタンスの種類 。オンプレミス システムに Data Classification をインストールする場合、料金はかかりません。
- データ分類では、コンソール エージェントを展開する必要があります。多くの場合、コンソールで使用している他のストレージやサービスがあるため、既にコンソール エージェントが存在します。コンソール エージェント インスタンスには、デプロイされているクラウド プロバイダーからの料金が発生します。参照 "各クラウドプロバイダーに展開されるインスタンスの種類"。オンプレミス システムにコンソール エージェントをインストールする場合、料金はかかりません。

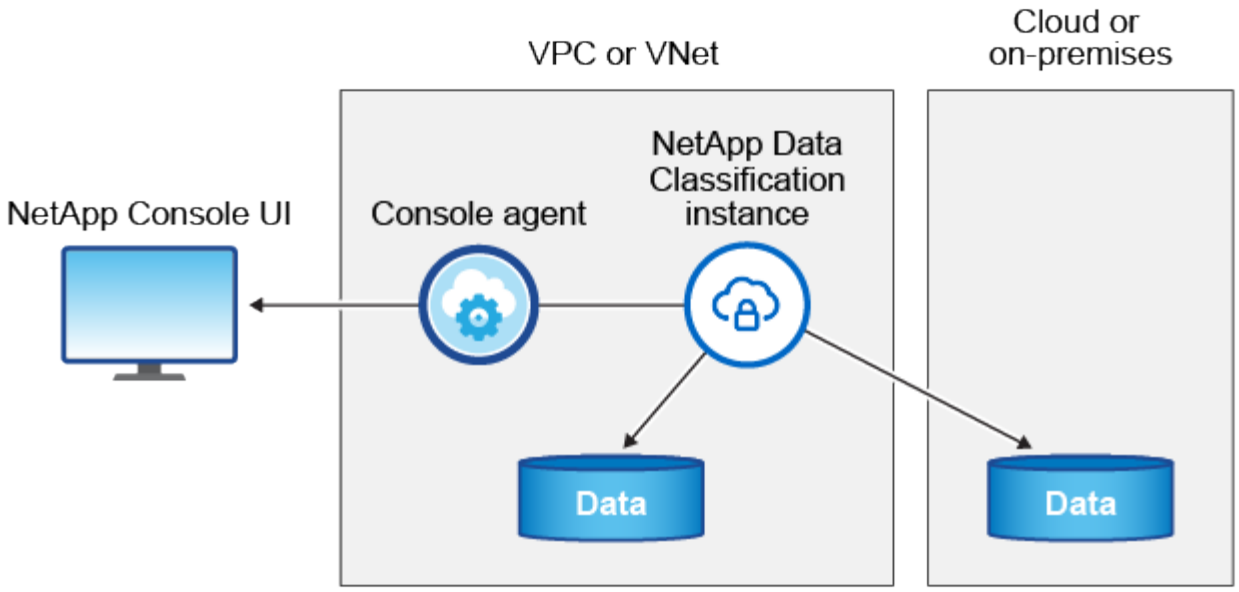
データ転送コスト

データ転送コストは設定によって異なります。データ分類インスタンスとデータ ソースが同じアベイラビリティー ゾーンとリージョンにある場合、データ転送コストは発生しません。ただし、 Cloud Volumes ONTAP システムなどのデータ ソースが別のアベイラビリティ ゾーンまたはリージョンにある場合は、クラウド プロバイダーからデータ転送コストが請求されます。詳細については、次のリンクを参照してください。

- "AWS: Amazon Elastic Compute Cloud (Amazon EC2) の料金"
- "Microsoft Azure: 帯域幅の料金詳細"
- "Google Cloud: ストレージ転送サービスの料金"

データ分類インスタンス

クラウドにデータ分類をデプロイすると、コンソールはコンソール エージェントと同じサブネットにインスタンスをデプロイします。 "コンソール エージェントの詳細について説明します。"



デフォルトインスタンスについては次の点に注意してください。

- AWSでは、データ分類は "m6i.4xlargeインスタンス"500 GiB GP2 ディスク付き。オペレーティング システム イメージは Amazon Linux 2 です。 AWS にデプロイする場合、少量のデータをスキャンする場合は、より小さいインスタンス サイズを選択できます。
- Azureでは、データ分類は"Standard_D16s_v3 VM"500 GiB のディスクを搭載。オペレーティング システム イメージは Ubuntu 22.04 です。
- GCPでは、データ分類は"n2-標準-16 VM"500 GiB の標準永続ディスクを備えています。オペレーティング システム イメージは Ubuntu 22.04 です。
- デフォルトのインスタンスが利用できないリージョンでは、データ分類は代替インスタンスで実行されます。"代替インスタンスタイプを参照" 。
- インスタンスの名前は *CloudCompliance* となり、生成されたハッシュ (UUID) が連結されます。例: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- コンソール エージェントごとに 1 つのデータ分類インスタンスのみが展開されます。

また、オンプレミスの Linux ホストまたは優先クラウド プロバイダーのホストにデータ分類を展開することもできます。どのインストール方法を選択しても、ソフトウェアはまったく同じように機能します。インスタンスがインターネットにアクセスできる限り、データ分類ソフトウェアのアップグレードは自動化されます。

データ分類は継続的にデータをスキャンするため、インスタンスは常に実行されたままにしておく必要があります。

異なるインスタンスタイプにデプロイ

インスタンス タイプの次の仕様を確認してください。

システムサイズ	仕様	制限事項
特大	32 個の CPU、128 GB の RAM、1 TiB の SSD	最大5億個のファイルをスキャンできます。

システムサイズ	仕様	制限事項
大（デフォルト）	16 CPU、64 GB RAM、500 GiB SSD	最大2億5000万個のファイルをスキャンできます。

Azure または GCP でデータ分類をデプロイするときに、より小さいインスタンス タイプを使用したい場合は、ng-contact-data-sense@netapp.com に電子メールでお問い合わせください。

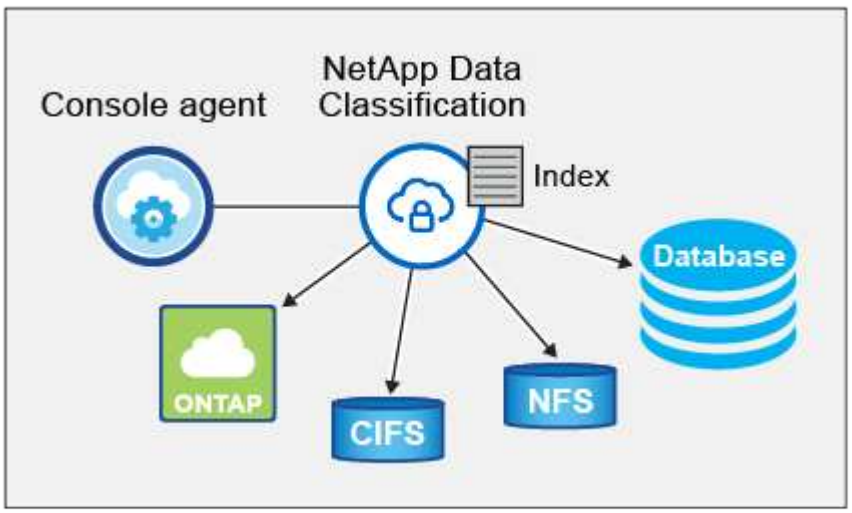
データ分類スキャンの仕組み

大まかに言えば、データ分類スキャンは次のように機能します。

1. コンソールでデータ分類のインスタンスをデプロイします。
2. 1 つ以上のデータ ソースに対して、高レベル マッピング (マッピングのみ スキャンと呼ばれる) または詳細レベル スキャン (マップと分類 スキャンと呼ばれる) を有効にします。
3. データ分類は、AI 学習プロセスを使用してデータをスキャンします。
4. 提供されているダッシュボードとレポート ツールを使用して、コンプライアンスとガバナンスの取り組みを支援します。

データ分類を有効にし、スキャンするリポジトリ (ボリューム、データベース スキーマ、またはその他のユーザー データ) を選択すると、すぐにデータのスキャンが開始され、個人データと機密データが識別されます。ほとんどの場合、バックアップ、ミラー、または DR サイトではなく、ライブの本番データのスキャンに重点を置く必要があります。次に、データ分類によって組織のデータがマッピングされ、各ファイルが分類され、データ内のエンティティと定義済みパターンが識別および抽出されます。スキャンの結果は、個人情報、機密個人情報、データ カテゴリ、およびファイル タイプのインデックスです。

Data Classification は、NFS および CIFS ボリュームをマウントすることで、他のクライアントと同様にデータに接続します。 NFS ボリュームは自動的に読み取り専用としてアクセスされますが、CIFS ボリュームをスキャンするには Active Directory の資格情報を提供する必要があります。



最初のスキャンの後、データ分類はラウンドロビン方式でデータを継続的にスキャンし、増分変更を検出します。そのため、インスタンスを実行し続けることが重要です。

ボリューム レベルまたはデータベース スキーマ レベルでスキャンを有効または無効にすることができます。

データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 + コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

マッピングスキャンと分類スキャンの違いは何ですか？

データ分類では、次の 2 種類のスキャンを実行できます。

- マッピングのみのスキャン は、データの概要のみを提供し、選択されたデータ ソースに対して実行されます。マッピングのみのスキャンでは、ファイルにアクセスして内部のデータを確認する必要がないため、マップおよび分類スキャンよりも時間がかかりません。最初にこれを実行して研究領域を特定し、次にそれらの領域に対してマップと分類のスキャンを実行することをお勧めします。
- マップと分類スキャン は、データの詳細なスキャンを提供します。

マッピングスキャンと分類スキャンの違いの詳細については、以下を参照してください。"マッピングスキャンと分類スキャンの違いは何ですか?" 。

データ分類が分類する情報

データ分類では、次のデータを収集し、インデックスを付け、カテゴリを割り当てます。

- ファイルに関する*標準メタデータ*: ファイルの種類、サイズ、作成日と変更日など。
- 個人データ: 電子メール アドレス、識別番号、クレジットカード番号などの個人を特定できる情報 (PII)。データ分類では、ファイル内の特定の単語、文字列、パターンを使用してこれを識別します。"個人データについて詳しくはこちら" 。
- 機密個人データ: 一般データ保護規則 (GDPR) やその他のプライバシー規制で定義されている、健康データ、民族的出身、政治的意見などの特別な種類の機密個人情報 (SPII)。"機密性の高い個人データについて詳しく見る" 。
- カテゴリ: データ分類は、スキャンしたデータを取得し、それをさまざまな種類のカテゴリに分割します。カテゴリは、各ファイルのコンテンツとメタデータの AI 分析に基づくトピックです。"カテゴリーについて詳しく見る"。
- 名前エンティティ認識: データ分類では、AI を使用して文書から人の自然な名前を抽出します。"データ主体のアクセス要求への対応について学ぶ" 。

ネットワークの概要

データ分類では、クラウドまたはオンプレミスなど、任意の場所に単一のサーバーまたはクラスターを展開します。サーバーは標準プロトコルを介してデータ ソースに接続し、同じサーバーにデプロイされている Elasticsearch クラスターで検出結果をインデックス化します。これにより、マルチクラウド、クロスクラウド、プライベートクラウド、オンプレミス環境のサポートが可能になります。

コンソールは、コンソール エージェントからの受信 HTTP 接続を有効にするセキュリティ グループを使用して、データ分類インスタンスをデプロイします。

コンソールを SaaS モードで使用する場合、コンソールへの接続は HTTPS 経由で提供され、ブラウザとデータ分類インスタンス間で送信されるプライベート データは TLS 1.2 を使用したエンドツーエンドの暗号化で

保護されるため、 NetAppやサードパーティが読み取ることはできません。

アウトバウンドルールは完全にオープンです。データ分類ソフトウェアをインストールおよびアップグレードし、使用状況メトリックを送信するには、インターネット アクセスが必要です。

厳しいネットワーク要件がある場合、"データ分類が接続するエンドポイントについて学習する" 。

NetApp Data Classificationにアクセス

NetApp Consoleを通じてNetApp Data Classification にアクセスできます。

コンソールにサインインするには、 NetAppサポート サイトの認証情報を使用するか、電子メールとパスワードを使用してNetApp Consoleログインにサインアップすることができます。"コンソールへのログインについて詳しくは" 。

特定のタスクには、特定のコンソール ユーザー ロールが必要です。"すべてのサービスのコンソールアクセスロールについて学習します" 。

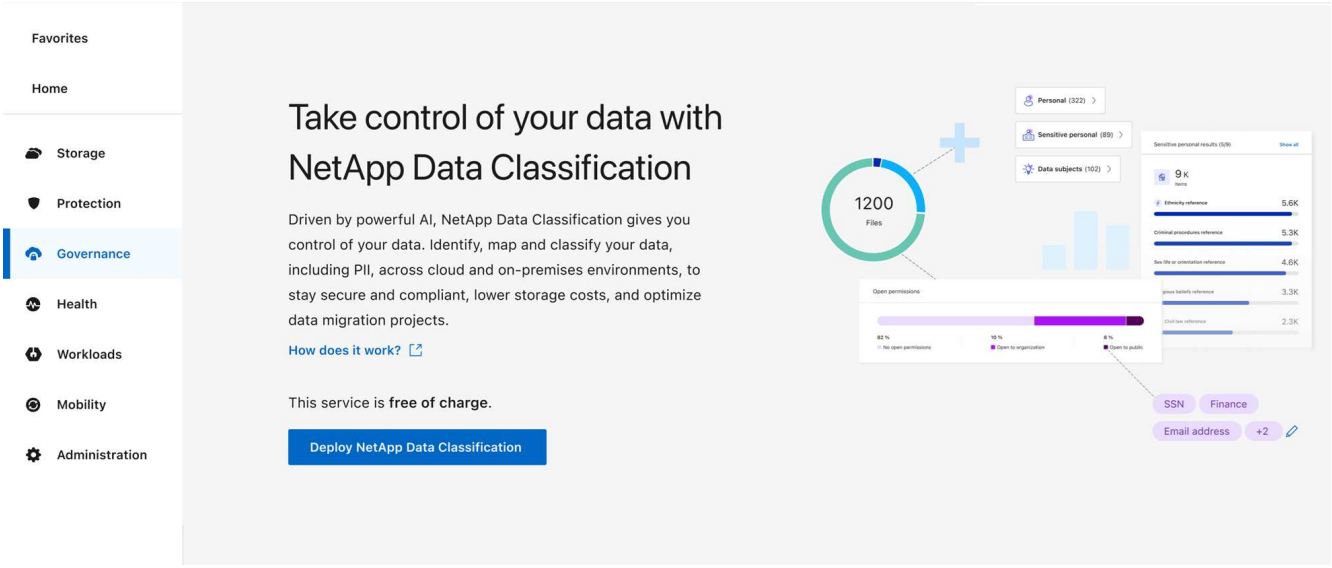
開始する前に

- "コンソール エージェントを追加する必要があります。"
- "どのデータ分類展開スタイルがワークロードに適しているかを理解します。"

手順

1. ウェブブラウザで、"コンソール" 。
2. コンソールにログインします。
3. NetApp Consoleのメイン ページから、ガバナンス > データ分類 を選択します。
4. データ分類に初めてアクセスする場合は、ランディング ページが表示されます。

   分類インスタンスの展開を開始するには、*オンプレミスまたはクラウドでの分類の展開*を選択します。詳細については、"どのデータ分類展開を使用する必要がありますか?"



   それ以外の場合は、データ分類ダッシュボードが表示されます。

データ分類を展開する

どの**NetApp Data Classification**展開を使用する必要がありますか**?**

NetApp Data Classification はさまざまな方法で導入できます。どの方法がニーズに合っているかを学びます。

データ分類は次の方法で展開できます。

- "コンソールを使用してクラウドにデプロイする" 。コンソールは、コンソール エージェントと同じクラウド プロバイダー ネットワークにデータ分類インスタンスを展開します。
- "インターネットにアクセスできる Linux ホストにインストールする" 。インターネットにアクセスできるネットワーク内の Linux ホスト、またはクラウド内の Linux ホストに Data Classification をインストールします。このタイプのインストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションですが、必須ではありません。
- "インターネットにアクセスできないオンプレミス サイトの Linux ホストにインストールする"プライベート モード とも呼ばれます。インストール スクリプトを使用するこのタイプのインストールでは、コンソール SaaS レイヤーへの接続はありません。

> BlueXPプライベート モード (レガシーBlueXPインターフェイス) は通常、インターネット接続がなく、AWS Secret Cloud、AWS Top Secret Cloud、Azure IL6 などの安全なクラウド領域があるオンプレミス環境で使用されます。NetApp は、従来のBlueXPインターフェースを使用してこれらの環境を引き続きサポートします。従来のBlueXPインターフェースのプライベートモードのドキュメントについては、"BlueXPプライベートモードの PDF ドキュメント" 。

インターネットにアクセスできる Linux ホストへのインストールと、インターネットにアクセスできない Linux ホストへのオンプレミス インストールの両方で、インストール スクリプトが使用されます。スクリプトは、システムと環境が前提条件を満たしているかどうかを確認することから始まります。前提条件が満たされている場合は、インストールが開始されます。データ分類のインストールを実行せずに前提条件を検証したい場合は、前提条件のみをテストする別のソフトウェア パッケージをダウンロードできます。

。 "Linuxホストがデータ分類をインストールする準備ができていることを確認します" 。

NetApp Consoleを使用してクラウドに**NetApp Data Classification**を導入する

NetApp Consoleを使用して、クラウドにNetApp Data Classification を導入できます。コンソールは、コンソール エージェントと同じクラウド プロバイダー ネットワークにデータ分類インスタンスを展開します。

また、"インターネットにアクセスできる Linux ホストにデータ分類をインストールする" 。このタイプのインストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションですが、必須ではありません。どのインストール方法を選択しても、ソフトウェアはまったく同じように機能します。

クイック スタート

以下の手順に従ってすぐに開始するか、残りのセクションまでスクロールして詳細を確認してください。

1 コンソールエージェントを作成する

コンソール エージェントがまだない場合は作成します。見る"AWSでコンソールエージェントを作成する"、"Azureでコンソールエージェントを作成する" 、 または"GCP でコンソール エージェントを作成する"。

また、"オンプレミスにコンソールエージェントをインストールする"ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。

2 前提条件

お使いの環境が前提条件を満たしていることを確認してください。これには、インスタンスの外向きインターネットアクセス、ConsoleエージェントとData Classification間のポート443での接続などが含まれます。完全なリストを見る.

3 データ分類を展開する

インストール ウィザードを起動して、データ分類インスタンスをクラウドにデプロイします。

コンソールエージェントを作成する

コンソール エージェントがまだない場合は、クラウド プロバイダーにコンソール エージェントを作成します。見る "AWSでコンソールエージェントを作成する"または "Azureでコンソールエージェントを作成する"、または "GCP でコンソール エージェントを作成する"。ほとんどの場合、データ分類を有効化する前にコンソールエージェントをセットアップしておく必要があります。 "コンソール機能にはコンソールエージェントが必要です"ただし、今すぐ設定する必要がある場合もあります。

特定のクラウド プロバイダーにデプロイされたコンソール エージェントを使用する必要があるシナリオがいくつかあります。

- AWS のCloud Volumes ONTAPまたはAmazon FSx for ONTAPバケット内のデータをスキャンする場合は、AWS のコンソールエージェントを使用します。
- Azure のCloud Volumes ONTAPまたはAzure NetApp Filesでデータをスキャンする場合は、Azure のコンソール エージェントを使用します。
  - Azure NetApp Filesの場合、スキャンするボリュームと同じリージョンにデプロイする必要があります。
- GCP のCloud Volumes ONTAPでデータをスキャンする場合は、GCP のコンソール エージェントを使用します。

これらのクラウド コンソール エージェントのいずれかを使用すると、オンプレミスのONTAPシステム、NetAppファイル共有、およびデータベースをスキャンできます。

また、 "オンプレミスにコンソールエージェントをインストールする"ネットワークまたはクラウド内の Linux ホスト上。オンプレミスでデータ分類をインストールする予定のユーザーの中には、オンプレミスでコンソール エージェントをインストールすることを選択する場合もあります。

使用する必要がある状況があるかもしれません "複数のコンソールエージェント"。

データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 + コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

政府地域支援

データ分類は、コンソール エージェントが政府リージョン (AWS GovCloud、Azure Gov、または Azure DoD) にデプロイされている場合にサポートされます。この方法で展開する場合、データ分類には次の制限があります。

"政府地域におけるコンソールエージェントの導入について学習します"。

前提条件

クラウドにデータ分類を展開する前に、次の前提条件を確認して、サポートされている構成があることを確認してください。クラウドにデータ分類を展開すると、コンソール エージェントと同じサブネットに配置されます。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。プロキシは非透過である必要があります。透過プロキシは現在サポートされていません。

データ分類を AWS、Azure、GCP のいずれに展開するかに応じて、以下の適切な表を確認してください。

AWSに必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供します。
https://kinesis.us-east-1.amazonaws.com	NetApp が監査レコードからデータをストリーミングできるようにします。
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	データ分類を有効にして、マニフェストとテンプレートにアクセスしてダウンロードし、ログとメトリックを送信できるようにします。

Azureに必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。

GCP に必要なエンドポイント

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。

エンドポイント	目的
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。

データ分類に必要な権限があることを確認する

データ分類に、リソースをデプロイし、データ分類インスタンスのセキュリティ グループを作成するための権限があることを確認します。

- "Google Cloud の権限"
- "AWS 権限"
- "Azure のアクセス許可"

コンソールエージェントがデータ分類にアクセスできることを確認する

コンソール エージェントとデータ分類インスタンス間の接続を確認します。コンソール エージェントのセキュリティ グループは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。この接続により、データ分類インスタンスのデプロイが可能になり、コンプライアンス タブとガバナンス タブで情報を表示できるようになります。データ分類は、AWS および Azure の政府リージョンでサポートされています。

AWS および AWS GovCloud のデプロイメントには、追加の受信および送信セキュリティ グループ ルールが必要です。見る "AWS のコンソールエージェントのルール"詳細については。

Azure および Azure Government の展開には、追加の受信および送信セキュリティ グループ ルールが必要です。見る "Azure のコンソール エージェントのルール"詳細については。

データ分類を継続して実行できるようにする

データを継続的にスキャンするには、データ分類インスタンスをオンのままにしておく必要があります。

データ分類への**Web**ブラウザ接続を確保する

データ分類を有効にした後、ユーザーがデータ分類インスタンスに接続しているホストからコンソール インターフェイスにアクセスすることを確認します。

データ分類インスタンスは、インデックス付けされたデータがインターネットからアクセスできないようにするためにプライベート IP アドレスを使用します。そのため、コンソールにアクセスするために使用する Web ブラウザは、そのプライベート IP アドレスに接続している必要があります。この接続は、クラウド プロバイダーへの直接接続 (VPN など) から行うことも、データ分類インスタンスと同じネットワーク内にあるホストから行うこともできます。

vCPUの制限を確認する

クラウド プロバイダーの vCPU 制限によって、必要な数のコアを持つインスタンスのデプロイが許可されていることを確認します。コンソールが実行されているリージョン内の関連するインスタンス ファミリの vCPU 制限を確認する必要があります。"必要なインスタンスタイプを確認する" 。

vCPU 制限の詳細については、次のリンクを参照してください。

- "AWS ドキュメント: Amazon EC2 サービスクォータ"
- "Azure ドキュメント: 仮想マシンの vCPU クォータ"
- "Google Cloud ドキュメント: リソース割り当て"

クラウドでデータ分類を展開

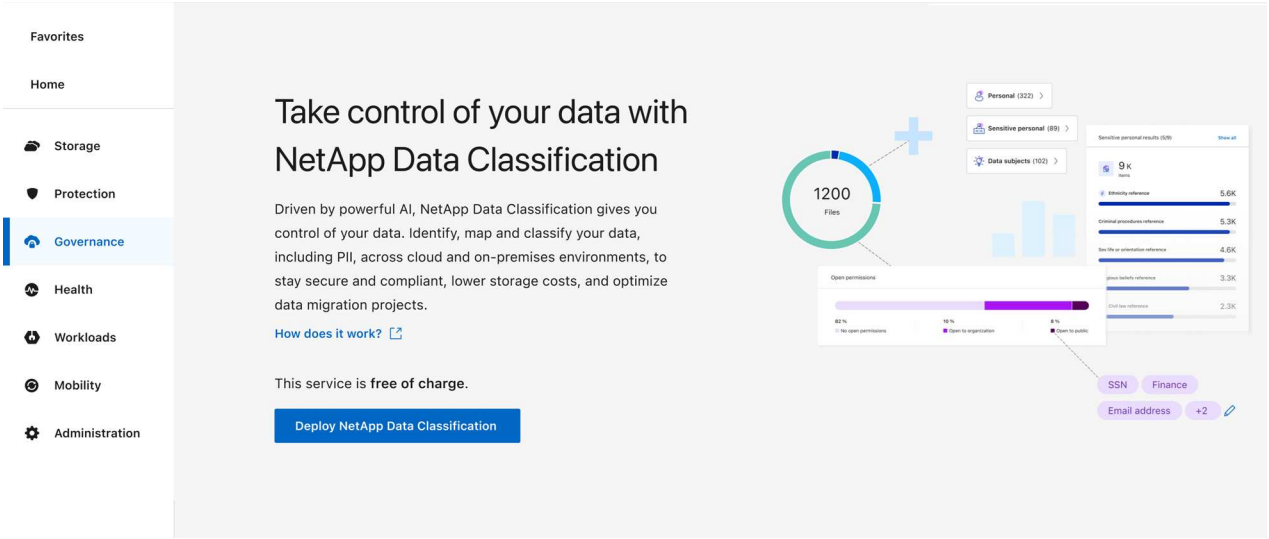
クラウドにデータ分類のインスタンスをデプロイするには、次の手順に従います。コンソール エージェントはクラウドにインスタンスを展開し、そのインスタンスにデータ分類ソフトウェアをインストールします。

デフォルトのインスタンスタイプが利用できない地域では、データ分類は"代替インスタンスタイプ"。

AWSにデプロイ

手順

1. データ分類のメイン ページから、オンプレミスまたはクラウドでの分類の展開 を選択します。

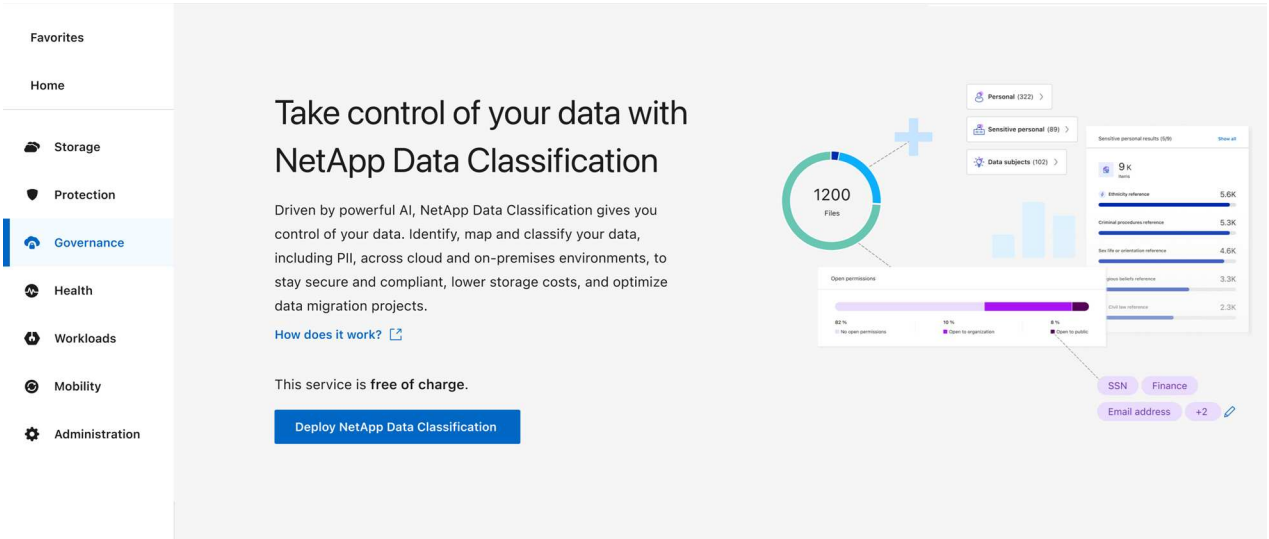


2. インストール ページで、デプロイ > デプロイ を選択して、「大」インスタンス サイズを使用し、クラウド デプロイ ウィザードを起動します。
3. ウィザードは、展開手順を実行する際の進行状況を表示します。入力が必要な場合、または問題が発生した場合には、プロンプトが表示されます。
4. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

Azureにデプロイする

手順

1. データ分類のメイン ページから、オンプレミスまたはクラウドでの分類の展開 を選択します。



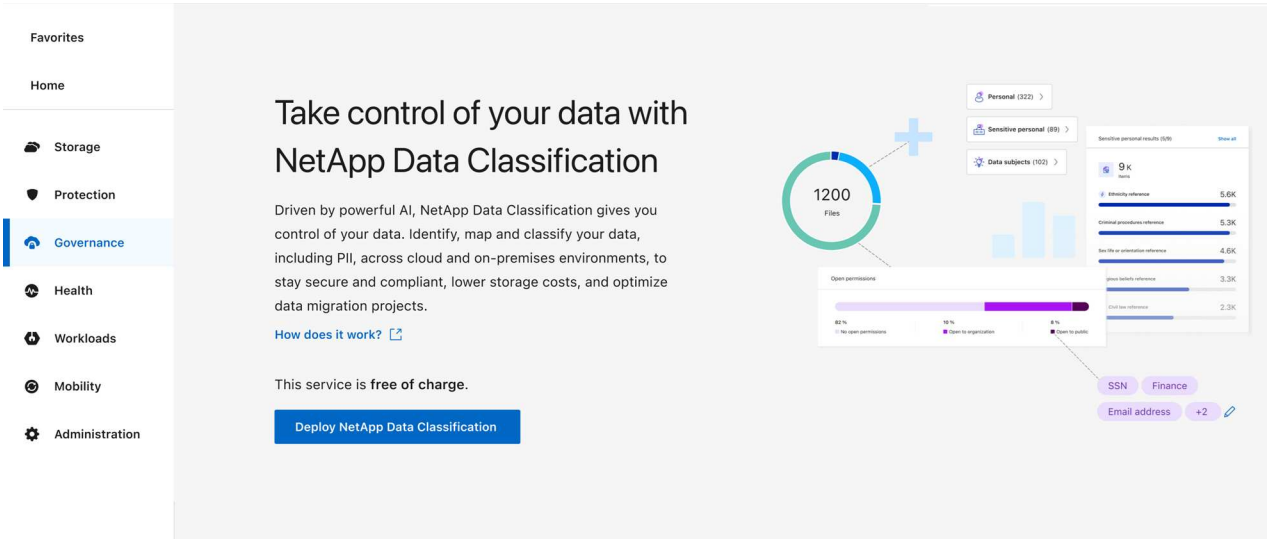
2. デプロイ を選択して、クラウド デプロイ ウィザードを起動します。

3. ウィザードは、展開手順を実行する際の進行状況を表示します。問題が発生すると停止し、入力を求められます。
4. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

Google Cloud にデプロイ

手順

1. データ分類のメイン ページから、ガバナンス > 分類 を選択します。
2. *オンプレミスまたはクラウドでの分類の展開*を選択します。



3. デプロイ を選択して、クラウド デプロイ ウィザードを起動します。
4. ウィザードは、展開手順を実行する際の進行状況を表示します。問題が発生すると停止し、入力を求められます。
5. インスタンスがデプロイされ、データ分類がインストールされたら、[構成に進む] を選択して [構成] ページに移動します。

結果

コンソールは、クラウド プロバイダーにデータ分類インスタンスをデプロイします。

インスタンスがインターネットに接続されている限り、コンソール エージェントとデータ分類ソフトウェアへのアップグレードは自動的に行われます。

次は何？

構成ページから、スキャンするデータ ソースを選択できます。

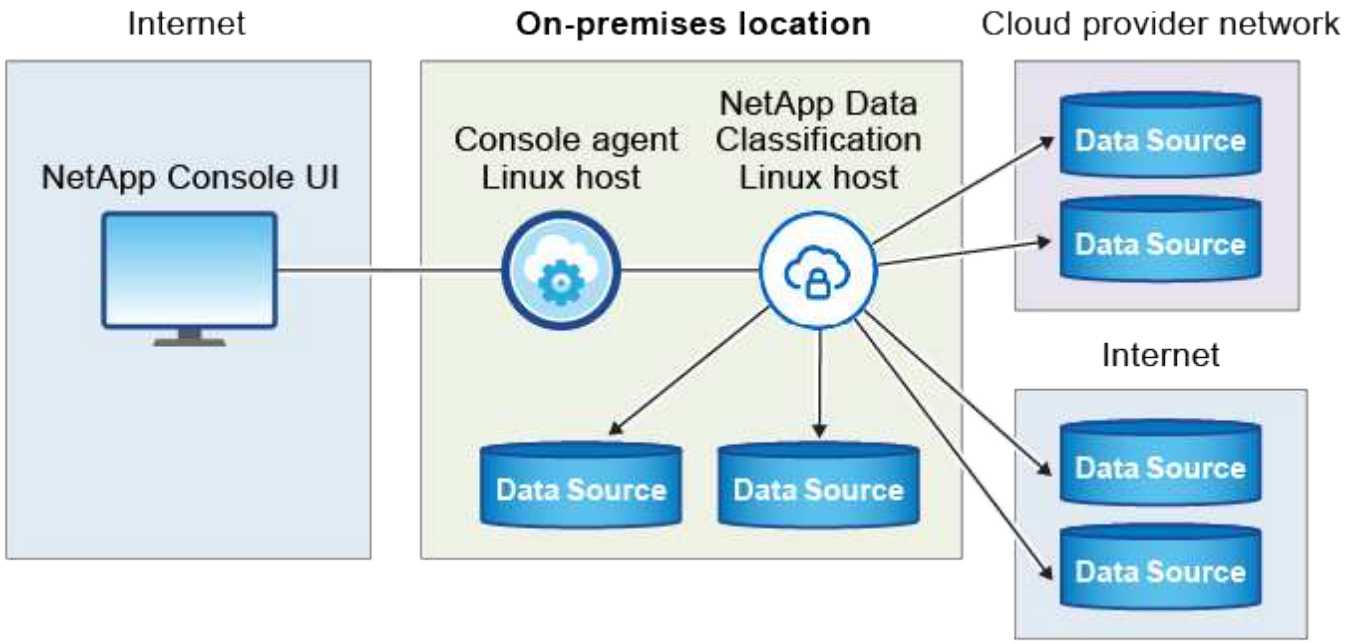
インターネットにアクセスできるホストに**NetApp Data Classification**をインストールする

ネットワーク内の Linux ホストまたはインターネットにアクセスできるクラウド内の Linux ホストにNetApp Data Classification を展開するには、ネットワークまたはクラウドに Linux ホストを手動で展開する必要があります。

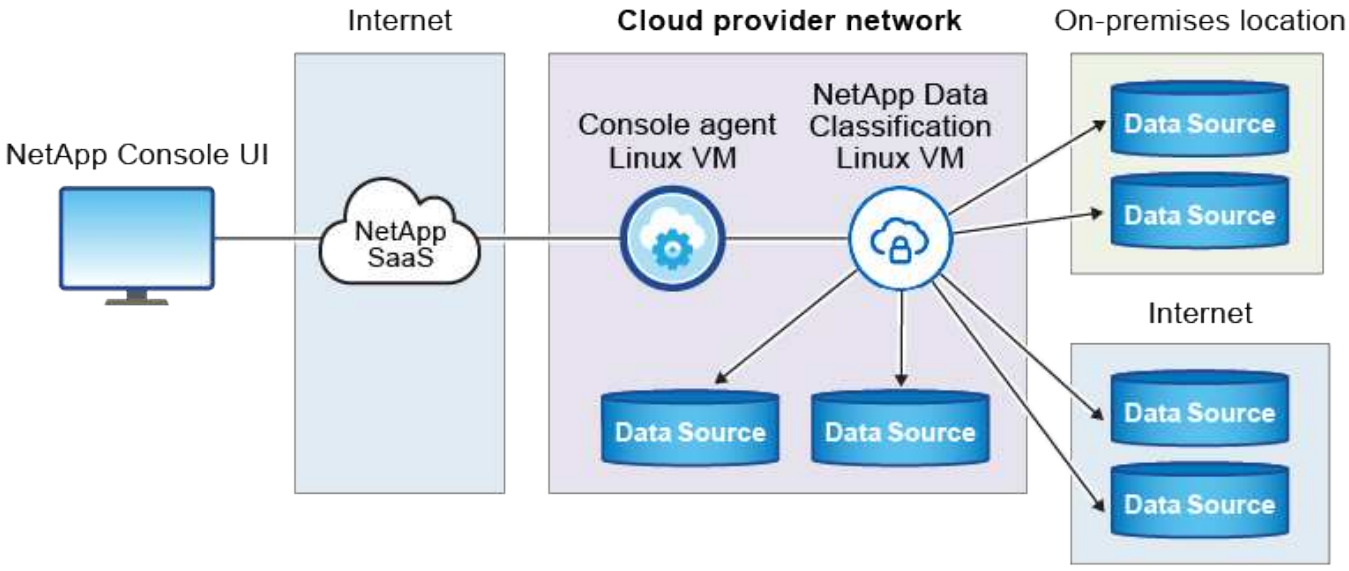
オンプレミス インストールは、オンプレミスにあるデータ分類インスタンスを使用してオンプレミスのONTAPシステムをスキャンする場合に適したオプションです。これは必須ではありません。どのインストール方法を選択しても、ソフトウェアは同じように機能します。

データ分類のインストール スクリプトは、システムと環境が必要な前提条件を満たしているかどうかを確認することから始まります。前提条件がすべて満たされている場合は、インストールが開始されます。データ分類のインストールの実行とは別に前提条件を検証したい場合は、前提条件のみをテストする別のソフトウェアパッケージをダウンロードできます。"Linuxホストがデータ分類をインストールする準備ができているかどうかを確認する方法をご覧ください" 。

社内の Linux ホストへの一般的なインストールには、次のコンポーネントと接続が含まれます。



クラウド内の Linux ホストへの一般的なインストールには、次のコンポーネントと接続が含まれます。



クイック スタート

以下の手順に従ってすぐに開始するか、残りのセクションまでスクロールして詳細を確認してください。

1 コンソールエージェントを作成する

コンソールエージェントがまだない場合は、 "コンソールエージェントをオンプレミスに展開する"ネットワーク内の Linux ホスト上、またはクラウド内の Linux ホスト上。

クラウド プロバイダーを使用してコンソール エージェントを作成することもできます。見る "AWSでコンソールエージェントを作成する"、 "Azureでコンソールエージェントを作成する" 、 または "GCP でコンソールエージェントを作成する"。

2 前提条件を確認する

ご使用の環境が前提条件を満たしていることを確認してください。これには、インスタンスのアウトバウンド インターネット アクセス、ポート 443 経由のコンソール エージェントとデータ分類間の接続などが含まれます。完全なリストを見る 。

また、以下の要件を満たすLinuxシステムも必要です。以下の要件 。

3 データ分類をダウンロードして展開する

NetAppサポート サイトから Cloud Data Classification ソフトウェアをダウンロードし、使用する予定の Linux ホストにインストーラ ファイルをコピーします。次に、インストール ウィザードを起動し、プロンプトに従ってデータ分類インスタンスをデプロイします。

コンソールエージェントを作成する

Data Classification をインストールして使用する前に、コンソール エージェントが必要です。ほとんどの場合、データ分類を有効化する前にコンソールエージェントをセットアップしておく必要があります。 "コンソール機能にはコンソールエージェントが必要です"ただし、今すぐ設定する必要がある場合もあります。

クラウドプロバイダー環境で作成するには、 "AWSでコンソールエージェントを作成する" 、 "Azureでコンソールエージェントを作成する" 、 または "GCP でコンソール エージェントを作成する"。

特定のクラウド プロバイダーにデプロイされたコンソール エージェントを使用する必要があるシナリオがいくつかあります。

- AWS のCloud Volumes ONTAPまたはAmazon FSx for ONTAPでデータをスキャンする場合は、AWS のコンソールエージェントを使用します。
- Azure のCloud Volumes ONTAPまたはAzure NetApp Filesでデータをスキャンする場合は、Azure のコンソール エージェントを使用します。

  Azure NetApp Filesの場合、スキャンするボリュームと同じリージョンにデプロイする必要があります。

- GCP のCloud Volumes ONTAPでデータをスキャンする場合は、GCP のコンソール エージェントを使用します。

オンプレミスのONTAPシステム、 NetAppファイル共有、データベース アカウントは、これらのクラウド コ

ンソール エージェントのいずれかを使用してスキャンできます。

また、 "コンソールエージェントをオンプレミスに展開する"ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。オンプレミスで Data Classification をインストールする予定のユーザーの中には、オンプレミスでコンソール エージェントをインストールすることを選択する場合もあります。

Data Classification をインストールするときは、コンソール エージェント システムの IP アドレスまたはホスト名が必要になります。オンプレミスでコンソール エージェントをインストールした場合は、この情報が得られます。コンソール エージェントがクラウドに展開されている場合は、コンソールからこの情報を確認できます。ヘルプ アイコンを選択し、次に サポート、コンソール エージェント を選択します。

Linuxホストシステムを準備する

データ分類ソフトウェアは、特定のオペレーティング システム要件、RAM 要件、ソフトウェア要件などを満たすホスト上で実行する必要があります。 Linux ホストはネットワーク内またはクラウド内に配置できます。

データ分類を実行し続けることができることを確認します。データ分類マシンは、データを継続的にスキャンするためにオンのままにしておく必要があります。

- データ分類は専用ホスト上で実行する必要があります。ホストは、他のアプリケーションやウイルス対策などのサードパーティ製ソフトウェアと共有することはできません。
- データ分類でスキャンする予定のデータ セットに合わせてサイズを選択します。

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
特大	CPU×32	128GBのRAM	• / に 1 TiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker で 895 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**
大きい	CPU×16	64GBのRAM	• / に 500 GiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker または Podman /var/lib/containers で 400 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

- データ分類インストール用にクラウドにコンピューティング インスタンスをデプロイする場合は、上記の「大規模」システム要件を満たすシステムを使用することをお勧めします。
  - **Amazon Elastic Compute Cloud (Amazon EC2)** インスタンスタイプ:「m6i.4xlarge」。"その他のAWSインスタンスタイプを見る" 。

- **Azure VM** サイズ:「Standard_D16s_v3」。"その他のAzureインスタンスタイプを見る" 。
- **GCP** マシンタイプ:「n2-standard-16」。"その他の GCP インスタンスタイプを見る" 。

- **UNIX** フォルダ権限: 次の最低限の UNIX 権限が必要です。

フォルダ	最小限の権限
/tmp	`rwxrwxrwt`
/opt	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/システム	`rwxr-xr-x`

- オペレーティング･システム：
  - 次のオペレーティング システムでは、Docker コンテナ エンジンを使用する必要があります。
    - Red Hat Enterprise Linux バージョン 7.8 および 7.9
    - Ubuntu 22.04 (データ分類バージョン 1.23 以上が必要)
    - Ubuntu 24.04 (データ分類バージョン 1.23 以上が必要)
  - 次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、データ分類バージョン 1.30 以上が必要です。
    - Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5、および 9.6。
  - ホスト システムで Advanced Vector Extensions (AVX2) を有効にする必要があります。
- **Red Hat** サブスクリプション管理: ホストは Red Hat サブスクリプション管理に登録されている必要があります。登録されていない場合、システムはリポジトリにアクセスできず、インストール中に必要なサードパーティ製ソフトウェアを更新できません。
- 追加ソフトウェア: Data Classification をインストールする前に、ホストに次のソフトウェアをインストールする必要があります。
  - 使用している OS に応じて、次のいずれかのコンテナ エンジンをインストールする必要があります。
    - Docker Engine バージョン 19.3.1 以上。 "インストール手順を見る" 。
    - Podman バージョン 4 以上。 Podmanをインストールするには、次のように入力します。(`sudo yum install podman netavark -y`）。
- Python バージョン 3.6 以上。 "インストール手順を見る" 。
  - **NTP** に関する考慮事項: NetApp、データ分類システムをネットワーク タイム プロトコル (NTP) サービスを使用するように構成することを推奨しています。データ分類システムとコンソール エージェント システムの間で時刻を同期する必要があります。
- **Firewalld**の考慮事項: 使用を計画している場合 `firewalld`、データ分類をインストールする前に有効にすることをお勧めします。設定するには次のコマンドを実行します `firewalld`データ分類と互換性があるように:

```
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
```

追加のデータ分類ホストをスキャナー ノードとして使用することを計画している場合は、この時点で次のルールをプライマリ システムに追加します。

```
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
```

有効化または更新するたびにDockerまたはPodmanを再起動する必要があることに注意してください。 `firewalld`設定。

データ分類ホスト システムの IP アドレスは、インストール後に変更することはできません。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソールとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.bluexp.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。
https://github.com/docker https://download.docker.com	docker インストールの前提条件パッケージを提供します。
http://packages.ubuntu.com/ http://archive.ubuntu.com	Ubuntu インストールの前提条件となるパッケージを提供します。

必要なポートがすべて有効になっていることを確認します

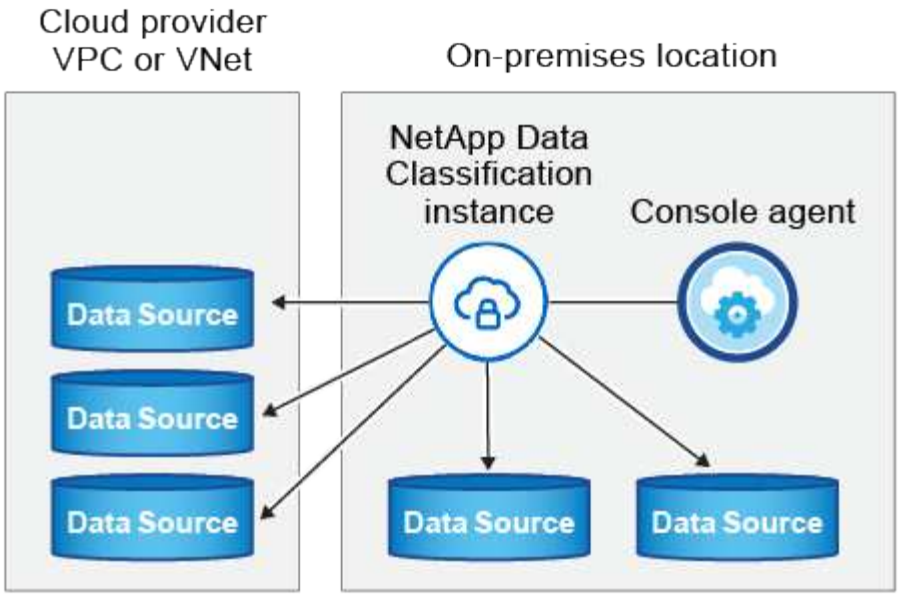
コンソール エージェント、データ分類、Active Directory、およびデータ ソース間の通信に必要なすべてのポートが開いていることを確認する必要があります。

接続タイプ	ポート	説明
コンソールエージェント <> データ分類	8080 (TCP)、443 (TCP)、および 80。9000	コンソール エージェントのファイアウォールまたはルーティング ルールでは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。コンソールでインストールの進行状況を確認できるように、ポート 8080 が開いていることを確認してください。 Linux ホストでファイアウォールが使用されている場合、Ubuntu サーバー内の内部プロセスにはポート 9000 が必要です。
コンソールエージェント <> ONTAPクラスタ (NAS)	443（TCP）	コンソールは、HTTPS を使用してONTAPクラスターを検出します。カスタム ファイアウォール ポリシーを使用する場合は、次の要件を満たす必要があります。 • コンソール エージェント ホストは、ポート 443 経由の送信 HTTPS アクセスを許可する必要があります。コンソール エージェントがクラウド内にある場合、すべての送信通信は事前定義されたファイアウォールまたはルーティング ルールによって許可されます。 • ONTAPクラスタは、ポート 443 経由の着信 HTTPS アクセスを許可する必要があります。デフォルトの「mgmt」ファイアウォール ポリシーでは、すべての IP アドレスからの受信 HTTPS アクセスが許可されます。このデフォルト ポリシーを変更した場合、または独自のファイアウォール ポリシーを作成した場合は、HTTPS プロトコルをそのポリシーに関連付け、コンソール エージェント ホストからのアクセスを有効にする必要があります。

接続タイプ	ポート	説明
データ分類 <> ONTAPクラスタ	• NFSの場合 - 111 (TCP\UDP) および 2049 (TCP\UDP) • CIFSの場合 - 139（TCP\UDP）および445（TCP\UDP）	データ分類には、各Cloud Volumes ONTAPサブネットまたはオンプレミスのONTAPシステムへのネットワーク接続が必要です。 Cloud Volumes ONTAPのファイアウォールまたはルーティング ルールは、データ分類インスタンスからの受信接続を許可する必要があります。 次のポートがデータ分類インスタンスに対して開いていることを確認します。 • NFSの場合 - 111と2049 • CIFSの場合 - 139および445 NFS ボリュームのエクスポート ポリシーでは、データ分類インスタンスからのアクセスを許可する必要があります。
データ分類 <> Active Directory	389 (TCP & UDP)、636 (TCP)、3268 (TCP)、および 3269 (TCP)	社内のユーザー用に Active Directory がすでに設定されている必要があります。さらに、データ分類では、CIFS ボリュームをスキャンするために Active Directory 資格情報が必要です。 Active Directory の情報が必要です: • DNSサーバーのIPアドレス、または複数のIPアドレス • サーバーのユーザー名とパスワード • ドメイン名（アクティブディレクトリ名） • セキュアLDAP（LDAPS）を使用しているかどうか • LDAP サーバー ポート (通常、LDAP の場合は 389、セキュア LDAP の場合は 636)

Linuxホストにデータ分類をインストールする

通常の構成では、ソフトウェアを単一のホスト システムにインストールします。ここでその手順をご覧ください 。



見るLinuxホストシステムの準備そして前提条件の確認データ分類を展開する前に、要件の完全なリストを確認してください。

インスタンスがインターネットに接続されている限り、データ分類ソフトウェアへのアップグレードは自動化されます。

> 現在、データ分類では、ソフトウェアがオンプレミスにインストールされている場合、S3 バケット、 Azure NetApp Files、または FSx for ONTAPをスキャンできません。このような場合には、クラウドに別のコンソールエージェントとデータ分類のインスタンスを展開し、 "コネクタ間の切り替え"さまざまなデータ ソース用。

一般的な構成の単一ホストインストール

単一のオンプレミス ホストにデータ分類ソフトウェアをインストールする場合は、要件を確認し、次の手順に従ってください。

"このビデオを見る"Data Classification のインストール方法を確認します。

Data Classification をインストールすると、すべてのインストール アクティビティがログに記録されることに注意してください。インストール中に問題が発生した場合は、インストール監査ログの内容を表示できます。それは、 `/opt/netapp/install_logs/` 。

開始する前に

- Linuxシステムがホスト要件。
- システムに 2 つの前提条件ソフトウェア パッケージ (Docker Engine または Podman、および Python 3) がインストールされていることを確認します。
- Linux システムでルート権限を持っていることを確認してください。
- インターネットへのアクセスにプロキシを使用している場合:
  - プロキシ サーバー情報 (IP アドレスまたはホスト名、接続ポート、接続スキーム: https または http、ユーザー名とパスワード) が必要になります。
  - プロキシが TLS インターセプションを実行している場合は、TLS CA 証明書が保存されている Data

Classification Linux システム上のパスを知っておく必要があります。

- プロキシは非透過である必要があります。データ分類は現在、透過プロキシをサポートしていません。
- ユーザーはローカル ユーザーである必要があります。ドメイン ユーザーはサポートされていません。

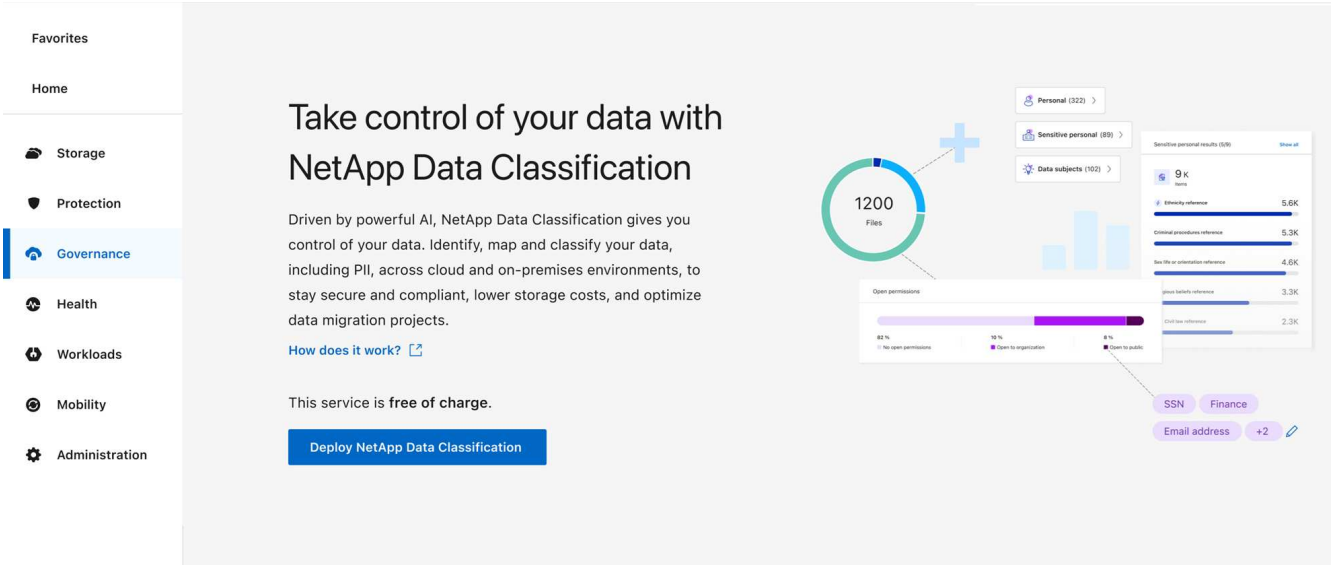
- オフライン環境が要件を満たしていることを確認する権限と接続性。

手順

1. データ分類ソフトウェアを以下からダウンロードしてください。 "NetAppサポート サイト" 。選択するファイルの名前は **DATASENSE-INSTALLER-<version>.tar.gz** です。
2. 使用する予定のLinuxホストにインストーラファイルをコピーします（ `scp`または他の方法）。
3. ホスト マシン上でインストーラ ファイルを解凍します。例:

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

4. コンソールで、*ガバナンス > 分類*を選択します。
5. *オンプレミスまたはクラウドでの分類の展開*を選択します。



6. クラウドで準備したインスタンスにデータ分類をインストールするか、オンプレミスで準備したインスタンスにデータ分類をインストールするかに応じて、適切な デプロイ オプションを選択してデータ分類のインストールを開始します。
7. オンプレミスでのデータ分類の展開 ダイアログが表示されます。提供されたコマンドをコピーします（例： `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`) を作成し、テキスト ファイルに貼り付けて、後で使用することもできます。次に、[閉じる] を選択してダイアログを閉じます。
8. ホスト マシンで、コピーしたコマンドを入力して一連のプロンプトに従うか、必要なすべてのパラメーターを含む完全なコマンドをコマンド ライン引数として指定することもできます。

   インストーラーは、インストールを正常に実行するためにシステムとネットワークの要件が満たされているかどうかを確認するための事前チェックを実行することに注意してください。 "このビデオを見る"事前チェックのメッセージとその意味を理解する。

プロンプトに従ってパラメータを入力します。	完全なコマンドを入力します。
a. 手順 7 からコピーしたコマンドを貼り付けます。 `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` クラウドインスタンス（オンプレミスではない）にインストールする場合は、 `--manual-cloud-install <cloud_provider>` 。 b. コンソール エージェント システムからアクセスできるように、データ分類ホスト マシンの IP アドレスまたはホスト名を入力します。 c. データ分類システムからアクセスできるように、コンソール エージェント ホスト マシンの IP アドレスまたはホスト名を入力します。 d. プロンプトに従ってプロキシの詳細を入力します。コンソール エージェントがすでにプロキシを使用している場合は、データ分類はコンソール エージェントが使用するプロキシを自動的に使用するため、ここでこの情報を再度入力する必要はありません。	あるいは、必要なホストとプロキシのパラメータを指定して、コマンド全体を事前に作成することもできます。 `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>`

変数値:

- *account_id* = NetAppアカウント ID
- *client_id* = コンソールエージェントのクライアントID（クライアントIDに「clients」というサフィックスがない場合は追加します）
- *user_token* = JWTユーザーアクセストークン
- *ds_host* = データ分類 Linux システムの IP アドレスまたはホスト名。
- *cm_host* = コンソール エージェント システムの IP アドレスまたはホスト名。
- *cloud_provider* = クラウドインスタンスにインストールする場合は、クラウドプロバイダーに応じて「AWS」、「Azure」、または「Gcp」を入力します。
- *proxy_host* = ホストがプロキシ サーバーの背後にある場合のプロキシ サーバーの IP またはホスト名。
- *proxy_port* = プロキシ サーバーに接続するためのポート (デフォルトは 80)。
- *proxy_scheme* = 接続スキーム: https または http (デフォルトは http)。
- *proxy_user* = 基本認証が必要な場合に、プロキシ サーバーに接続するための認証済みユーザー。ユーザーはローカル ユーザーである必要があります。ドメイン ユーザーはサポートされていません。
- *proxy_password* = 指定したユーザー名のパスワード。
- *ca_cert_dir* = 追加の TLS CA 証明書バンドルを含むデータ分類 Linux システム上のパス。プロキシが TLS インターセプションを実行している場合にのみ必要です。

結果

Data Classification インストーラーは、パッケージをインストールし、インストールを登録し、Data Classification をインストールします。インストールには10～20分かかります。

ホスト マシンとコンソール エージェント インスタンスの間にポート 8080 経由の接続がある場合は、コンソールの [データ分類] タブにインストールの進行状況が表示されます。

次は何？

構成ページから、スキャンするデータ ソースを選択できます。

インターネットにアクセスできない **Linux** ホストに**NetApp Data Classification**をインストールする

インターネットにアクセスできないオンプレミス サイトの Linux ホストにNetApp Data Classification をインストールすることを、プライベート モード と呼びます。インストール スクリプトを使用するこのタイプのインストールでは、 NetApp ConsoleSaaS レイヤーに接続できません。

> BlueXPプライベート モード (レガシーBlueXPインターフェイス) は通常、インターネット接続がなく、AWS Secret Cloud、AWS Top Secret Cloud、Azure IL6 などの安全なクラウド領域があるオンプレミス環境で使用されます。NetApp は、従来のBlueXPインターフェースを使用してこれらの環境を引き続きサポートします。従来のBlueXPインターフェースのプライベートモードのドキュメントについては、"BlueXPプライベートモードの PDF ドキュメント" 。

Linuxホストが**NetApp Data Classification**をインストールする準備ができていることを確認します

Linux ホストにNetApp Data Classification を手動でインストールする前に、オプションでホスト上でスクリプトを実行し、Data Classification をインストールするための前提条件がすべて満たされていることを確認します。このスクリプトは、ネットワーク内の Linux ホストまたはクラウド内の Linux ホストで実行できます。ホストはインターネットに接続することも、インターネットにアクセスできないサイト (ダーク サイト) に存在することもできます。

データ分類インストール スクリプトには、環境が要件を満たしていることを確認するためのテスト スクリプトが含まれています。インストール スクリプトを実行する前に、このスクリプトを個別に実行して、Linux ホストの準備状況を確認できます。

はじめに

以下のタスクを実行します。

- 必要に応じて、コンソール エージェントがまだインストールされていない場合はインストールします。コンソール エージェントがインストールされていなくてもテスト スクリプトを実行できますが、スクリプトはコンソール エージェントとデータ分類ホスト マシン間の接続をチェックするため、コンソール エージェントを使用することをお勧めします。
- ホスト マシンを準備し、すべての要件を満たしていることを確認します。
- データ分類ホスト マシンからのアウトバウンド インターネット アクセスを有効にします。

- すべてのシステムで必要なすべてのポートが有効になっていることを確認します。
- 前提条件テスト スクリプトをダウンロードして実行します。

コンソールエージェントを作成する

Data Classification をインストールして使用する前に、コンソール エージェントが必要です。ただし、コンソール エージェントなしで前提条件スクリプトを実行することは可能です。

あなたはできる "オンプレミスにコンソールエージェントをインストールする" ネットワーク内の Linux ホストまたはクラウド内の Linux ホスト上。コンソール エージェントがオンプレミスにインストールされている場合は、データ分類をオンプレミスにインストールすることもできます。

クラウド プロバイダー環境でコンソール エージェントを作成するには、以下を参照してください。

- "AWSでコンソールエージェントを作成する"
- "Azureでコンソールエージェントを作成する"
- "GCP でコンソール エージェントを作成する"

前提条件スクリプトを実行するときは、コンソール エージェント システムの IP アドレスまたはホスト名が必要です。オンプレミスでコンソール エージェントをインストールした場合は、この情報が得られます。コンソール エージェントがクラウドに展開されている場合は、コンソールからこの情報を確認できます。ヘルプ アイコンを選択してから [サポート] を選択し、[エージェントと監査] セクションで [エージェントに移動] を選択します。

ホストの要件を確認する

データ分類ソフトウェアは、特定のオペレーティング システム要件、RAM 要件、およびソフトウェア要件を満たすホスト上で実行する必要があります。

- データ分類は専用ホスト上で実行する必要があります。ホストは、他のアプリケーションやウイルス対策などのサードパーティ製ソフトウェアと共有することはできません。
- データ分類でスキャンする予定のデータ セットに合わせてサイズを選択します。

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
特大	CPU×32	128GBのRAM	• / に 1 TiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker で 895 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

システムサイズ	**CPU**	**RAM**（スワップメモリを無効にする必要があります）	ディスク
大きい	CPU×16	64GBのRAM	• / に 500 GiB SSD、または /opt に 100 GiB 利用可能 • /var/lib/docker または Podman /var/lib/containers で 400 GiB が利用可能 • /tmp に 5 GiB • **Podman** の場合、**/var/tmp** に **30 GB**

- データ分類インストール用にクラウドにコンピューティング インスタンスをデプロイする場合は、上記の「大規模」システム要件を満たすシステムを使用することをお勧めします。
  - **Amazon Elastic Compute Cloud (Amazon EC2)** インスタンスタイプ:「m6i.4xlarge」。"その他のAWSインスタンスタイプを見る" 。
  - **Azure VM** サイズ:「Standard_D16s_v3」。"その他のAzureインスタンスタイプを見る" 。
  - **GCP** マシンタイプ:「n2-standard-16」。"その他の GCP インスタンスタイプを見る" 。
- **UNIX** フォルダ権限: 次の最低限の UNIX 権限が必要です。

フォルダ	最小限の権限
/tmp	`rwxrwxrwt`
/opt	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/システム	`rwxr-xr-x`

- オペレーティング･システム：
  - 次のオペレーティング システムでは、Docker コンテナ エンジンを使用する必要があります。
    - Red Hat Enterprise Linux バージョン 7.8 および 7.9
    - Ubuntu 22.04 (データ分類バージョン 1.23 以上が必要)
    - Ubuntu 24.04 (データ分類バージョン 1.23 以上が必要)
  - 次のオペレーティング システムでは、Podman コンテナー エンジンを使用する必要があり、データ分類バージョン 1.30 以上が必要です。
    - Red Hat Enterprise Linux バージョン 8.8、8.10、9.0、9.1、9.2、9.3、9.4、9.5、および 9.6。
  - ホスト システムで Advanced Vector Extensions (AVX2) を有効にする必要があります。
- **Red Hat** サブスクリプション管理: ホストは Red Hat サブスクリプション管理に登録されている必要があります。登録されていない場合、システムはリポジトリにアクセスできず、インストール中に必要なサードパーティ製ソフトウェアを更新できません。
- 追加ソフトウェア: Data Classification をインストールする前に、ホストに次のソフトウェアをインストールする必要があります。

- 使用している OS に応じて、次のいずれかのコンテナ エンジンをインストールする必要があります。
  - Docker Engine バージョン 19.3.1 以上。 "インストール手順を見る" 。
  - Podman バージョン 4 以上。 Podmanをインストールするには、次のように入力します。(`sudo yum install podman netavark -y`）。

- Python バージョン 3.6 以上。 "インストール手順を見る" 。
  - **NTP** に関する考慮事項: NetApp、データ分類システムをネットワーク タイム プロトコル (NTP) サービスを使用するように構成することを推奨しています。データ分類システムとコンソール エージェント システムの間で時刻を同期する必要があります。

- **Firewalld**の考慮事項: 使用を計画している場合 `firewalld`、データ分類をインストールする前に有効にすることをお勧めします。設定するには次のコマンドを実行します `firewalld`データ分類と互換性があるように:

```
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
```

追加のデータ分類ホストをスキャナー ノードとして使用することを計画している場合 (分散モデル)、この時点で次のルールをプライマリ システムに追加します。

```
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
```

有効化または更新するたびにDockerまたはPodmanを再起動する必要があることに注意してください。 `firewalld`設定。

データ分類からのアウトバウンドインターネットアクセスを有効にする

データ分類には、アウトバウンドのインターネット アクセスが必要です。仮想ネットワークまたは物理ネットワークでインターネット アクセスにプロキシ サーバーを使用している場合は、データ分類インスタンスに次のエンドポイントに接続するための送信インターネット アクセスがあることを確認してください。

このセクションは、インターネットに接続できないサイトにインストールされたホスト システムでは必要ありません。

エンドポイント	目的
https://api.console.netapp.com	NetAppアカウントを含むコンソール サービスとの通信。
https://netapp-cloud-account.auth0.com https://auth0.com	集中ユーザー認証のためのコンソール Web サイトとの通信。

エンドポイント	目的
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/	ソフトウェア イメージ、マニフェスト、テンプレートへのアクセスを提供し、ログとメトリックを送信します。
https://support.compliance.api.console.netapp.com/	NetApp が監査レコードからデータをストリーミングできるようにします。
https://github.com/docker https://download.docker.com	docker インストールの前提条件パッケージを提供します。
http://packages.ubuntu.com/ http://archive.ubuntu.com	Ubuntu インストールの前提条件となるパッケージを提供します。

必要なポートがすべて有効になっていることを確認します

コンソール エージェント、データ分類、Active Directory、およびデータ ソース間の通信に必要なすべてのポートが開いていることを確認する必要があります。

接続タイプ	ポート	説明
コンソールエージェント <> データ分類	8080 (TCP)、443 (TCP)、および 80。9000	コンソール エージェントのファイアウォールまたはルーティング ルールでは、ポート 443 経由のデータ分類インスタンスとの間の受信トラフィックと送信トラフィックを許可する必要があります。コンソールでインストールの進行状況を確認できるように、ポート 8080 が開いていることを確認してください。 Linux ホストでファイアウォールが使用されている場合、Ubuntu サーバー内の内部プロセスにはポート 9000 が必要です。
コンソールエージェント <> ONTAPクラスタ (NAS)	443（TCP）	コンソールは、HTTPS を使用してONTAPクラスターを検出します。カスタム ファイアウォール ポリシーを使用する場合、コンソール エージェント ホストはポート 443 経由の送信 HTTPS アクセスを許可する必要があります。コンソール エージェントがクラウド内にある場合、すべての送信通信は事前定義されたファイアウォールまたはルーティング ルールによって許可されます。

データ分類の前提条件スクリプトを実行する

データ分類の前提条件スクリプトを実行するには、次の手順に従います。

"このビデオを見る"前提条件スクリプトを実行して結果を解釈する方法を確認します。

開始する前に

- Linuxシステムがホスト要件。
- システムに 2 つの前提条件ソフトウェア パッケージ (Docker Engine または Podman、および Python 3)

がインストールされていることを確認します。

- Linux システムでルート権限を持っていることを確認してください。

手順

1. データ分類の前提条件スクリプトを以下からダウンロードします。 "NetAppサポート サイト" 。選択するファイルの名前は **standalone-pre-requisite-tester-<version>** です。
2. 使用する予定のLinuxホストにファイルをコピーします（ `scp`または他の方法）。
3. スクリプトを実行するための権限を割り当てます。

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

4. 次のコマンドを使用してスクリプトを実行します。

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

インターネットにアクセスできないホストでスクリプトを実行する場合にのみ、オプション "--darksite" を追加します。ホストがインターネットに接続されていない場合、特定の前提条件テストはスキップされます。

5. スクリプトは、データ分類ホスト マシンの IP アドレスの入力を求めます。
   - IP アドレスまたはホスト名を入力します。
6. スクリプトは、コンソール エージェントがインストールされているかどうかを尋ねます。
   - コンソール エージェントがインストールされていない場合は、**N** と入力します。
   - コンソール エージェントがインストールされている場合は、「**Y**」と入力します。次に、テスト スクリプトがこの接続をテストできるように、コンソール エージェントの IP アドレスまたはホスト名を入力します。
7. スクリプトはシステム上でさまざまなテストを実行し、進行中に結果を表示します。終了すると、セッションのログを次のファイルに書き込みます。 `prerequisites-test-<timestamp>.log`ディレクトリ内 `/opt/netapp/install_logs`。

結果

すべての前提条件テストが正常に実行された場合、準備ができたらホストに Data Classification をインストールできます。

問題が発見された場合は、修正が「推奨」または「必須」として分類されます。推奨される問題は通常、データ分類のスキャンと分類のタスクの実行速度を低下させる項目です。これらの項目は修正する必要はありませんが、対処することをお勧めします。

「必須」の問題がある場合は、問題を修正して、前提条件テスト スクリプトを再度実行する必要があります。

データソースのスキャンを有効にする

NetApp Data Classificationでデータソースをスキャン

NetApp Data Classification は、選択したリポジトリ内のデータ (ボリューム、データベース スキーマ、またはその他のユーザー データ) をスキャンして、個人データや機密データを識別します。次に、データ分類により組織のデータがマッピングされ、各ファイルが分類され、データ内の定義済みパターンが識別されます。スキャンの結果は、個人情報、機密個人情報、データ カテゴリ、およびファイル タイプのインデックスです。

最初のスキャンの後、データ分類はラウンドロビン方式でデータを継続的にスキャンし、増分変更を検出します。そのため、インスタンスを実行し続けることが重要です。

ボリューム レベルまたはデータベース スキーマ レベルでスキャンを有効または無効にすることができます。

マッピングスキャンと分類スキャンの違いは何ですか？

データ分類では、次の 2 種類のスキャンを実行できます。

- マッピングのみのスキャン は、データの概要のみを提供し、選択されたデータ ソースに対して実行されます。マッピングのみのスキャンでは、ファイルにアクセスして内部のデータを確認する必要がないため、マップおよび分類スキャンよりも時間がかかりません。最初にこれを実行して研究領域を特定し、次にそれらの領域に対してマップと分類のスキャンを実行することをお勧めします。
- マップと分類スキャン は、データの詳細なスキャンを提供します。

以下の表にいくつかの違いを示します。

特徴	スキャンをマップして分類する	マッピングのみのスキャン
スキャン速度	遅い	速い
料金	空き	空き
容量	500 TiB に制限されます*	500 TiB に制限されます*
ファイルの種類と使用容量の一覧	はい	はい
ファイル数と使用容量	はい	はい
ファイルの古さとサイズ	はい	はい
実行する能力"データマッピングレポート"	はい	はい
ファイルの詳細を表示するためのデータ調査ページ	はい	いいえ
ファイル内の名前を検索する	はい	いいえ
作成する"保存されたクエリ"カスタム検索結果を提供する	はい	いいえ
他のレポートを実行する機能	はい	いいえ
ファイルのメタデータを表示する機能**	いいえ	はい

{アスタリスク} データ分類では、スキャンできるデータの量に制限はありません。各コンソール エージェントは、500 TiB のデータのスキャンと表示をサポートします。500TiB以上のデータをスキャンするには、"別のコンソールエージェントをインストールする"それから"別のデータ分類インスタンスをデプロイする"。 +

コンソール UI には、単一のコネクタからのデータが表示されます。複数のコンソールエージェントからデータを表示するヒントについては、"複数のコンソールエージェントを操作する" 。

{アスタリスク}{アスタリスク} マッピング スキャン中にファイルから次のメタデータが抽出されます。

- システム
- システムタイプ
- ストレージリポジトリ
- ファイル タイプ
- 使用済み容量
- ファイル数
- ファイル サイズ
- ファイル作成
- ファイルの最終アクセス
- ファイルの最終更新日時
- ファイル発見時刻
- 権限の抽出

ガバナンス ダッシュボードの違い:

特徴	マップと分類	マップ
古いデータ	はい	はい
非ビジネスデータ	はい	はい
重複したファイル	はい	はい
定義済みの保存済みクエリ	はい	いいえ
デフォルトの保存クエリ	はい	はい
DDAレポート	はい	はい
マッピングレポート	はい	はい
感度レベル検出	はい	いいえ
幅広い権限を持つ機密データ	はい	いいえ
オープン権限	はい	はい
データの時代	はい	はい
データのサイズ	はい	はい
カテゴリ	はい	いいえ
ファイルの種類	はい	はい

コンプライアンス ダッシュボードの違い:

特徴	マップと分類	マップ
個人情報	はい	いいえ
機密個人情報	はい	いいえ
プライバシーリスク評価レポート	はい	いいえ
HIPAAレポート	はい	いいえ
PCI DSSレポート	はい	いいえ

調査フィルターの違い:

特徴	マップと分類	マップ
保存されたクエリ	はい	はい
システムタイプ	はい	はい
システム	はい	はい
ストレージリポジトリ	はい	はい
ファイル タイプ	はい	はい
ファイル サイズ	はい	はい
作成時間	はい	はい
発見された時間	はい	はい
最終更新日	はい	はい
最終アクセス	はい	はい
オープン権限	はい	はい
ファイルディレクトリパス	はい	はい
カテゴリ	はい	いいえ
感度レベル	はい	いいえ
識別子の数	はい	いいえ
個人データ	はい	いいえ
機密性の高い個人データ	はい	いいえ
データ主体	はい	いいえ
重複	はい	はい
分類ステータス	はい	ステータスは常に「限られた洞察」です
スキャン分析イベント	はい	はい
ファイルハッシュ	はい	はい
アクセス権を持つユーザーの数	はい	はい
ユーザー/グループの権限	はい	はい
ファイルの所有者	はい	はい
ディレクトリタイプ	はい	はい

NetApp Data Classificationを使用して**Amazon FSx for ONTAP**ボリュームをスキャンする

NetApp Data Classificationを使用してAmazon FSx for ONTAPボリュームをスキャンするには、いくつかの手順を実行します。

開始する前に

- データ分類を展開および管理するには、AWS にアクティブなコンソールエージェントが必要です。
- システムの作成時に選択したセキュリティ グループは、データ分類インスタンスからのトラフィックを許可する必要があります。 FSx for ONTAPファイルシステムに接続された ENI を使用して関連付けられたセキュリティグループを見つけ、AWS マネジメントコンソールを使用して編集できます。

  "Linuxインスタンス用のAWSセキュリティグループ"

  "Windowsインスタンス用のAWSセキュリティグループ"

  "AWS エラスティックネットワークインターフェース (ENI)"

- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Data Classification は、AWS のコンソールエージェントおよびスキャンする FSx ボリュームと同じ AWS ネットワークにデプロイする必要があります。

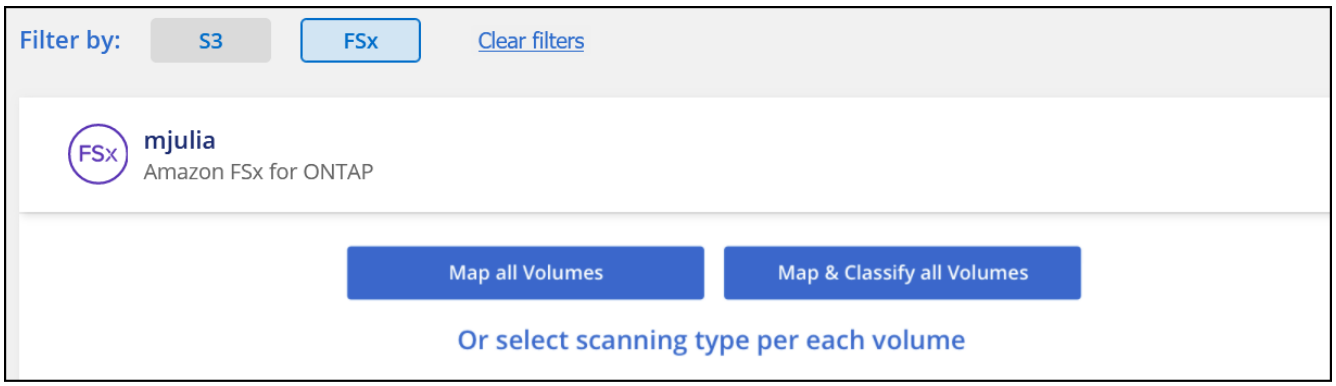
注意: FSx ボリュームをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

インスタンスがインターネットに接続されている限り、データ分類ソフトウェアへのアップグレードは自動化されます。

システムでデータ分類を有効にする

FSx for ONTAPボリュームのデータ分類を有効にすることができます。

1. NetApp Consoleから、*ガバナンス > 分類*を選択します。
2. データ分類メニューから、*構成*を選択します。



タブのスクリーンショット。"]

3. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ"

:

- すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
- すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
- 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップおよび/または分類するボリュームを選択します。

4. 確認ダイアログボックスで「承認」を選択すると、データ分類によってボリュームのスキャンが開始されます。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できるようになります。かかる時間はデータの量によって異なり、数分または数時間かかる場合があります。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。進行状況バーで各スキャンの進行状況を追跡します。進行状況バーの上にマウスを置くと、ボリューム内のファイルの合計数に対するスキャンされたファイルの数が表示されます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"このデータ分類の制限に関する詳細を見る" 。

データ分類がボリュームにアクセスできることを確認する

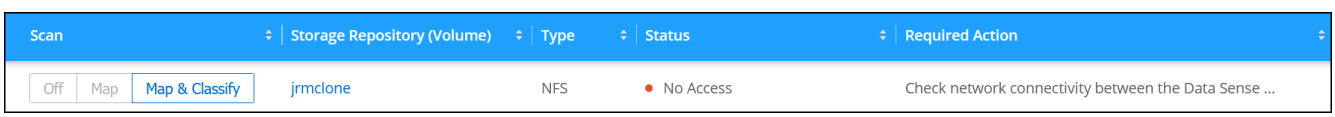
ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。

CIFS ボリュームにアクセスできるようにするには、データ分類に CIFS 資格情報を提供する必要があります。

手順

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで [詳細の表示] を選択してステータスを確認し、エラーを修正します。

   たとえば、次の画像は、データ分類インスタンスとボリューム間のネットワーク接続の問題により、データ分類がスキャンできないボリュームを示しています。



3. Data Classification インスタンスと、FSx for ONTAPのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。

FSx for ONTAPの場合、データ分類はコンソールと同じリージョン内のボリュームのみをスキャンできます。

4. NFS ボリュームのエクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていることを確認して、各ボリュームのデータにアクセスできるようにします。
5. CIFS を使用する場合は、Data Classification に Active Directory 資格情報を提供して、CIFS ボリュームをスキャンできるようにします。

   a. データ分類メニューから、*構成*を選択します。

   b. 各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

      データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

   資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。

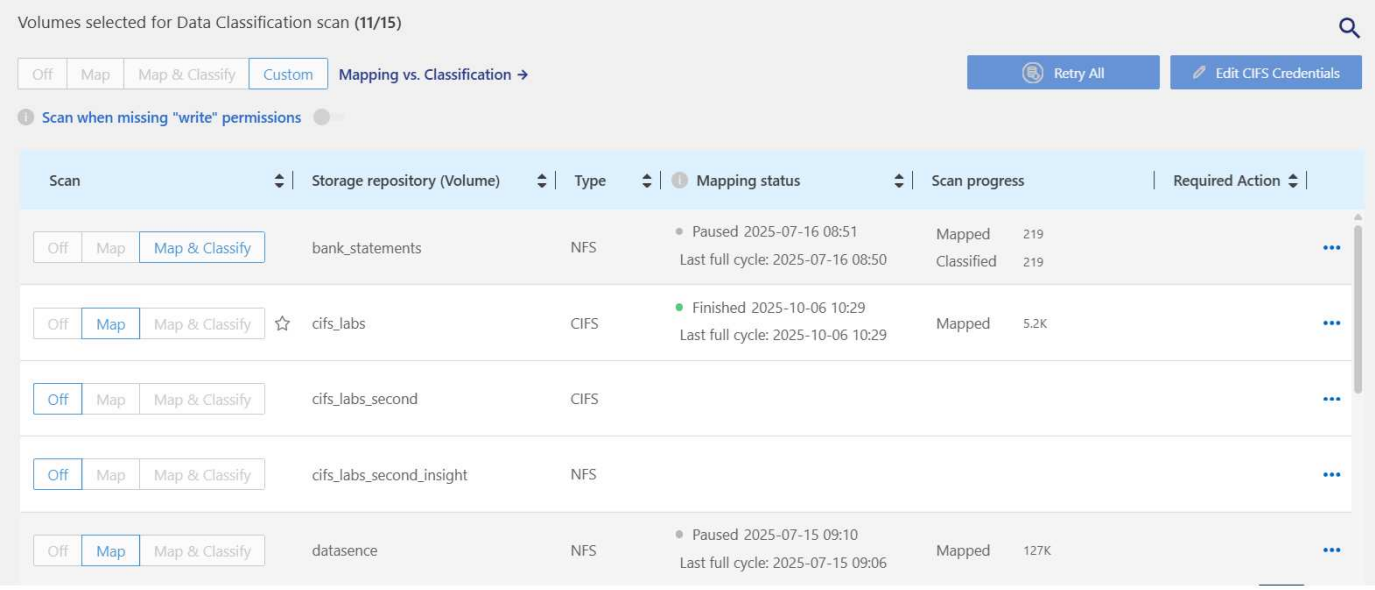
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

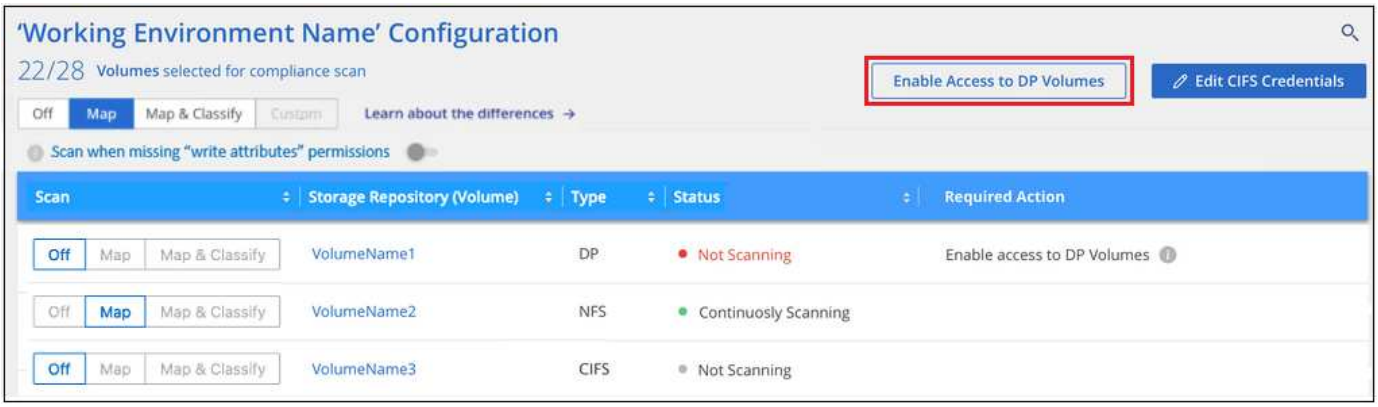
結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

データ保護ボリュームをスキャンする

デフォルトでは、データ保護 (DP) ボリュームは外部に公開されておらず、データ分類ではアクセスできないため、スキャンされません。これらは、FSx for ONTAPファイル システムからのSnapMirror操作の宛先ボリュームです。

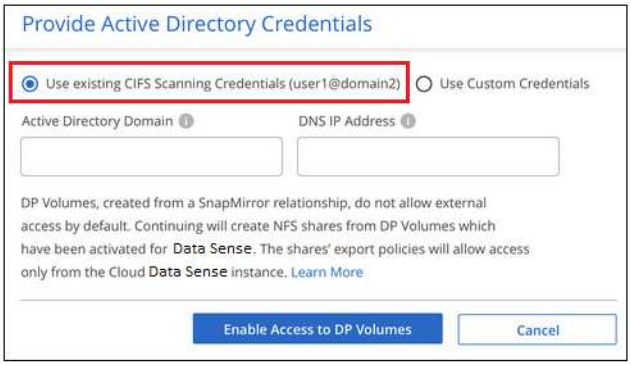
最初、ボリューム リストでは、これらのボリュームが、タイプ **DP**、ステータス スキャンなし、必要なアクション **DP** ボリュームへのアクセスを有効にする として識別されます。

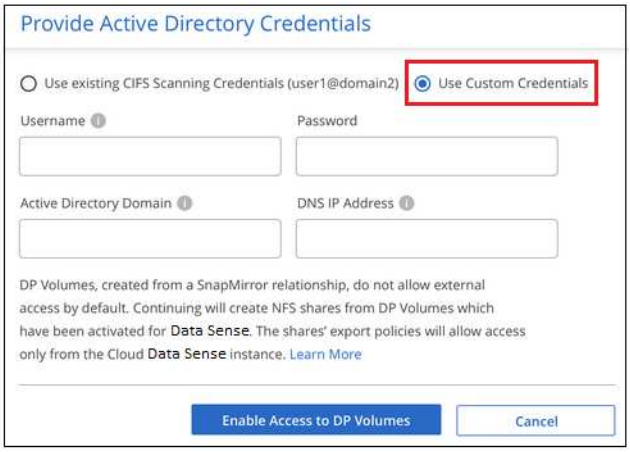


手順

これらのデータ保護ボリュームをスキャンする場合:

1. データ分類メニューから、*構成*を選択します。
2. ページの上部にある*DP ボリュームへのアクセスを有効にする*を選択します。
3. 確認メッセージを確認し、*DP ボリュームへのアクセスを有効にする*を再度選択します。
   - ソース FSx for ONTAPファイル システムで最初に NFS ボリュームとして作成されたボリュームが有効になります。
   - ソース FSx for ONTAPファイル システムで最初に CIFS ボリュームとして作成されたボリュームでは、それらの DP ボリュームをスキャンするために CIFS 認証情報を入力する必要があります。データ分類が CIFS ボリュームをスキャンできるように Active Directory 資格情報をすでに入力している場合は、その資格情報を使用することも、別の管理者資格情報セットを指定することもできます。





4. スキャンする各 DP ボリュームをアクティブ化します。

結果

有効にすると、データ分類はスキャン用にアクティブ化された各 DP ボリュームから NFS 共有を作成します。共有エクスポート ポリシーでは、データ分類インスタンスからのアクセスのみが許可されます。

最初に DP ボリュームへのアクセスを有効にしたときに CIFS データ保護ボリュームが存在せず、後でボリュームを追加すると、[構成] ページの上部に [**CIFS DP** へのアクセスを有効にする] ボタンが表示されます。このボタンを選択し、CIFS 資格情報を追加して、これらの CIFS DP ボリュームへのアクセスを有効にします。

Active Directory の資格情報は最初の CIFS DP ボリュームのストレージ VM にのみ登録されるため、その SVM 上のすべての DP ボリュームがスキャンされます。他の SVM に存在するボリュームには Active Directory 資格情報が登録されていないため、それらの DP ボリュームはスキャンされません。

NetApp Data Classificationを使用してAzure NetApp Filesボリュームをスキャンする

Azure NetApp FilesのNetApp Data Classificationを開始するには、いくつかの手順を完了します。

スキャンする**Azure NetApp Files**システムを検出します

スキャンするAzure NetApp FilesシステムがNetApp Consoleにシステムとしてまだ存在しない場合は、"システムページに追加します" 。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Azure NetApp Filesボリュームをスキャンするときは、データ分類をクラウドにデプロイする必要があり、スキャンするボリュームと同じリージョンにデプロイする必要があります。

注: Azure NetApp Filesボリュームをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

システムでデータ分類を有効にする

Azure NetApp Filesボリュームでデータ分類を有効にすることができます。

1. データ分類メニューから、*構成*を選択します。

タブのスクリーンショット。"]

2. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ" ：
   - すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
   - すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
   - 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップするボリューム、またはマップして分類するボリュームを選択します。

見るボリュームのスキャンを有効または無効にする詳細については。

3. 確認ダイアログボックスで*承認*を選択します。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できます。かかる時間はデータの量によって異なり、数分または数時間かかる場合があります。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。データ分類では、スキャンごとに進行状況バーが表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数が表示されます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"データ分類の制限について学ぶ" 。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームにアクセスできるようにするには、Data Classification に CIFS 資格情報を提供する必要があります。

Azure NetApp Filesの場合、データ分類ではコンソールと同じリージョン内のボリュームのみをスキャンできます。

チェックリスト

- データ分類インスタンスと、 Azure NetApp Filesのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。
- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにします。

手順

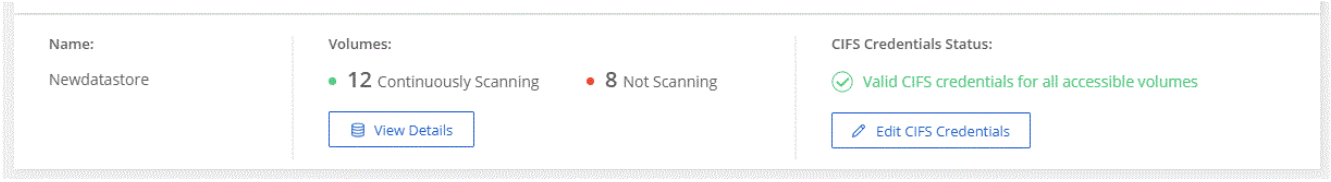
1. データ分類メニューから、*構成*を選択します。

   a. CIFS (SMB) を使用している場合は、Active Directory の資格情報が正しいことを確認してください。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができます。管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。



2. [構成] ページで [詳細の表示] を選択して、各 CIFS ボリュームと NFS ボリュームのステータスを確認します。必要に応じて、ネットワーク接続の問題などのエラーを修正します。

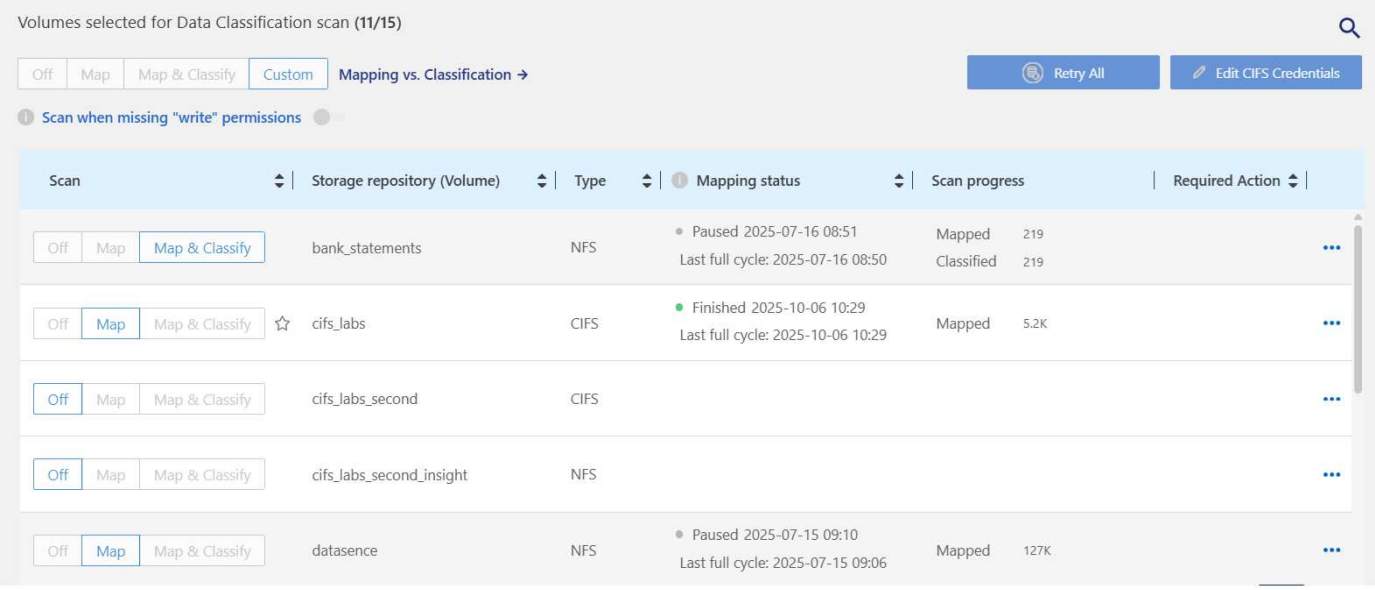
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

NetApp Data Classificationを使用してCloud Volumes ONTAPとオンプレミスのONTAPボリュームをスキャンします。

NetApp Data Classificationを使用してCloud Volumes ONTAPとオンプレミスのONTAPボリュームのスキャンを開始するには、いくつかの手順を実行します。

前提条件

データ分類を有効にする前に、サポートされている構成があることを確認してください。

- インターネット経由でアクセス可能なCloud Volumes ONTAPおよびオンプレミスのONTAPシステムをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所"。
- インターネットにアクセスできないダークサイトにインストールされているオンプレミスのONTAPシステムをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントを同じオンプレミスの場所に展開する必要があります。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーをチェックして、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームにアクセスできるようにするには、データ分類に CIFS 資格情報を提供する必要があります。

チェックリスト

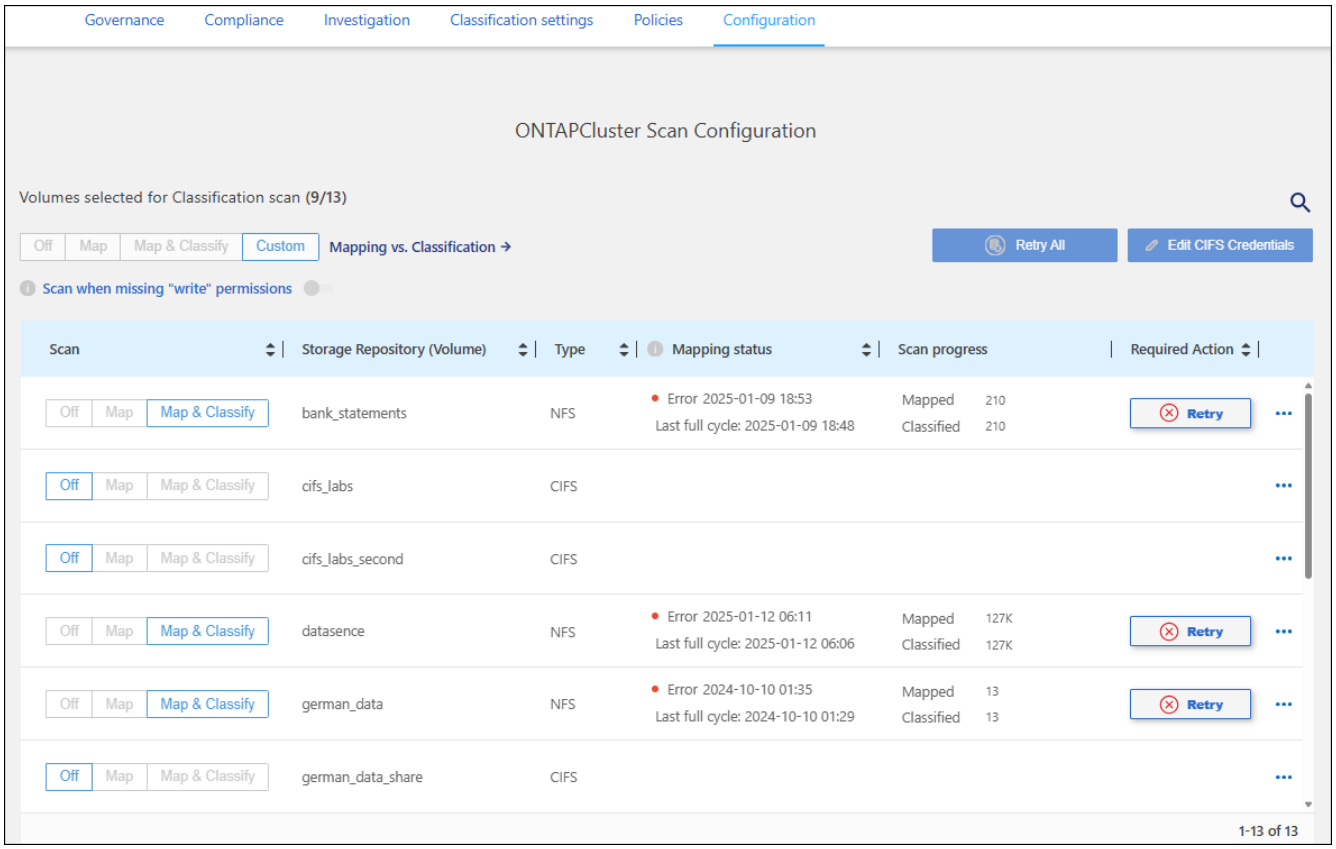
- Data Classification インスタンスと、 Cloud Volumes ONTAPまたはオンプレミスのONTAPクラスターのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- Cloud Volumes ONTAPのセキュリティ グループが、データ分類インスタンスからの受信トラフィックを許可していることを確認します。

  データ分類インスタンスの IP アドレスからのトラフィックに対してセキュリティ グループを開くことも、仮想ネットワーク内からのすべてのトラフィックに対してセキュリティ グループを開くこともできます。

- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにしていることを確認します。

手順

1. データ分類メニューから、*構成*を選択します。



タブのスクリーンショット。"]

2. CIFS を使用する場合は、Data Classification に Active Directory 資格情報を提供して、CIFS ボリュームをスキャンできるようにします。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

資格情報を正しく入力した場合、すべての CIFS ボリュームが正常に認証されたことを確認するメッセージが表示されます。

3. [構成] ページで [構成] を選択し、各 CIFS ボリュームと NFS ボリュームのステータスを確認し、エラーを修正します。

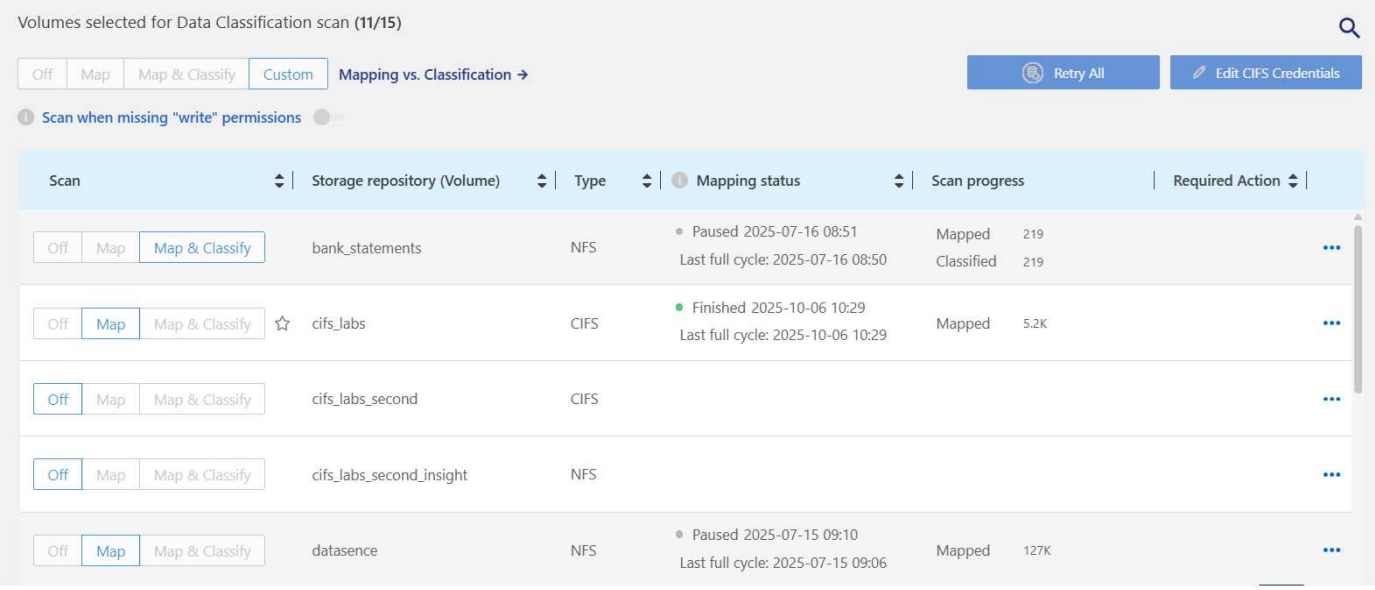
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"このデータ分類の制限に関する詳細を見る" 。

NetApp Data Classificationでデータベーススキーマをスキャンする

NetApp Data Classificationを使用してデータベース スキーマのスキャンを開始するには、いくつかの手順を実行します。

前提条件を確認する

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

サポートされているデータベース

データ分類では、次のデータベースからスキーマをスキャンできます。

- Amazon リレーショナルデータベースサービス (Amazon RDS)
- MongoDB
- MySQL
- Oracle
- PostgreSQL
- SAP HANA
- SQL サーバー (MSSQL)

データベースで統計収集機能が有効になっている必要があります。

データベース要件

ホストされている場所に関係なく、データ分類インスタンスに接続できるデータベースであればスキャンできます。データベースに接続するには、次の情報が必要です。

- IPアドレスまたはホスト名
- ポート
- サービス名（Oracle データベースへのアクセスのみ）
- スキーマへの読み取りアクセスを許可する資格情報

  ユーザー名とパスワードを選択するときは、スキャンするすべてのスキーマとテーブルに対する完全な読み取り権限を持つユーザー名とパスワードを選択することが重要です。必要なすべての権限を持つデータ分類システム専用のユーザーを作成することをお勧めします。

MongoDB の場合、読み取り専用の管理者ロールが必要です。

データ分類インスタンスをデプロイする

インスタンスがまだデプロイされていない場合は、データ分類をデプロイします。

インターネット経由でアクセス可能なデータベーススキーマをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所にデータ分類を展開する"。

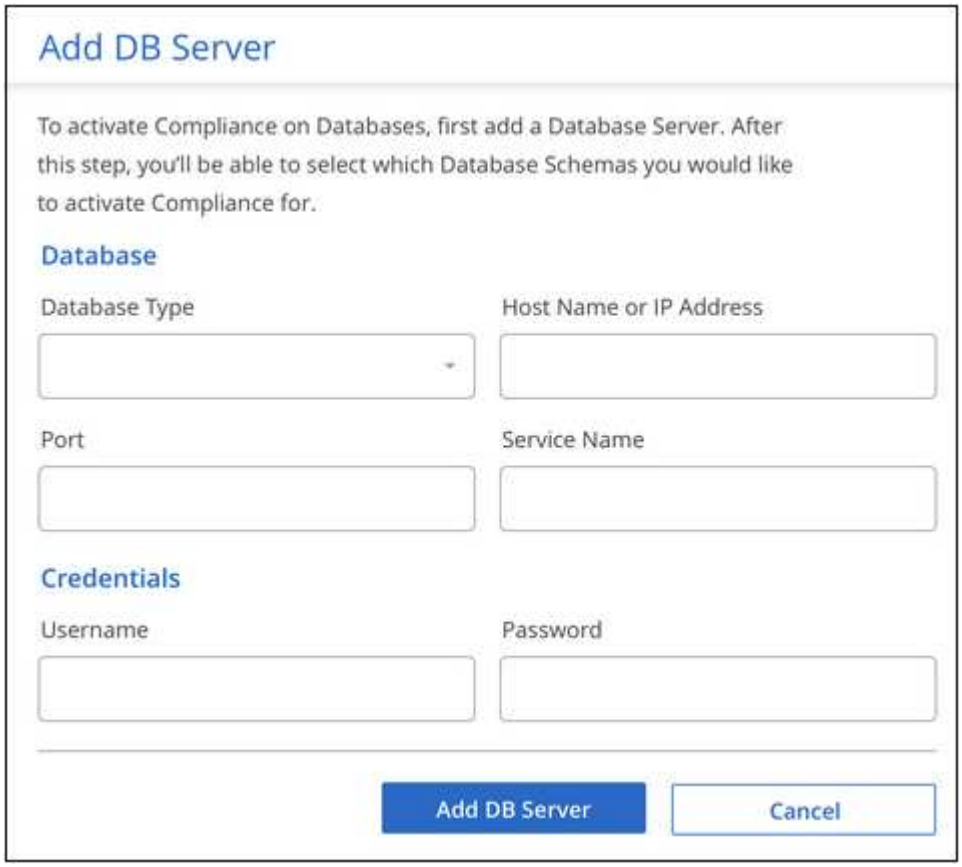
インターネットにアクセスできないダークサイトにインストールされているデータベーススキーマをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントが同じオンプレミスの場所に展開されていることも必要です。

データベースサーバーを追加する

スキーマが存在するデータベース サーバーを追加します。

1. データ分類メニューから、*構成*を選択します。
2. 構成ページで、システムの追加 > *データベース サーバーの追加*を選択します。
3. データベース サーバーを識別するために必要な情報を入力します。
   a. データベースの種類を選択します。

b. データベースに接続するためのポートとホスト名または IP アドレスを入力します。
c. Oracle データベースの場合は、サービス名を入力します。
d. データ分類がサーバーにアクセスできるように資格情報を入力します。
e. *DB サーバーの追加*を選択します。



データベースがシステムのリストに追加されます。

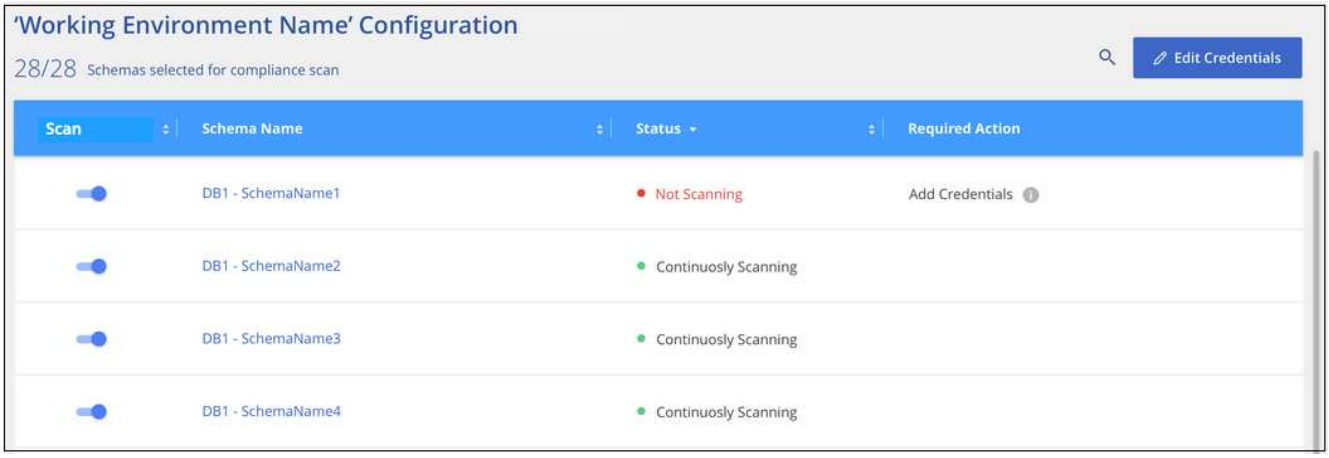
データベーススキーマのスキャンを有効または無効にする

スキーマの完全スキャンはいつでも停止または開始できます。

データベース スキーマのマッピングのみのスキャンを選択するオプションはありません。

1. 構成ページで、構成するデータベースの*構成*ボタンを選択します。

2. スライダーを右に移動して、スキャンするスキーマを選択します。



結果

データ分類は、有効にしたデータベース スキーマのスキャンを開始します。構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。各スキャンの進行状況は進行状況バーとして表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。エラーがある場合は、エラーを修正するために必要なアクションとともにステータス列に表示されます。

データ分類では、データベースを 1 日に 1 回スキャンします。データベースは他のデータ ソースのように継続的にスキャンされるわけではありません。

NetApp Data Classificationを使用してGoogle Cloud NetApp Volumesをスキャンする

NetApp Data Classification は、システムとしてGoogle Cloud NetApp Volumes をサポートします。 Google Cloud NetApp Volumesシステムをスキャンする方法を学びます。

スキャンするGoogle Cloud NetApp Volumesシステムを検出します

スキャンしたいGoogle Cloud NetApp VolumesシステムがNetApp Consoleにシステムとしてまだ登録されていない場合は、"システムページに追加する" 。

データ分類インスタンスをデプロイする

"データ分類を展開する"インスタンスがまだデプロイされていない場合。

Google Cloud NetApp Volumesをスキャンするときは、データ分類をクラウドにデプロイする必要があり、スキャンするボリュームと同じリージョンにデプロイする必要があります。

注: Google Cloud NetApp Volumesをスキャンする場合、オンプレミスの場所でのデータ分類の展開は現在サポートされていません。

システムでデータ分類を有効にする

Google Cloud NetApp Volumesシステムでデータ分類を有効にすることができます。

1. データ分類メニューから、*構成*を選択します。

2. 各システム内のボリュームをスキャンする方法を選択します。"マッピングと分類スキャンについて学ぶ"：
   - すべてのボリュームをマップするには、「すべてのボリュームをマップ」を選択します。
   - すべてのボリュームをマップして分類するには、[すべてのボリュームをマップして分類] を選択します。
   - 各ボリュームのスキャンをカスタマイズするには、[または各ボリュームのスキャン タイプを選択] を選択し、マップおよび/または分類するボリュームを選択します。

     見るボリュームのスキャンを有効または無効にする詳細については。
3. 確認ダイアログボックスで*承認*を選択します。

結果

データ分類は、システムで選択したボリュームのスキャンを開始します。データ分類が初期スキャンを完了するとすぐに、コンプライアンス ダッシュボードで結果を確認できます。かかる時間はデータの量によって異なり、数分から数時間かかります。初期スキャンの進行状況は、[構成] メニューの [システム構成] セクションで追跡できます。データ分類では、スキャンごとに進行状況バーが表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。

- デフォルトでは、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のファイルをスキャンしません。最終アクセス時刻がリセットされても構わない場合は、「またはボリュームごとにスキャン タイプを選択」を選択します。表示されるページには、権限に関係なくデータ分類がボリュームをスキャンするように有効にできる設定があります。
- データ分類では、ボリューム下の 1 つのファイル共有のみをスキャンします。ボリューム内に複数の共有がある場合は、他の共有を共有グループとして個別にスキャンする必要があります。"データ分類の制限について学ぶ"。

データ分類がボリュームにアクセスできることを確認する

ネットワーク、セキュリティ グループ、エクスポート ポリシーを確認して、データ分類がボリュームにアクセスできることを確認します。 CIFS ボリュームの場合、データ分類に CIFS 資格情報を提供する必要があります。

Google Cloud NetApp Volumesの場合、データ分類ではコンソールと同じリージョン内のボリュームのみをスキャンできます。

チェックリスト

- データ分類インスタンスと、 Google Cloud NetApp Volumesのボリュームを含む各ネットワークとの間にネットワーク接続があることを確認します。
- データ分類インスタンスに対して次のポートが開いていることを確認します。
  - NFS の場合 - ポート 111 および 2049。
  - CIFS の場合 - ポート 139 および 445。
- NFS ボリューム エクスポート ポリシーにデータ分類インスタンスの IP アドレスが含まれていて、各ボリュームのデータにアクセスできるようにします。

手順

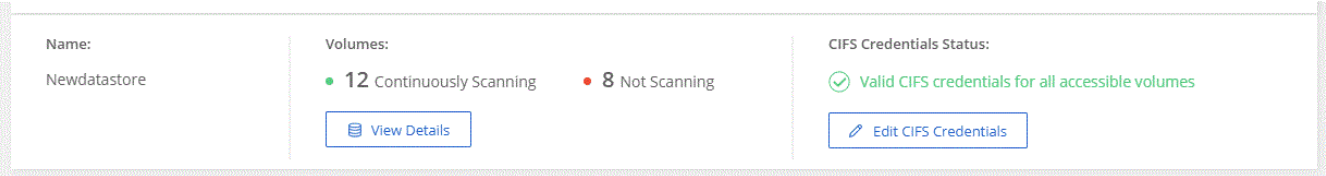
1. データ分類メニューから、*構成*を選択します。

   a. CIFS (SMB) を使用している場合は、Active Directory の資格情報が正しいことを確認してください。各システムについて、「**CIFS** 資格情報の編集」を選択し、データ分類がシステム上の CIFS ボリュームにアクセスするために必要なユーザー名とパスワードを入力します。

      資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

      データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

      資格情報を入力すると、すべての CIFS ボリュームが正常に認証されたことを示すメッセージが表示されます。



2. [構成] ページで [詳細の表示] を選択し、各 CIFS ボリュームと NFS ボリュームのステータスを確認し、エラーを修正します。

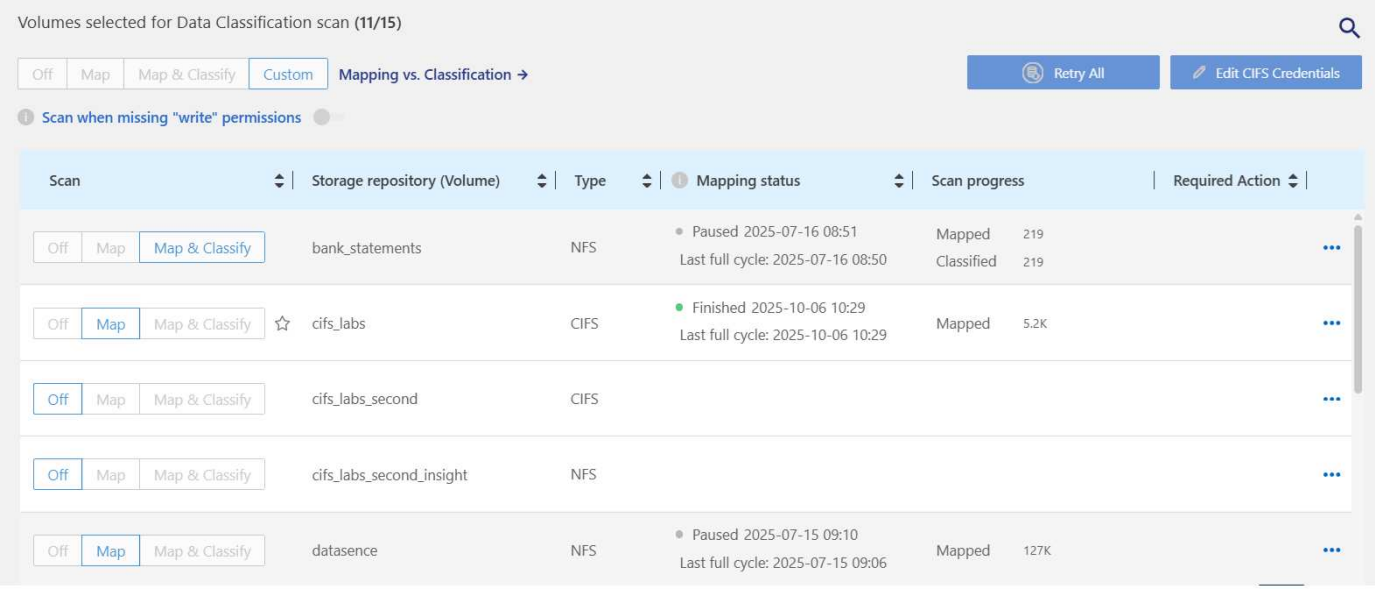
ボリュームのスキャンを有効または無効にする

構成ページからいつでも任意のシステムのスキャンを開始または停止できます。スキャンをマッピングのみのスキャンからマッピングと分類のスキャンに切り替えることも、その逆に切り替えることもできます。システム内のすべてのボリュームをスキャンすることをお勧めします。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を選択した場合にのみ自動的にスキャンされます。見出し領域で カスタム または オフ に設定すると、システムに追加する新しいボリュームごとにマッピングや完全スキャンをアクティブ化する必要があります。

ページ上部の「書き込み権限がない場合にスキャンする」スイッチは、デフォルトで無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。最終アクセス時間がリセットされても構わない場合は、スイッチをオンにすると、権限に関係なくすべてのファイルがスキャンされます。"詳細情報"。

システムに追加された新しいボリュームは、見出し領域で マップ または マップと分類 設定を行った場合にのみ自動的にスキャンされます。すべてのボリュームの設定が カスタム または オフ の場合、追加する新しいボリュームごとにスキャンを手動で有効にする必要があります。



手順

1. データ分類メニューから、*構成*を選択します。
2. システムを選択し、*構成*を選択します。
3. すべてのボリュームのスキャンを有効または無効にするには、すべてのボリュームの上にある見出しで [マップ]、[マップと分類]、または [オフ] を選択します。

   個々のボリュームのスキャンを有効または無効にするには、リストでボリュームを見つけて、ボリューム名の横にある [マップ]、[マップと分類]、または [オフ] を選択します。

結果

スキャンを有効にすると、データ分類はシステムで選択したボリュームのスキャンを開始します。データ分類がスキャンを開始するとすぐに、コンプライアンス ダッシュボードに結果が表示され始めます。スキャンの完了時間はデータの量に応じて数分から数時間の範囲になります。

NetApp Data Classificationでファイル共有をスキャンする

ファイル共有をスキャンするには、まずNetApp Data Classificationでファイル共有グループを作成する必要があります。ファイル共有グループは、オンプレミスまたはクラウドでホストされる NFS または CIFS (SMB) 共有用です。

データ分類コア バージョンでは、 NetApp以外のファイル共有からのデータのスキャンはサポートされていません。

前提条件

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

- 共有は、クラウドやオンプレミスなど、どこにでもホストできます。古いNetApp 7-Mode ストレージ システムの CIFS 共有は、ファイル共有としてスキャンできます。
  - データ分類では、7-Mode システムから権限や「最終アクセス時刻」を抽出できません。

- 一部の Linux バージョンと 7-Mode システム上の CIFS 共有の間に既知の問題があるため、NTLM 認証が有効になっている SMBv1 のみを使用するように共有を構成する必要があります。

- データ分類インスタンスと共有の間にはネットワーク接続が必要です。
- DFS (分散ファイル システム) 共有を通常の CIFS 共有として追加できます。データ分類では、共有が複数のサーバー/ボリューム上に構築され、単一の CIFS 共有として結合されていることを認識しないため、メッセージが実際には別のサーバー/ボリュームにあるフォルダー/共有の 1 つにのみ適用される場合でも、共有に関するアクセス許可または接続エラーが発生する可能性があります。
- CIFS (SMB) 共有の場合は、共有への読み取りアクセスを提供する Active Directory 資格情報があることを確認します。データ分類で昇格された権限を必要とするデータをスキャンする必要がある場合は、管理者の資格情報が優先されます。

  データ分類スキャンによってファイルの「最終アクセス時刻」が変更されないようにしたい場合は、ユーザーに CIFS での属性書き込み権限または NFS での書き込み権限を与えることをお勧めします。可能であれば、Active Directory ユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部として構成します。

- グループ内のすべての CIFS ファイル共有では、同じ Active Directory 資格情報を使用する必要があります。
- NFS と CIFS (Kerberos または NTLM のいずれかを使用) の共有を混在させることができます。共有をグループに個別に追加する必要があります。つまり、プロトコルごとに 1 回ずつ、プロセスを 2 回完了する必要があります。
  - CIFS 認証タイプ (Kerberos と NTLM) が混在するファイル共有グループを作成することはできません。
- Kerberos 認証で CIFS を使用している場合は、提供された IP アドレスがデータ分類にアクセスできることを確認してください。 IP アドレスに到達できない場合は、ファイル共有を追加できません。

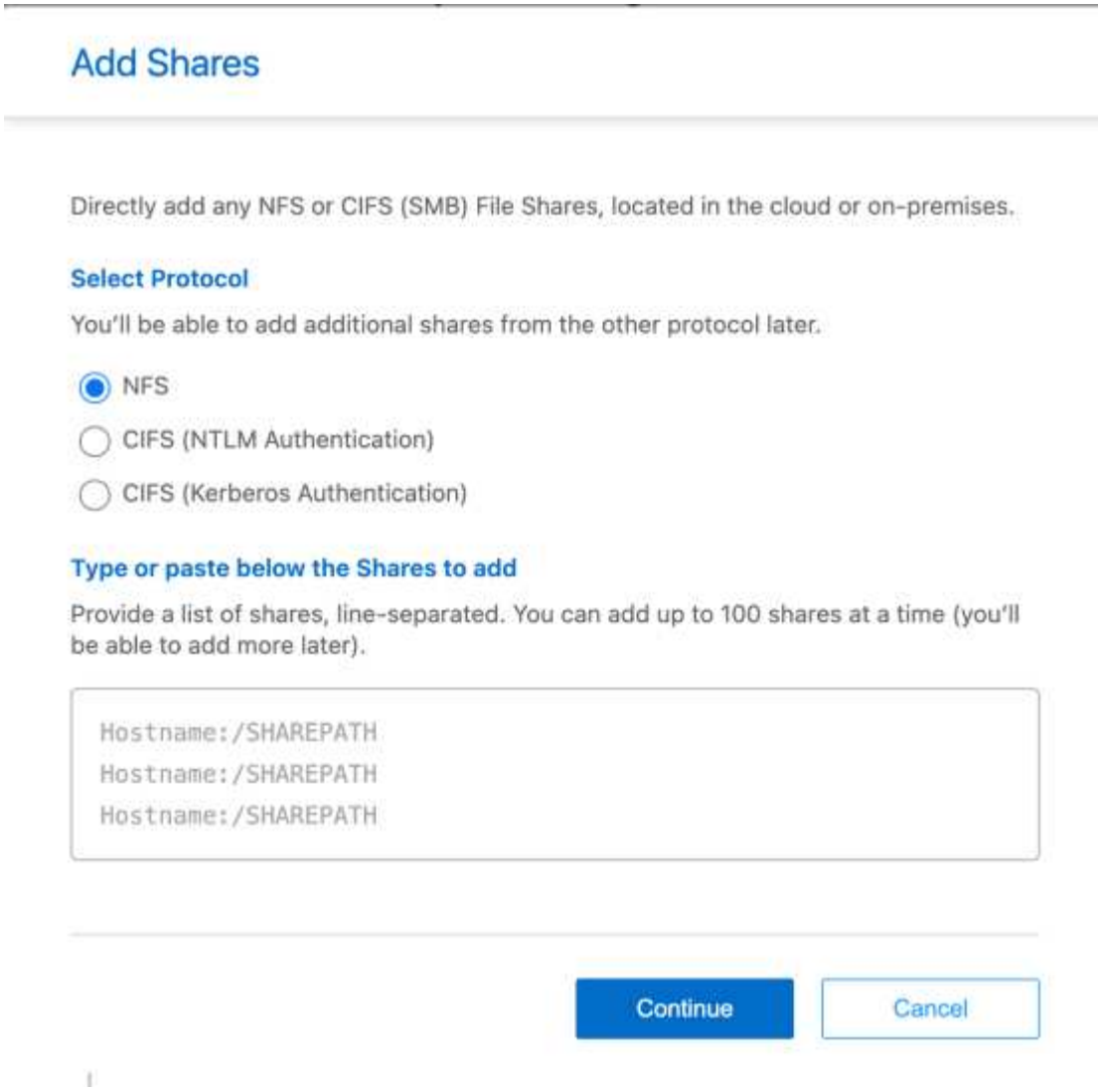
ファイル共有グループを作成する

グループにファイル共有を追加するときは、次の形式を使用する必要があります。
`<host_name>:/<share_path>` 。

ファイル共有を個別に追加することも、スキャンするファイル共有の行区切りリストを入力することもできます。一度に追加できる株式数は最大 100 です。

手順

1. データ分類メニューから、*構成*を選択します。
2. 構成ページで、システムの追加 > *ファイル共有グループの追加*を選択します。
3. [ファイル共有グループの追加] ダイアログで、共有グループの名前を入力し、[続行] を選択します。
4. 追加するファイル共有のプロトコルを選択します。



a. NTLM 認証を使用して CIFS 共有を追加する場合は、Active Directory の資格情報を入力して CIFS ボリュームにアクセスします。読み取り専用の資格情報はサポートされていますが、管理者の資格情報を使用してフルアクセスを提供することをお勧めします。 保存を選択します。

5. スキャンするファイル共有を追加します (1 行につき 1 つのファイル共有)。次に、[続行] を選択します。
6. 確認ダイアログに追加された共有数が表示されます。

   ダイアログに追加できなかった共有がリストされている場合は、問題を解決できるようにこの情報を取得します。問題が命名規則に関係する場合は、修正した名前で共有を再度追加できます。

7. ボリュームのスキャンを構成します。
   - ファイル共有でマッピングのみのスキャンを有効にするには、[マップ] を選択します。
   - ファイル共有の完全スキャンを有効にするには、[マップと分類] を選択します。
   - ファイル共有のスキャンを無効にするには、[オフ] を選択します。

ページ上部の「「属性の書き込み」権限がない場合にスキャンする」スイッチは、デフォルトでは無効になっています。つまり、データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはファイルをスキャンしません。+ *「属性の書き込み」権限がない場合にスキャン*を*オン*に切り替えると、スキャンは最終アクセス時刻をリセットし、権限に関係なくすべてのファイルをスキャンします。+ 最終アクセスタイムスタンプの詳細については、以下を参照してください。"データ分類のデータソースから収集されたメタデータ" 。

結果

データ分類は、追加したファイル共有内のファイルのスキャンを開始します。あなたはできるxref:#track-the-scanning-progressダッシュボードでスキャンの結果を表示します。

Kerberos 認証を使用した CIFS 構成のスキャンが正常に完了しない場合は、[構成] タブでエラーを確認してください。

ファイル共有グループを編集する

ファイル共有グループを作成した後、CIFS プロトコルを編集したり、ファイル共有を追加および削除したりできます。

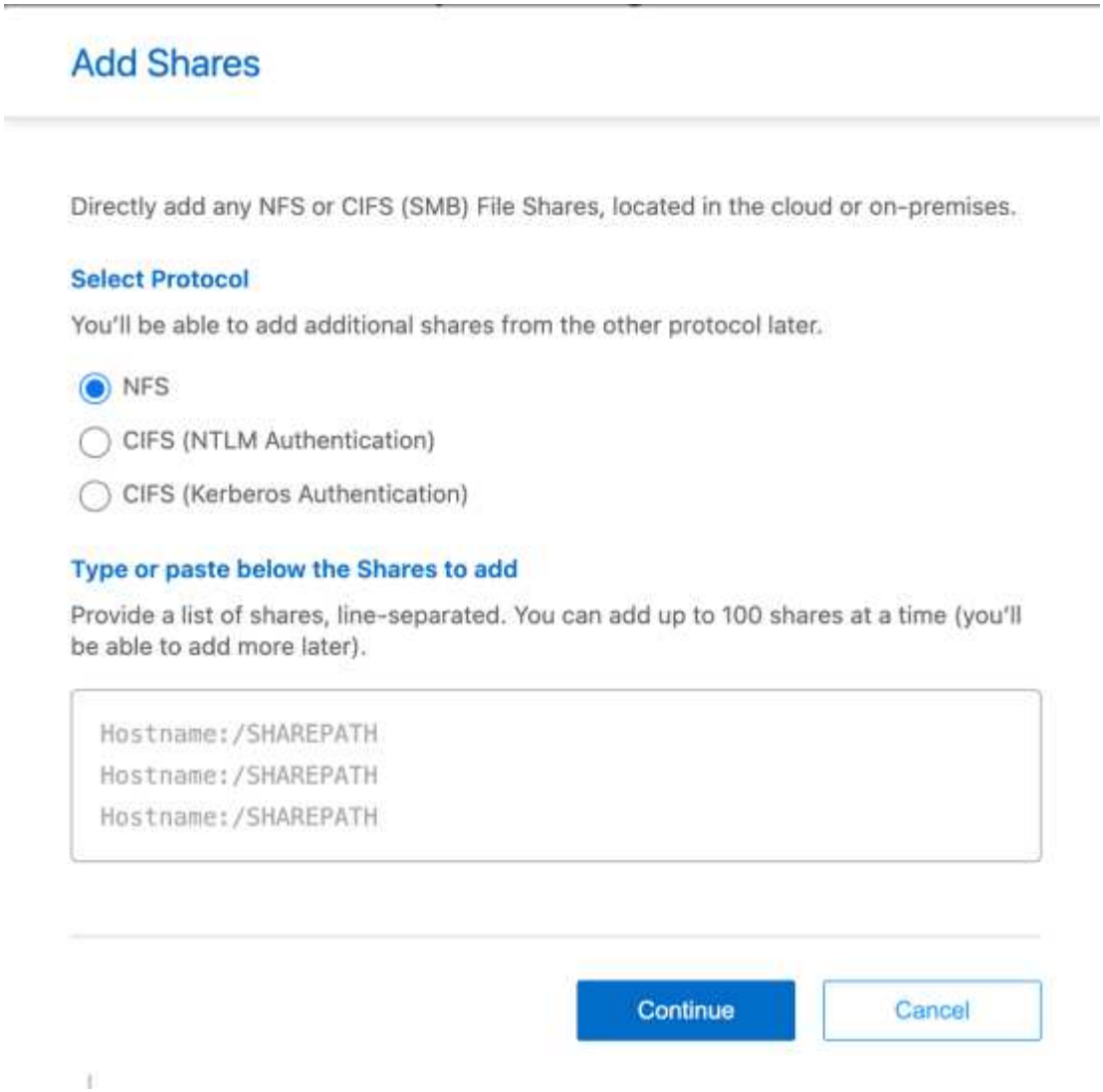
CIFSプロトコル設定を編集する

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで、変更するファイル共有グループを選択します。
3. **CIFS** 資格情報の編集 を選択します。

4. 認証方法を選択します: **NTLM** または **Kerberos**。
5. Active Directory の ユーザー名 と パスワード を入力します。
6. プロセスを完了するには、[保存] を選択します。

スキャンにファイル共有を追加する

1. データ分類メニューから、*構成*を選択します。
2. [構成] ページで、変更するファイル共有グループを選択します。
3. **+** 共有を追加 を選択します。
4. 追加するファイル共有のプロトコルを選択します。



すでに構成済みのプロトコルにファイル共有を追加する場合、変更は必要ありません。

2番目のプロトコルでファイル共有を追加する場合は、認証が適切に設定されていることを確認してください。"前提条件" 。

5. スキャンするファイル共有（1行につき1つのファイル共有）を次の形式で追加します。 `<host_name>:/<share_path>` 。
6. ファイル共有の追加を完了するには、[続行] を選択します。

スキャンからファイル共有を削除する

1. データ分類メニューから、*構成*を選択します。
2. ファイル共有を削除するシステムを選択します。
3. *構成*を選択します。
4. 設定ページからアクションを選択します ••• 削除するファイル共有の。
5. [アクション] メニューから、[共有を削除] を選択します。

スキャンの進行状況を追跡する

初期スキャンの進行状況を追跡できます。

1. 構成 メニューを選択します。
2. システム構成を選択します。
3. ストレージ リポジトリの場合は、スキャン進行状況列をチェックしてステータスを表示します。

NetApp Data ClassificationでStorageGRIDデータをスキャン

いくつかの手順を完了すると、 NetApp Data Classificationを使用してStorageGRID内のデータを直接スキャンし始めることができます。

StorageGRIDの要件を確認する

データ分類を有効にする前に、次の前提条件を確認して、サポートされている構成があることを確認してください。

- オブジェクト ストレージ サービスに接続するには、エンドポイント URL が必要です。
- データ分類がバケットにアクセスできるようにするには、 StorageGRIDからのアクセス キーとシークレット キーが必要です。

データ分類インスタンスをデプロイする

インスタンスがまだデプロイされていない場合は、データ分類をデプロイします。

インターネット経由でアクセス可能なStorageGRIDからデータをスキャンする場合は、"クラウドでデータ分類を展開する"または"インターネットにアクセスできるオンプレミスの場所にデータ分類を展開する"。

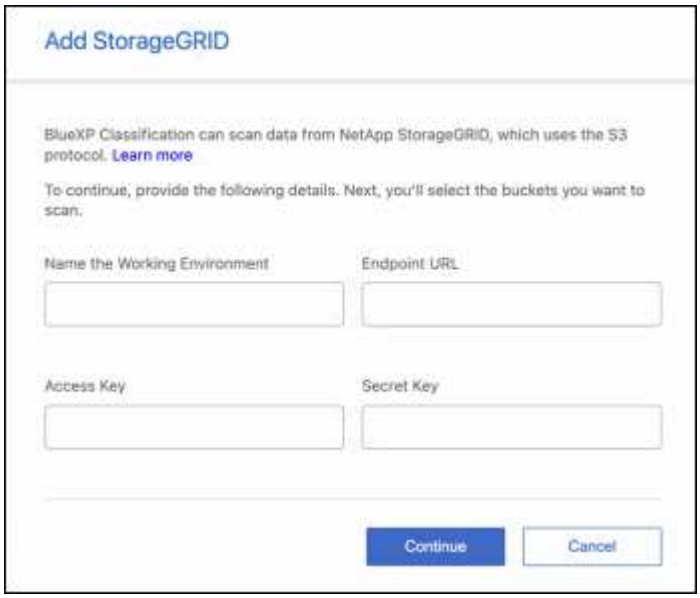
インターネットにアクセスできない暗い場所にインストールされているStorageGRIDからデータをスキャンする場合は、"インターネットにアクセスできないオンプレミスの同じ場所にデータ分類を展開する" 。これには、コンソール エージェントが同じオンプレミスの場所に展開されていることも必要です。

データ分類にStorageGRIDサービスを追加する

StorageGRIDサービスを追加します。

手順

1. データ分類メニューから、*構成*オプションを選択します。
2. 構成ページで、システムの追加 > * StorageGRIDの追加*を選択します。
3. 「StorageGRIDサービスの追加」ダイアログで、 StorageGRIDサービスの詳細を入力し、「続行」を選択します。
   a. システムに使用する名前を入力します。この名前は、接続先のStorageGRIDサービスの名前を反映する必要があります。
   b. オブジェクト ストレージ サービスにアクセスするためのエンドポイント URL を入力します。
   c. データ分類がStorageGRID内のバケットにアクセスできるように、アクセス キーとシークレット キーを入力します。



結果

StorageGRIDがシステムのリストに追加されます。

StorageGRIDバケットのスキャンを有効または無効にする

StorageGRIDでデータ分類を有効にした後、次のステップはスキャンするバケットを構成することです。データ分類はこれらのバケットを検出し、作成したシステムに表示します。

手順

1. 構成ページで、 StorageGRIDシステムを見つけます。
2. StorageGRIDシステム タイルで、[構成] を選択します。
3. スキャンを有効または無効にするには、次のいずれかの手順を実行します。
   - バケットでマッピングのみのスキャンを有効にするには、[マップ] を選択します。
   - バケットの完全スキャンを有効にするには、[マップと分類] を選択します。
   - バケットのスキャンを無効にするには、[オフ] を選択します。

結果

データ分類は、有効にしたバケットのスキャンを開始します。 構成 メニューに移動し、システム構成 を選択すると、初期スキャンの進行状況を追跡できます。各スキャンの進行状況は進行状況バーとして表示されます。進行状況バーにマウスを移動すると、ボリューム内のファイルの合計数に対するスキャンされたファイルの数を確認することもできます。エラーがある場合は、エラーを修正するために必要なアクションとともにステータス列に表示されます。

Active Directory とNetApp Data Classification を統合

グローバル Active Directory をNetApp Data Classificationと統合して、ファイル所有者やファイルにアクセスできるユーザーとグループに関する Data Classification のレポート結果を強化できます。

特定のデータ ソース (以下にリスト) を設定する場合、データ分類で CIFS ボリュームをスキャンするには、Active Directory の資格情報を入力する必要があります。この統合により、データ分類では、それらのデータ ソースに存在するデータのファイル所有者と権限の詳細が提供されます。これらのデータ ソースに入力された Active Directory は、ここで入力するグローバル Active Directory 資格情報とは異なる場合があります。データ分類では、統合されたすべての Active Directory でユーザーと権限の詳細を検索します。

この統合により、データ分類の次の場所に追加情報が提供されます。

- 「ファイル所有者」を使用することができます"フィルター"調査ペインのファイルのメタデータで結果を確認します。 SID (セキュリティ ID) を含むファイル所有者の代わりに、実際のユーザー名が入力されます。

  また、ファイル所有者の詳細情報（アカウント名、メール アドレス、SAM アカウント名）を表示したり、そのユーザーが所有するアイテムを表示したりすることもできます。

- 見ることができます"完全なファイル権限"「すべての権限を表示」ボタンをクリックすると、各ファイルとディレクトリに対して権限が表示されます。
- の中で"ガバナンスダッシュボード"、Open Permissions パネルにデータに関するより詳細な情報が表示されます。

ローカル ユーザーの SID および不明なドメインの SID は、実際のユーザー名に変換されません。

サポートされているデータソース

Active Directory とデータ分類の統合により、次のデータ ソース内のデータを識別できます。

- オンプレミスのONTAPシステム
- Cloud Volumes ONTAP
- Azure NetApp Files
- ONTAP向け FSx

Active Directoryサーバーに接続する

データ分類を展開し、データ ソースのスキャンを有効にしたら、データ分類を Active Directory と統合できます。 Active Directory には、DNS サーバー IP アドレスまたは LDAP サーバー IP アドレスを使用してアクセスできます。

Active Directory の資格情報は読み取り専用にすることができますが、管理者の資格情報を提供することで、データ分類は昇格された権限を必要とするすべてのデータを読み取ることができるようになります。資格情報は、データ分類インスタンスに保存されます。

CIFS ボリューム/ファイル共有の場合、ファイルの「最終アクセス時刻」がデータ分類分類スキャンによって変更されていないことを確認するには、ユーザーに属性の書き込み権限が必要です。可能であれば、Active Directory で構成されたユーザーを、すべてのファイルへの権限を持つ組織内の親グループの一部にすることをお勧めします。

要件

- 社内のユーザー用に Active Directory がすでに設定されている必要があります。

- Active Directory の情報が必要です:
  - DNSサーバーのIPアドレス、または複数のIPアドレス

    または

  LDAPサーバーのIPアドレス、または複数のIPアドレス
  - サーバーにアクセスするためのユーザー名とパスワード
  - ドメイン名（アクティブディレクトリ名）
  - セキュアLDAP（LDAPS）を使用しているかどうか
  - LDAP サーバー ポート (通常、LDAP の場合は 389、セキュア LDAP の場合は 636)
- データ分類インスタンスによる送信通信のために、次のポートが開いている必要があります。

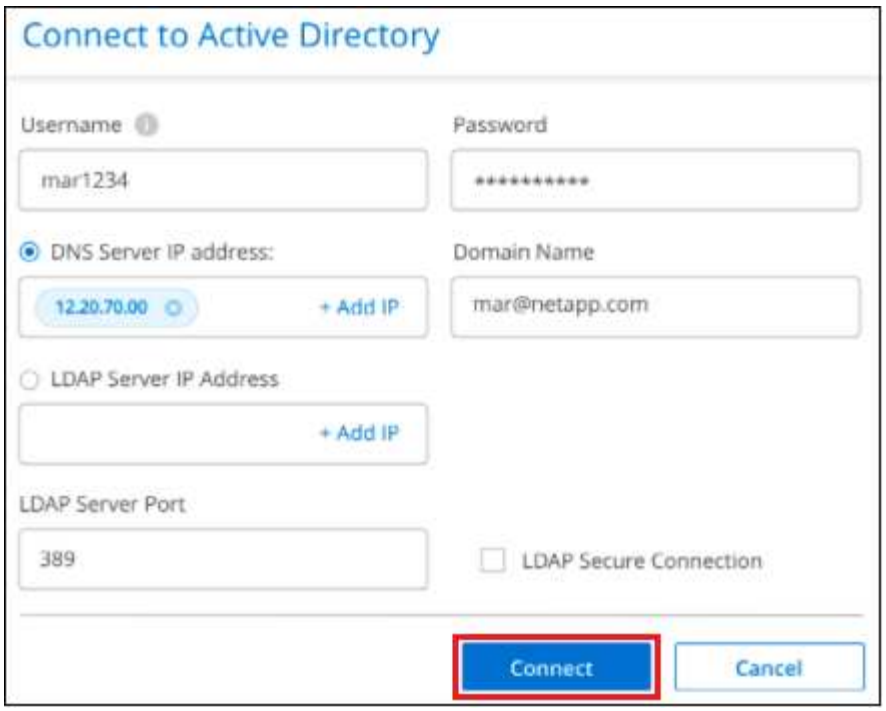
プロトコル	ポート	デスティネーション	目的
TCPとUDP	389	Active Directory	LDAP
TCP	636	Active Directory	LDAP over SSL
TCP	3268	Active Directory	グローバルカタログ
TCP	3269	Active Directory	SSL経由のグローバルカタログ

手順

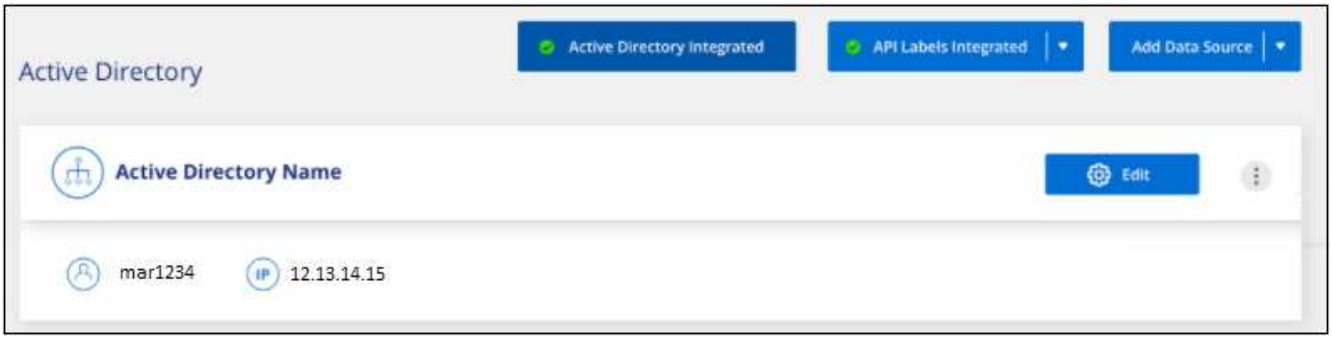
1. データ分類構成ページで、*Active Directory の追加*をクリックします。

2. 「Active Directory への接続」ダイアログで、Active Directory の詳細を入力し、「接続」をクリックします。

   必要に応じて、「**IP** を追加」を選択して複数の IP アドレスを追加できます。



データ分類は Active Directory に統合され、構成ページに新しいセクションが追加されます。



Active Directory統合を管理する

Active Directory 統合の値を変更する必要がある場合は、[編集] ボタンをクリックして変更を加えます。

統合を削除するには、 ボタンをクリックし、[Active Directory を削除] をクリックします。

著作権に関する情報

商標に関する情報