Documentation sur la NetApp Data Classification

NetApp Data Classification

NetApp
October 22, 2025

Sommaire

Documentation sur la NetApp Data Classification

Notes de version

Quoi de neuf dans la NetApp Data Classification

Découvrez les nouveautés de la NetApp Data Classification.

06 octobre 2025

Version 1.47

La BlueXP classification est désormais la NetApp Data Classification

La BlueXP classification a été renommée NetApp Data Classification. En plus du changement de nom, l'interface utilisateur a été améliorée.

BlueXP est désormais NetApp Console

BlueXP a été renommé et repensé pour mieux refléter son rôle dans la gestion de votre infrastructure de données.

La NetApp Console offre une gestion centralisée des services de stockage et de données dans les environnements sur site et dans le cloud à l'échelle de l'entreprise, offrant des informations en temps réel, des flux de travail plus rapides et une administration simplifiée.

Pour plus de détails sur ce qui a changé, consultez le "Notes de version de la NetApp Console" .

Expérience d'enquête améliorée

Recherchez et comprenez vos données plus rapidement grâce à de nouveaux filtres consultables, des décomptes de résultats par valeur, des informations en temps réel résumant les principales conclusions et un tableau de résultats actualisé avec des colonnes personnalisables et un volet de détails coulissant.

Pour plus d'informations, consultez la section "Enquêter sur les données" .

Nouveaux tableaux de bord de gouvernance et de conformité

Obtenez des informations essentielles plus rapidement grâce à des widgets intuitifs, des visuels plus clairs et des performances de chargement améliorées. Pour plus d'informations, voir"Consultez les informations de gouvernance sur vos données" et"Afficher les informations de conformité concernant vos données" .

Politiques pour les requêtes enregistrées (aperçu)

La classification des données vous permet désormais d'automatiser la gouvernance avec des actions conditionnelles. Vous pouvez créer des règles de conservation avec suppression automatique, configurer des notifications par e-mail périodiques, le tout géré à partir d'une page de requêtes enregistrées mise à jour.

Pour plus d'informations, consultez la section "Créer des politiques" .

Actions (aperçu)

Prenez le contrôle direct depuis la page Investigation : supprimez, déplacez, copiez ou étiquetez les fichiers individuellement ou en masse, pour une gestion et une correction efficaces des données.

Pour plus d'informations, consultez la section "Enquêter sur les données" .

Prise en charge des Google Cloud NetApp Volumes

La classification des données prend désormais en charge l'analyse sur les Google Cloud NetApp Volumes.

Ajoutez facilement des Google Cloud NetApp Volumes à partir de la NetApp Console pour une analyse et une classification transparentes des données. Pour plus d'informations, voir "Analyser les Google Cloud NetApp Volumes".

11 août 2025

Version 1.46

Cette version de classification des données inclut des corrections de bogues et les mises à jour suivantes :

Informations améliorées sur les événements d'analyse dans la page d'audit

La page Audit prend désormais en charge des informations améliorées sur les événements d'analyse pour la BlueXP classification. La page Audit affiche désormais le moment où l'analyse d'un système commence, les statuts des systèmes et les problèmes éventuels. Les statuts des partages et des systèmes ne sont disponibles que pour les analyses de mappage.

Pour plus d'informations sur la page Audit, voir"Surveiller les opérations de la NetApp Console" .

Prise en charge de RHEL 9.6

Cette version ajoute la prise en charge de Red Hat Enterprise Linux v9.6 pour l'installation manuelle sur site de la BlueXP classification, y compris les déploiements de sites sombres.

Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version de BlueXP classification 1.30 ou supérieure : Red Hat Enterprise Linux versions 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4 et 9.5.

14 juillet 2025

Version 1.45

Cette version de BlueXP classification inclut des modifications de code qui optimisent l'utilisation des ressources et :

Flux de travail amélioré pour ajouter des partages de fichiers pour la numérisation

Le flux de travail permettant d'ajouter des partages de fichiers à un groupe de partage de fichiers a été simplifié. Le processus différencie désormais également la prise en charge du protocole CIFS en fonction du type d'authentification (Kerberos ou NTLM).

Pour plus d'informations, consultez la section "Analyser les partages de fichiers" .

Informations améliorées sur le propriétaire du fichier

Vous pouvez désormais afficher plus d'informations sur les propriétaires de fichiers capturés dans l'onglet Enquête. Lorsque vous affichez les métadonnées d'un fichier dans l'onglet Enquête, recherchez le propriétaire du fichier, puis sélectionnez **Afficher les détails** pour voir le nom d'utilisateur, l'e-mail et le nom du compte SAM. Vous pouvez également afficher d'autres éléments appartenant à cet utilisateur. Cette fonctionnalité est uniquement disponible pour les environnements de travail avec Active Directory.

Pour plus d'informations, consultez la section "Examinez les données stockées dans votre organisation" .

10 juin 2025

Version 1.44

Cette version de BlueXP classification comprend :

Amélioration des temps de mise à jour du tableau de bord de gouvernance

Les temps de mise à jour des composants individuels du tableau de bord de gouvernance ont été améliorés. Le tableau suivant affiche la fréquence des mises à jour pour chaque composant.

Composant	Heures de mise à jour
L'ère des données	24 heures
Catégories	24 heures
Aperçu des données	5 minutes
Fichiers en double	2 heures
Types de fichiers	24 heures
Données non commerciales	2 heures
Autorisations d'ouverture	24 heures
Recherches enregistrées	2 heures
Données sensibles et autorisations étendues	24 heures
Taille des données	24 heures
Données obsolètes	2 heures
Principaux référentiels de données par niveau de sensibilité	2 heures

Vous pouvez afficher l'heure de la dernière mise à jour et mettre à jour manuellement les composants Fichiers en double, Données non commerciales, Recherches enregistrées, Données obsolètes et Principaux référentiels de données par niveau de sensibilité. Pour plus d'informations sur le tableau de bord de gouvernance, voir"Afficher les détails de gouvernance sur les données stockées dans votre organisation" .

Améliorations des performances et de la sécurité

Des améliorations ont été apportées pour améliorer les performances, la consommation de mémoire et la sécurité de la classification BlueXP .

Corrections de bugs

Redis a été mis à niveau pour améliorer la fiabilité de la BlueXP classification. La BlueXP classification utilise désormais Elasticsearch pour améliorer la précision des rapports sur le nombre de fichiers lors des analyses.

12 mai 2025

Version 1.43

Cette version de classification des données comprend :

Prioriser les analyses de classification

La classification des données prend en charge la possibilité de hiérarchiser les analyses de cartographie et de classification en plus des analyses de cartographie uniquement, vous permettant de sélectionner les analyses à effectuer en premier. La priorisation des analyses Map & Classify est prise en charge pendant et avant le

début des analyses. Si vous choisissez de donner la priorité à une analyse pendant qu'elle est en cours, les analyses de mappage et de classification sont toutes deux prioritaires.

Pour plus d'informations, consultez la section "Prioriser les analyses" .

Prise en charge des catégories de données d'informations personnelles identifiables (PII) canadiennes

Les analyses de classification des données identifient les catégories de données PII canadiennes. Ces catégories comprennent les renseignements bancaires, les numéros de passeport, les numéros d'assurance sociale, les numéros de permis de conduire et les numéros de carte d'assurance-maladie pour toutes les provinces et tous les territoires canadiens.

Pour plus d'informations, consultez la section "Catégories de données personnelles" .

Classification personnalisée (aperçu)

La classification des données prend en charge les classifications personnalisées pour les analyses Map & Classify. Grâce aux classifications personnalisées, vous pouvez personnaliser les analyses de classification des données pour capturer des données spécifiques à votre organisation à l'aide d'expressions régulières. Cette fonctionnalité est actuellement en version préliminaire.

Pour plus d'informations, consultez la section "Ajouter des classifications personnalisées" .

Onglet Recherches enregistrées

L'onglet **Politiques** a été renommé**"Recherches enregistrées"** . La fonctionnalité reste inchangée.

Envoyer les événements d'analyse à la page Audit

La classification des données prend en charge l'envoi d'événements de classification (lorsqu'une analyse est lancée et lorsqu'elle se termine) au"Page d'audit du conseil NetApp" .

Mises à jour de sécurité

- Le package Keras a été mis à jour, atténuant les vulnérabilités (BDSA-2025-0107 et BDSA-2025-1984).
- La configuration des conteneurs Docker a été mise à jour. Le conteneur n'a plus accès aux interfaces réseau de l'hôte pour créer des paquets réseau bruts. En réduisant les accès inutiles, la mise à jour atténue les risques potentiels de sécurité.

Améliorations des performances

Des améliorations de code ont été implémentées pour réduire l'utilisation de la RAM et améliorer les performances globales de la classification des données.

Corrections de bugs

Les bugs qui entraînaient l'échec des analyses StorageGRID , le non-chargement des options de filtrage de la page d'investigation et le non-téléchargement de l'évaluation de découverte de données pour les évaluations à volume élevé ont été corrigés.

14 avril 2025

Version 1.42

Cette version de BlueXP classification comprend :

Analyse en masse pour les environnements de travail

La BlueXP classification prend en charge les opérations en masse pour les environnements de travail. Vous pouvez choisir d'activer les analyses de mappage, d'activer les analyses de mappage et de classification, de

désactiver les analyses ou de créer une configuration personnalisée sur les volumes dans l'environnement de travail. Si vous effectuez une sélection pour un volume individuel, elle remplace la sélection en bloc. Pour effectuer une opération en masse, accédez à la page **Configuration** et faites votre sélection.

Télécharger le rapport d'enquête localement

La BlueXP classification prend en charge la possibilité de télécharger des rapports d'enquête sur les données localement pour les afficher dans le navigateur. Si vous choisissez l'option locale, l'enquête sur les données n'est disponible qu'au format CSV et n'affiche que les 10 000 premières lignes de données.

Pour plus d'informations, consultez la section "Examinez les données stockées dans votre organisation avec la BlueXP classification" .

10 mars 2025

Version 1.41

Cette version de BlueXP classification inclut des améliorations générales et des corrections de bugs. Il comprend également :

État de l'analyse

La BlueXP classification suit la progression en temps réel des analyses de mappage et de classification *initiales* sur un volume. Des barres progressives distinctes suivent les analyses de cartographie et de classification, présentant un pourcentage du total des fichiers analysés. Vous pouvez également survoler une barre de progression pour afficher le nombre de fichiers analysés et le nombre total de fichiers. Le suivi de l'état de vos analyses crée des informations plus approfondies sur la progression de l'analyse, vous permettant de mieux planifier vos analyses et de comprendre l'allocation des ressources.

Pour afficher l'état de vos analyses, accédez à **Configuration** dans la BlueXP classification puis sélectionnez la **Configuration de l'environnement de travail**. La progression est affichée en ligne pour chaque volume.

19 février 2025

Version 1.40

Cette version de BlueXP classification inclut les mises à jour suivantes.

Prise en charge de RHEL 9.5

Cette version prend en charge Red Hat Enterprise Linux v9.5 en plus des versions précédemment prises en charge. Ceci s'applique à toute installation manuelle sur site de la BlueXP classification, y compris les déploiements de sites sombres.

Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version de BlueXP classification 1.30 ou supérieure : Red Hat Enterprise Linux versions 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4 et 9.5.

Donner la priorité aux analyses de cartographie uniquement

Lorsque vous effectuez des analyses de cartographie uniquement, vous pouvez prioriser les analyses les plus importantes. Cette fonctionnalité est utile lorsque vous disposez de nombreux environnements de travail et que vous souhaitez vous assurer que les analyses hautement prioritaires sont effectuées en premier.

Par défaut, les analyses sont mises en file d'attente en fonction de l'ordre dans lequel elles sont lancées. Grâce à la possibilité de hiérarchiser les analyses, vous pouvez déplacer les analyses vers l'avant de la file d'attente. Plusieurs analyses peuvent être priorisées. La priorité est désignée selon un ordre premier entré,

premier sorti, ce qui signifie que la première analyse que vous priorisez passe en tête de la file d'attente ; la deuxième analyse que vous priorisez devient la deuxième dans la file d'attente, et ainsi de suite.

La priorité est accordée une seule fois. Les réanalyses automatiques des données de cartographie se produisent dans l'ordre par défaut.

La priorisation est limitée à"analyses de cartographie uniquement" ; il n'est pas disponible pour les analyses de cartographie et de classification.

Pour plus d'informations, consultez la section "Prioriser les analyses" .

Réessayer toutes les analyses

La BlueXP classification prend en charge la possibilité de réessayer par lots toutes les analyses ayant échoué.

Vous pouvez réessayer les analyses dans une opération par lots avec la fonction **Réessayer tout**. Si les analyses de classification échouent en raison d'un problème temporaire tel qu'une panne de réseau, vous pouvez réessayer toutes les analyses en même temps avec un seul bouton au lieu de les réessayer individuellement. Les analyses peuvent être relancées autant de fois que nécessaire.

Pour réessayer toutes les analyses :

1. Dans le menu de BlueXP classification , sélectionnez **Configuration**.
2. Pour réessayer toutes les analyses ayant échoué, sélectionnez **Réessayer toutes les analyses**.

Amélioration de la précision du modèle de catégorisation

La précision du modèle d'apprentissage automatique pour"catégories prédéfinies" s'est améliorée de 11%.

22 janvier 2025

Version 1.39

Cette version de BlueXP classification met à jour le processus d'exportation du rapport d'enquête sur les données. Cette mise à jour d'exportation est utile pour effectuer des analyses supplémentaires sur vos données, créer des visualisations supplémentaires sur les données ou partager les résultats de votre enquête sur les données avec d'autres.

Auparavant, l'exportation du rapport d'enquête sur les données était limitée à 10 000 lignes. Avec cette version, la limite a été supprimée afin que vous puissiez exporter toutes vos données. Cette modification vous permet d'exporter davantage de données à partir de vos rapports d'investigation de données, vous offrant ainsi plus de flexibilité dans votre analyse de données.

Vous pouvez choisir l'environnement de travail, les volumes, le dossier de destination et le format JSON ou CSV. Le nom du fichier exporté inclut un horodatage pour vous aider à identifier quand les données ont été exportées.

Les environnements de travail pris en charge incluent :

- Cloud Volumes ONTAP
- FSx pour ONTAP
- ONTAP
- Groupe de partage

L'exportation des données du rapport d'enquête sur les données présente les limitations suivantes :

- Le nombre maximal d'enregistrements à télécharger est de 500 millions. par type (fichiers, répertoires et tables)
- Il est prévu qu'un million d'enregistrements soient exportés en environ 35 minutes.

Pour plus de détails sur l'enquête sur les données et le rapport, voir "Enquêter sur les données stockées dans votre organisation" .

16 décembre 2024

Version 1.38

Cette version de BlueXP classification inclut des améliorations générales et des corrections de bugs.

4 novembre 2024

Version 1.37

Cette version de BlueXP classification inclut les mises à jour suivantes.

Prise en charge de RHEL 8.10

Cette version prend en charge Red Hat Enterprise Linux v8.10 en plus des versions précédemment prises en charge. Ceci s'applique à toute installation manuelle sur site de la BlueXP classification, y compris les déploiements de sites sombres.

Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version de BlueXP classification 1.30 ou supérieure : Red Hat Enterprise Linux versions 8.8, 8.10, 9.0, 9.1, 9.2, 9.3 et 9.4.

En savoir plus sur "BlueXP classification" .

Prise en charge de NFS v4.1

Cette version prend en charge NFS v4.1 en plus des versions précédemment prises en charge.

En savoir plus sur "BlueXP classification" .

10 octobre 2024

Version 1.36

Prise en charge de RHEL 9.4

Cette version prend en charge Red Hat Enterprise Linux v9.4 en plus des versions précédemment prises en charge. Ceci s'applique à toute installation manuelle sur site de la BlueXP classification, y compris les déploiements de sites sombres.

Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version de BlueXP classification 1.30 ou supérieure : Red Hat Enterprise Linux versions 8.8, 9.0, 9.1, 9.2, 9.3 et 9.4.

En savoir plus sur "Présentation des déploiements de BlueXP classification" .

Amélioration des performances d'analyse

Cette version offre des performances d'analyse améliorées.

2 septembre 2024

Version 1.35

Analyser les données StorageGRID

La BlueXP classification prend en charge l'analyse des données dans StorageGRID.

Pour plus de détails, reportez-vous à"Analyser les données StorageGRID" .

05 août 2024

Version 1.34

Cette version de BlueXP classification inclut la mise à jour suivante.

Passer de CentOS à Ubuntu

La BlueXP classification a mis à jour son système d'exploitation Linux pour Microsoft Azure et Google Cloud Platform (GCP) de CentOS 7.9 à Ubuntu 22.04.

Pour plus de détails sur le déploiement, reportez-vous à "Installer sur un hôte Linux avec accès Internet et préparer le système hôte Linux" .

01 juillet 2024

Version 1.33

Ubuntu pris en charge

Cette version prend en charge la plate-forme Linux Ubuntu 24.04.

Les analyses cartographiques collectent des métadonnées

Les métadonnées suivantes sont extraites des fichiers lors des analyses de cartographie et sont affichées dans les tableaux de bord de gouvernance, de conformité et d'enquête :

- Environnement de travail
- Type d'environnement de travail
- Référentiel de stockage
- Type de fichier
- Capacité utilisée
- Nombre de fichiers
- Taille du fichier
- Création de fichier
- Dernier accès au fichier
- Fichier modifié pour la dernière fois
- Heure de découverte du fichier
- Extraction des autorisations

Données supplémentaires dans les tableaux de bord

Cette version met à jour les données qui apparaissent dans les tableaux de bord de gouvernance, de

conformité et d'enquête lors des analyses de mappage.

Pour plus de détails, consultez la section "Quelle est la différence entre les analyses de cartographie et de classification" .

05 juin 2024

Version 1.32

Nouvelle colonne d'état de mappage dans la page de configuration

Cette version affiche désormais une nouvelle colonne d'état de mappage dans la page de configuration. La nouvelle colonne vous aide à identifier si le mappage est en cours d'exécution, en file d'attente, en pause ou plus.

Pour des explications sur les statuts, voir "Modifier les paramètres de numérisation" .

15 mai 2024

Version 1.31

La classification est disponible en tant que service principal dans BlueXP

La BlueXP classification est désormais disponible en tant que fonctionnalité principale de BlueXP sans frais supplémentaires pour un maximum de 500 Tio de données numérisées par connecteur. Aucune licence de classification ni abonnement payant n'est requis. Comme nous concentrons la fonctionnalité de BlueXP classification sur l'analyse des systèmes de stockage NetApp avec cette nouvelle version, certaines fonctionnalités héritées ne seront disponibles que pour les clients qui avaient précédemment payé pour une licence. L'utilisation de ces fonctionnalités héritées expirera lorsque le contrat payant atteindra sa date de fin.

> La classification des données n'impose pas de limite à la quantité de données qu'elle peut analyser. Chaque agent de console prend en charge l'analyse et l'affichage de 500 Tio de données. Pour scanner plus de 500 Tio de données,"installer un autre agent de console" alors"déployer une autre instance de classification des données" . + L'interface utilisateur de la console affiche les données d'un seul connecteur. Pour obtenir des conseils sur l'affichage des données de plusieurs agents de console, consultez"Travailler avec plusieurs agents de console" .

1er avril 2024

Version 1.30

Prise en charge ajoutée pour la BlueXP classification

Cette version prend en charge Red Hat Enterprise Linux v8.8 et v9.3 en plus de la version 9.x précédemment prise en charge, qui nécessite Podman plutôt que le moteur Docker. Ceci s'applique à toute installation manuelle sur site de la BlueXP classification.

Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version de BlueXP classification 1.30 ou supérieure : Red Hat Enterprise Linux versions 8.8, 9.0, 9.1, 9.2 et 9.3.

En savoir plus sur "Présentation des déploiements de BlueXP classification" .

La BlueXP classification est prise en charge si vous installez le connecteur sur un hôte RHEL 8 ou 9 résidant sur site. Cette option n'est pas prise en charge si l'hôte RHEL 8 ou 9 réside dans AWS, Azure ou Google

Cloud.

Option permettant d'activer la collecte des journaux d'audit supprimée

L'option permettant d'activer la collecte des journaux d'audit a été désactivée.

Vitesse de numérisation améliorée

Les performances d'analyse sur les nœuds de scanner secondaires ont été améliorées. Vous pouvez ajouter davantage de nœuds de scanner si vous avez besoin d'une puissance de traitement supplémentaire pour vos numérisations. Pour plus de détails, reportez-vous à "Installer la BlueXP classification sur un hôte disposant d'un accès Internet" .

Mises à niveau automatiques

Si vous avez déployé la BlueXP classification sur un système avec accès Internet, le système est mis à niveau automatiquement. Auparavant, la mise à niveau se produisait après un certain temps écoulé depuis la dernière activité de l'utilisateur. Avec cette version, la BlueXP classification est mise à niveau automatiquement si l'heure locale est comprise entre 1h00 et 5h00 du matin. Si l'heure locale est en dehors de ces heures, la mise à niveau se produit après un délai spécifique écoulé depuis la dernière activité de l'utilisateur. Pour plus de détails, reportez-vous à "Installer sur un hôte Linux avec accès Internet" .

Si vous avez déployé la BlueXP classification sans accès Internet, vous devrez effectuer la mise à niveau manuellement. Pour plus de détails, reportez-vous à "Installer la BlueXP classification sur un hôte Linux sans accès Internet" .

04 mars 2024

Version 1.29

Vous pouvez désormais exclure les données d'analyse qui résident dans certains répertoires de sources de données

Si vous souhaitez que la BlueXP classification exclue les données d'analyse qui résident dans certains répertoires de sources de données, vous pouvez ajouter ces noms de répertoire à un fichier de configuration traité par la BlueXP classification . Cette fonctionnalité vous permet d'éviter d'analyser des répertoires inutiles ou qui pourraient renvoyer des résultats de données personnelles faussement positifs.

"Apprendre encore plus" .

La prise en charge des instances Extra Large est désormais qualifiée

Si vous avez besoin BlueXP classification pour analyser plus de 250 millions de fichiers, vous pouvez utiliser une instance Extra Large dans votre déploiement cloud ou votre installation sur site. Ce type de système peut analyser jusqu'à 500 millions de fichiers.

"Apprendre encore plus" .

10 janvier 2024

Version 1.27

Les résultats de la page d'enquête affichent la taille totale en plus du nombre total d'éléments

Les résultats filtrés dans la page Enquête affichent la taille totale des éléments en plus du nombre total de fichiers. Cela peut être utile lors du déplacement de fichiers, de la suppression de fichiers, etc.

Configurer des identifiants de groupe supplémentaires comme « Ouvrir à l'organisation »

Vous pouvez désormais configurer les ID de groupe dans NFS pour qu'ils soient considérés comme « Ouverts

à l'organisation » directement à partir de la BlueXP classification si le groupe n'avait pas été initialement défini avec cette autorisation. Tous les fichiers et dossiers auxquels ces identifiants de groupe sont associés s'afficheront comme « Ouvert à l'organisation » dans la page Détails de l'enquête. Découvrez comment"ajouter des identifiants de groupe supplémentaires comme « ouverts à l'organisation »" .

14 décembre 2023

Version 1.26.6

Cette version comprend quelques améliorations mineures.

La version a également supprimé les options suivantes :

- L'option permettant d'activer la collecte des journaux d'audit a été désactivée.
- Lors de l'enquête sur les annuaires, l'option permettant de calculer le nombre de données d'informations personnelles identifiables (PII) par annuaires n'est pas disponible. "Examinez les données stockées dans votre organisation" .
- L'option permettant d'intégrer des données à l'aide des étiquettes Azure Information Protection (AIP) a été désactivée.

06 novembre 2023

Version 1.26.3

Les problèmes suivants ont été résolus dans cette version

- Correction d'une incohérence lors de la présentation du nombre de fichiers scannés par le système dans les tableaux de bord.
- Amélioration du comportement d'analyse en gérant et en signalant les fichiers et répertoires avec des caractères spéciaux dans le nom et les métadonnées.

4 octobre 2023

Version 1.26

Prise en charge des installations sur site de la BlueXP classification sur RHEL version 9

Les versions 8 et 9 de Red Hat Enterprise Linux ne prennent pas en charge le moteur Docker, qui était requis pour l'installation de la BlueXP classification . Nous prenons désormais en charge l'installation de la BlueXP classification sur RHEL 9.0, 9.1 et 9.2 en utilisant Podman version 4 ou supérieure comme infrastructure de conteneur. Si votre environnement nécessite l'utilisation des versions les plus récentes de RHEL, vous pouvez désormais installer la BlueXP classification (version 1.26 ou supérieure) lorsque vous utilisez Podman.

À l'heure actuelle, nous ne prenons pas en charge les installations de sites sombres ou les environnements d'analyse distribués (utilisant un nœud de scanner maître et distant) lors de l'utilisation de RHEL 9.x.

05 septembre 2023

Version 1.25

Les déploiements de petite et moyenne taille sont temporairement indisponibles

Lorsque vous déployez une instance de BlueXP classification dans AWS, l'option permettant de sélectionner **Déployer > Configuration** et de choisir une instance de petite ou moyenne taille n'est pas disponible pour le

moment. Vous pouvez toujours déployer l'instance en utilisant la grande taille d'instance en sélectionnant **Déployer > Déployer**.

Appliquez des balises sur un maximum de 100 000 éléments à partir de la page Résultats de l'enquête

Auparavant, vous ne pouviez appliquer des balises qu'à une seule page à la fois dans la page Résultats de l'enquête (20 éléments). Vous pouvez désormais sélectionner **tous** les éléments dans les pages de résultats d'enquête et appliquer des balises à tous les éléments, jusqu'à 100 000 éléments à la fois.

Identifier les fichiers dupliqués avec une taille de fichier minimale de 1 Mo

La BlueXP classification était utilisée pour identifier les fichiers dupliqués uniquement lorsque les fichiers faisaient 50 Mo ou plus. Les fichiers dupliqués commençant par 1 Mo peuvent désormais être identifiés. Vous pouvez utiliser les filtres de la page Investigation « Taille du fichier » ainsi que « Doublons » pour voir quels fichiers d'une certaine taille sont dupliqués dans votre environnement.

17 juillet 2023

Version 1.24

Deux nouveaux types de données personnelles allemandes sont identifiés par la BlueXP classification

La BlueXP classification peut identifier et catégoriser les fichiers contenant les types de données suivants :

- Carte d'identité allemande (Personalausweisnummer)
- Numéro de sécurité sociale allemand (Sozialversicherungsnummer)

"Découvrez tous les types de données personnelles que la BlueXP classification peut identifier dans vos données" .

La BlueXP classification est entièrement prise en charge en mode restreint et en mode privé

La BlueXP classification est désormais entièrement prise en charge sur les sites sans accès Internet (mode privé) et avec un accès Internet sortant limité (mode restreint). "En savoir plus sur les modes de déploiement BlueXP pour le connecteur" .

Possibilité d'ignorer les versions lors de la mise à niveau d'une installation en mode privé de la BlueXP classification

Vous pouvez désormais mettre à niveau vers une version plus récente de la BlueXP classification même si elle n'est pas séquentielle. Cela signifie que la limitation actuelle de la mise à niveau de la BlueXP classification d'une version à la fois n'est plus nécessaire. Cette fonctionnalité est pertinente à partir de la version 1.24.

L'API de BlueXP classification est désormais disponible

L'API de BlueXP classification vous permet d'effectuer des actions, de créer des requêtes et d'exporter des informations sur les données que vous analysez. La documentation interactive est disponible via Swagger. La documentation est divisée en plusieurs catégories, notamment Enquête, Conformité, Gouvernance et Configuration. Chaque catégorie est une référence aux onglets de l'interface utilisateur de BlueXP classification .

"En savoir plus sur les API de BlueXP classification" .

06 juin 2023

Version 1.23

Le japonais est désormais pris en charge lors de la recherche de noms de personnes concernées

Les noms japonais peuvent désormais être saisis lors de la recherche du nom d'un sujet en réponse à une demande d'accès aux données personnelles (DSAR). Vous pouvez générer un"Rapport de demande d'accès aux données personnelles" avec les informations qui en résultent. Vous pouvez également saisir des noms japonais dans le champ"Filtre « Personne concernée » dans la page Enquête sur les données" pour identifier les fichiers qui contiennent le nom du sujet.

Ubuntu est désormais une distribution Linux prise en charge sur laquelle vous pouvez installer la BlueXP classification

Ubuntu 22.04 a été qualifié comme système d'exploitation pris en charge pour la BlueXP classification. Vous pouvez installer la BlueXP classification sur un hôte Ubuntu Linux de votre réseau ou sur un hôte Linux dans le cloud lorsque vous utilisez la version 1.23 du programme d'installation. "Découvrez comment installer la BlueXP classification sur un hôte avec Ubuntu installé" .

Red Hat Enterprise Linux 8.6 et 8.7 ne sont plus pris en charge avec les nouvelles installations de BlueXP classification

Ces versions ne sont pas prises en charge avec les nouveaux déploiements car Red Hat ne prend plus en charge Docker, ce qui est une condition préalable. Si vous disposez d'une machine de BlueXP classification existante exécutée sur RHEL 8.6 ou 8.7, NetApp continuera à prendre en charge votre configuration.

La BlueXP classification peut être configurée comme un collecteur FPolicy pour recevoir les événements FPolicy des systèmes ONTAP

Vous pouvez activer la collecte des journaux d'audit d'accès aux fichiers sur votre système de BlueXP classification pour les événements d'accès aux fichiers détectés sur les volumes de vos environnements de travail. La BlueXP classification peut capturer les types d'événements FPolicy suivants et les utilisateurs qui ont effectué les actions sur vos fichiers : Créer, Lire, Écrire, Supprimer, Renommer, Modifier le propriétaire/les autorisations et Modifier la SACL/DACL.

Les licences BYOL Data Sense sont désormais prises en charge sur les sites sombres

Vous pouvez désormais télécharger votre licence Data Sense BYOL dans le BlueXP digital wallet sur un site sombre afin d'être averti lorsque votre licence devient faible.

03 avril 2023

Version 1.22

Nouveau rapport d'évaluation de la découverte de données

Le rapport d'évaluation de la découverte de données fournit une analyse de haut niveau de votre environnement analysé pour mettre en évidence les résultats du système et montrer les zones de préoccupation et les étapes de correction potentielles. L'objectif de ce rapport est de sensibiliser aux problèmes de gouvernance des données, aux expositions en matière de sécurité des données et aux lacunes en matière de conformité des données de votre ensemble de données. "Découvrez comment générer et utiliser le rapport d'évaluation de la découverte de données" .

Possibilité de déployer la BlueXP classification sur des instances plus petites dans le cloud

Lors du déploiement de la BlueXP classification à partir d'un connecteur BlueXP dans un environnement AWS, vous pouvez désormais choisir entre deux types d'instances plus petits que ceux disponibles avec l'instance par défaut. Si vous numérisez un petit environnement, cela peut vous aider à économiser sur les coûts du cloud. Cependant, il existe certaines restrictions lors de l'utilisation de l'instance plus petite. "Voir les types d'instances disponibles et les limitations" .

Un script autonome est désormais disponible pour qualifier votre système Linux avant l'installation de la BlueXP classification

Si vous souhaitez vérifier que votre système Linux répond à toutes les conditions préalables indépendamment

de l'exécution de l'installation de la BlueXP classification , vous pouvez télécharger un script distinct qui teste uniquement les conditions préalables. "Découvrez comment vérifier si votre hôte Linux est prêt à installer la BlueXP classification" .

07 mars 2023

Version 1.21

Nouvelle fonctionnalité pour ajouter vos propres catégories personnalisées à partir de l'interface utilisateur de BlueXP classification

La BlueXP classification vous permet désormais d'ajouter vos propres catégories personnalisées afin que la BlueXP classification identifie les fichiers qui correspondent à ces catégories. La BlueXP classification comporte de nombreuses "catégories prédéfinies" , cette fonctionnalité vous permet donc d'ajouter des catégories personnalisées pour identifier où se trouvent les informations propres à votre organisation dans vos données.

Vous pouvez désormais ajouter des mots-clés personnalisés à partir de l'interface utilisateur de BlueXP classification

La BlueXP classification a la possibilité d'ajouter des mots-clés personnalisés que la BlueXP classification identifiera dans les analyses futures pendant un certain temps. Cependant, vous devez vous connecter à l'hôte Linux de BlueXP classification et utiliser une interface de ligne de commande pour ajouter les mots-clés. Dans cette version, la possibilité d'ajouter des mots-clés personnalisés est disponible dans l'interface utilisateur de BlueXP classification , ce qui facilite grandement l'ajout et la modification de ces mots-clés.

Possibilité de faire en sorte que la BlueXP classification ne scanne pas les fichiers lorsque l'« heure du dernier accès » est modifiée

Par défaut, si la BlueXP classification ne dispose pas des autorisations « d'écriture » adéquates, le système n'analysera pas les fichiers de vos volumes, car la BlueXP classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Cependant, si vous ne vous souciez pas de savoir si l'heure du dernier accès est réinitialisée à l'heure d'origine dans vos fichiers, vous pouvez remplacer ce comportement dans la page de configuration afin que la BlueXP classification analyse les volumes quelles que soient les autorisations.

En conjonction avec cette fonctionnalité, un nouveau filtre nommé « Événement d'analyse d'analyse » a été ajouté afin que vous puissiez afficher les fichiers qui n'ont pas été classés parce que la BlueXP classification n'a pas pu revenir à l'heure du dernier accès, ou les fichiers qui ont été classés même si la BlueXP classification n'a pas pu revenir à l'heure du dernier accès.

"En savoir plus sur l'horodatage du dernier accès et les autorisations requises par la BlueXP classification" .

Trois nouveaux types de données personnelles sont identifiés par la BlueXP classification

La BlueXP classification peut identifier et catégoriser les fichiers contenant les types de données suivants :

- Numéro de carte d'identité du Botswana (Omang)
- Numéro de passeport du Botswana
- Carte d'identité nationale de Singapour (NRIC)

"Découvrez tous les types de données personnelles que la BlueXP classification peut identifier dans vos données" .

Fonctionnalités mises à jour pour les répertoires

- L'option « Rapport CSV léger » pour les rapports d'investigation de données inclut désormais des informations provenant d'annuaires.

- Le filtre horaire « Dernier accès » affiche désormais l'heure du dernier accès pour les fichiers et les répertoires.

Améliorations de l'installation

- L'installateur de BlueXP classification pour les sites sans accès Internet (sites sombres) effectue désormais une pré-vérification pour s'assurer que les exigences de votre système et de votre réseau sont en place pour une installation réussie.
- Les fichiers journaux d'audit d'installation sont désormais enregistrés ; ils sont écrits dans `/ops/netapp/install_logs` .

05 février 2023

Version 1.20

Possibilité d'envoyer des e-mails de notification basés sur des politiques à n'importe quelle adresse e-mail

Dans les versions antérieures de la BlueXP classification, vous pouviez envoyer des alertes par e-mail aux utilisateurs BlueXP de votre compte lorsque certaines politiques critiques renvoyaient des résultats. Cette fonctionnalité vous permet de recevoir des notifications pour protéger vos données lorsque vous n'êtes pas en ligne. Vous pouvez désormais également envoyer des alertes par e-mail à partir des politiques à tous les autres utilisateurs (jusqu'à 20 adresses e-mail) qui ne figurent pas dans votre compte BlueXP .

"En savoir plus sur l'envoi d'alertes par e-mail en fonction des résultats de la politique" .

Vous pouvez désormais ajouter des modèles personnels à partir de l'interface de BlueXP classification

La BlueXP classification a la possibilité d'ajouter des « données personnelles » personnalisées que la BlueXP classification identifiera dans les analyses futures pendant un certain temps. Cependant, vous devez vous connecter à l'hôte Linux de BlueXP classification et utiliser une ligne de commande pour ajouter les modèles personnalisés. Dans cette version, la possibilité d'ajouter des modèles personnels à l'aide d'une expression régulière est présente dans l'interface utilisateur de BlueXP classification , ce qui facilite grandement l'ajout et la modification de ces modèles personnalisés.

Possibilité de déplacer 15 millions de fichiers à l'aide de la BlueXP classification

Par le passé, la BlueXP classification pouvait déplacer un maximum de 100 000 fichiers sources vers n'importe quel partage NFS. Vous pouvez désormais déplacer jusqu'à 15 millions de fichiers à la fois.

Possibilité de voir le nombre d'utilisateurs ayant accès aux fichiers SharePoint Online

Le filtre « Nombre d'utilisateurs avec accès » prend désormais en charge les fichiers stockés dans les référentiels SharePoint Online. Auparavant, seuls les fichiers sur les partages CIFS étaient pris en charge. Notez que les groupes SharePoint qui ne sont pas basés sur Active Directory ne seront pas comptabilisés dans ce filtre pour le moment.

Un nouveau statut « Succès partiel » a été ajouté au panneau Statut de l'action

Le nouveau statut « Succès partiel » indique qu'une action de BlueXP classification est terminée et que certains éléments ont échoué et que d'autres ont réussi, par exemple lorsque vous déplacez ou supprimez 100 fichiers. De plus, le statut « Terminé » a été renommé « Succès ». Par le passé, le statut « Terminé » pouvait répertorier les actions qui avaient réussi et celles qui avaient échoué. Désormais, le statut « Succès » signifie que toutes les actions ont réussi sur tous les éléments. "Découvrez comment afficher le panneau d'état des actions" .

09 janvier 2023

Version 1.19

Possibilité de visualiser un tableau des fichiers contenant des données sensibles et trop permissifs

Le tableau de bord de gouvernance a ajouté une nouvelle zone *Données sensibles et autorisations étendues* qui fournit une carte thermique des fichiers contenant des données sensibles (y compris des données personnelles sensibles et sensibles) et qui sont trop permissifs. Cela peut vous aider à voir où vous pourriez avoir des risques avec des données sensibles. "Apprendre encore plus" .

Trois nouveaux filtres sont disponibles sur la page Enquête sur les données

De nouveaux filtres sont disponibles pour affiner les résultats qui s'affichent dans la page Investigation des données :

- Le filtre « Nombre d'utilisateurs avec accès » indique quels fichiers et dossiers sont ouverts à un certain nombre d'utilisateurs. Vous pouvez choisir une plage de nombres pour affiner les résultats, par exemple pour voir quels fichiers sont accessibles par 51 à 100 utilisateurs.
- Les filtres « Heure de création », « Heure de découverte », « Dernière modification » et « Dernier accès » vous permettent désormais de créer une plage de dates personnalisée au lieu de simplement sélectionner une plage de jours prédéfinie. Par exemple, vous pouvez rechercher des fichiers dont la « Heure de création » est « ancienne » ou dont la date de « Dernière modification » est comprise dans les « 10 derniers jours ».
- Le filtre « Chemin de fichier » vous permet désormais de spécifier les chemins que vous souhaitez exclure des résultats de requête filtrés. Si vous entrez des chemins pour inclure et exclure certaines données, la BlueXP classification recherche d'abord tous les fichiers dans les chemins inclus, puis supprime les fichiers des chemins exclus, puis affiche les résultats.

"Consultez la liste de tous les filtres que vous pouvez utiliser pour analyser vos données" .

La BlueXP classification peut identifier le numéro individuel japonais

La BlueXP classification peut identifier et catégoriser les fichiers contenant le numéro individuel japonais (également connu sous le nom de Mon numéro). Cela inclut à la fois le numéro personnel et le numéro d'entreprise My Number. "Découvrez tous les types de données personnelles que la BlueXP classification peut identifier dans vos données" .

Limitations connues dans la NetApp Data Classification

Les limitations connues identifient les fonctions qui ne sont pas prises en charge ou qui n'interagissent pas correctement dans cette version. Examinez attentivement ces limitations.

Options désactivées de la NetApp Data Classification

La version de décembre 2023 (version 1.26.6) a supprimé les options suivantes :

- L'option permettant d'activer la collecte des journaux d'audit a été désactivée.
- Lors de l'enquête sur les annuaires, l'option permettant de calculer le nombre de données d'informations personnelles identifiables (PII) par annuaires n'est pas disponible.
- L'option permettant d'intégrer des données à l'aide des étiquettes Azure Information Protection (AIP) a été désactivée.

Analyse de la classification des données

Les limitations suivantes se produisent avec les analyses de classification des données.

La classification des données analyse un seul partage sous un volume

Si vous disposez de plusieurs partages de fichiers sous un seul volume, la classification des données analyse le partage avec la hiérarchie la plus élevée. Par exemple, si vous avez des actions comme les suivantes :

- /UN
- /A/B
- /C
- /D/E

Dans cette configuration, seules les données dans /A sont analysées. Les données dans /C et /D ne sont pas analysées.

Solution de contournement

Il existe une solution de contournement pour vous assurer que vous analysez les données de tous les partages de votre volume. Suivez ces étapes :

1. Dans le système, ajoutez le volume à analyser.
2. Une fois que la classification des données a terminé l'analyse du volume, accédez à la page *Investigation des données* et créez un filtre pour voir quel partage est analysé :

   Filtrez les données par « Nom du système » et « Type de répertoire = Partage » pour voir quel partage est analysé.

3. Obtenez la liste complète des partages qui existent dans le volume afin de pouvoir voir quels partages ne sont pas analysés.
4. "Ajouter les actions restantes à un groupe de partage" .

   Ajoutez toutes les actions individuellement, par exemple :

   ```
   /C
   /D
   ```

5. Effectuez ces étapes pour chaque volume du système qui possède plusieurs partages.

Horodatage du dernier accès

Lorsque Data Classification effectue une analyse d'un répertoire, l'analyse affecte le champ **Dernier accès** du répertoire. Lorsque vous affichez le champ **Dernier accès**, ces métadonnées reflètent soit la date et l'heure de l'analyse, soit la dernière fois qu'un utilisateur a accédé au répertoire.

Commencer

En savoir plus sur la NetApp Data Classification

NetApp Data Classification est un service de gouvernance des données pour la NetApp Console qui analyse vos sources de données d'entreprise sur site et dans le cloud pour mapper et classer les données et identifier les informations privées. Cela peut vous aider à réduire vos risques de sécurité et de conformité, à diminuer vos coûts de stockage et à vous aider dans vos projets de migration de données.

À partir de la version 1.31, la classification des données est disponible en tant que fonctionnalité principale dans la NetApp Console. Il n'y a pas de frais supplémentaires. Aucune licence de classification ni abonnement n'est requis. + Si vous avez utilisé la version héritée 1.30 ou une version antérieure, cette version est disponible jusqu'à l'expiration de votre abonnement.

NetApp Console

La classification des données est accessible via la NetApp Console.

La NetApp Console fournit une gestion centralisée des services de stockage et de données NetApp dans les environnements sur site et cloud à l'échelle de l'entreprise. La console est requise pour accéder aux services de données NetApp et les utiliser. En tant qu'interface de gestion, il vous permet de gérer de nombreuses ressources de stockage à partir d'une seule interface. Les administrateurs de console peuvent contrôler l'accès au stockage et aux services pour tous les systèmes de l'entreprise.

Vous n'avez pas besoin de licence ni d'abonnement pour commencer à utiliser NetApp Console et vous n'encourez des frais que lorsque vous devez déployer des agents de console dans votre cloud pour garantir la connectivité à vos systèmes de stockage ou à vos services de données NetApp . Cependant, certains services de données NetApp accessibles depuis la console sont sous licence ou basés sur un abonnement.

En savoir plus sur le"NetApp Console" .

Caractéristiques

La classification des données utilise l'intelligence artificielle (IA), le traitement du langage naturel (NLP) et l'apprentissage automatique (ML) pour comprendre le contenu qu'elle analyse afin d'extraire des entités et de catégoriser le contenu en conséquence. Cela permet à la classification des données de fournir les domaines de fonctionnalités suivants.

"En savoir plus sur les cas d'utilisation de la classification des données" .

Maintenir la conformité

La classification des données fournit plusieurs outils qui peuvent vous aider dans vos efforts de conformité. Vous pouvez utiliser la classification des données pour :

- Identifier les informations personnelles identifiables (PII).
- Identifiez un large éventail d'informations personnelles sensibles comme l'exigent les réglementations de confidentialité GDPR, CCPA, PCI et HIPAA.
- Répondre aux demandes d'accès aux données des personnes concernées (DSAR) en fonction du nom ou de l'adresse e-mail.

Renforcer la sécurité

La classification des données permet d'identifier les données potentiellement susceptibles d'être consultées à des fins criminelles. Vous pouvez utiliser la classification des données pour :

- Identifiez tous les fichiers et répertoires (partages et dossiers) avec des autorisations ouvertes qui sont exposés à l'ensemble de votre organisation ou au public.
- Identifiez les données sensibles qui résident en dehors de l'emplacement initial dédié.
- Respecter les politiques de conservation des données.
- Utilisez *Policies* pour détecter automatiquement les nouveaux problèmes de sécurité afin que le personnel de sécurité puisse agir immédiatement.

Optimiser l'utilisation du stockage

La classification des données fournit des outils qui peuvent vous aider à déterminer le coût total de possession (TCO) de votre stockage. Vous pouvez utiliser la classification des données pour :

- Augmentez l'efficacité du stockage en identifiant les données en double ou non liées à l'entreprise.
- Réduisez les coûts de stockage en identifiant les données inactives que vous pouvez hiérarchiser vers un stockage d'objets moins coûteux. "En savoir plus sur la hiérarchisation des systèmes Cloud Volumes ONTAP" . "En savoir plus sur la hiérarchisation des systèmes ONTAP sur site" .

Systèmes et sources de données pris en charge

La classification des données peut scanner et analyser des données structurées et non structurées provenant des types de systèmes et de sources de données suivants :

Systèmes

- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- Cloud Volumes ONTAP (déployé dans AWS, Azure ou GCP)
- Clusters ONTAP sur site
- StorageGRID
- Google Cloud NetApp Volumes

Sources de données

- Partages de fichiers NetApp
- Bases de données:
  - Service de base de données relationnelle Amazon (Amazon RDS)
  - MongoDB
  - MySQL
  - Oracle
  - PostgreSQL
  - SAP HANA
  - Serveur SQL (MSSQL)

La classification des données prend en charge les versions NFS 3.x, 4.0 et 4.1, ainsi que les versions CIFS 1.x, 2.0, 2.1 et 3.0.

Coût

La classification des données est gratuite. Aucune licence de classification ni abonnement payant n'est requis.

Coûts d'infrastructure

- L'installation de Data Classification dans le cloud nécessite le déploiement d'une instance cloud, ce qui entraîne des frais de la part du fournisseur cloud où elle est déployée. Voir le type d'instance déployé pour chaque fournisseur de cloud . L'installation de Data Classification sur un système local est gratuite.
- La classification des données nécessite que vous ayez déployé un agent de console. Dans de nombreux cas, vous disposez déjà d'un agent de console en raison d'autres stockages et services que vous utilisez dans la console. L'instance de l'agent de console entraîne des frais auprès du fournisseur de cloud où elle est déployée. Voir le "type d'instance déployée pour chaque fournisseur de cloud" . L'installation de l'agent de console sur un système local est gratuite.

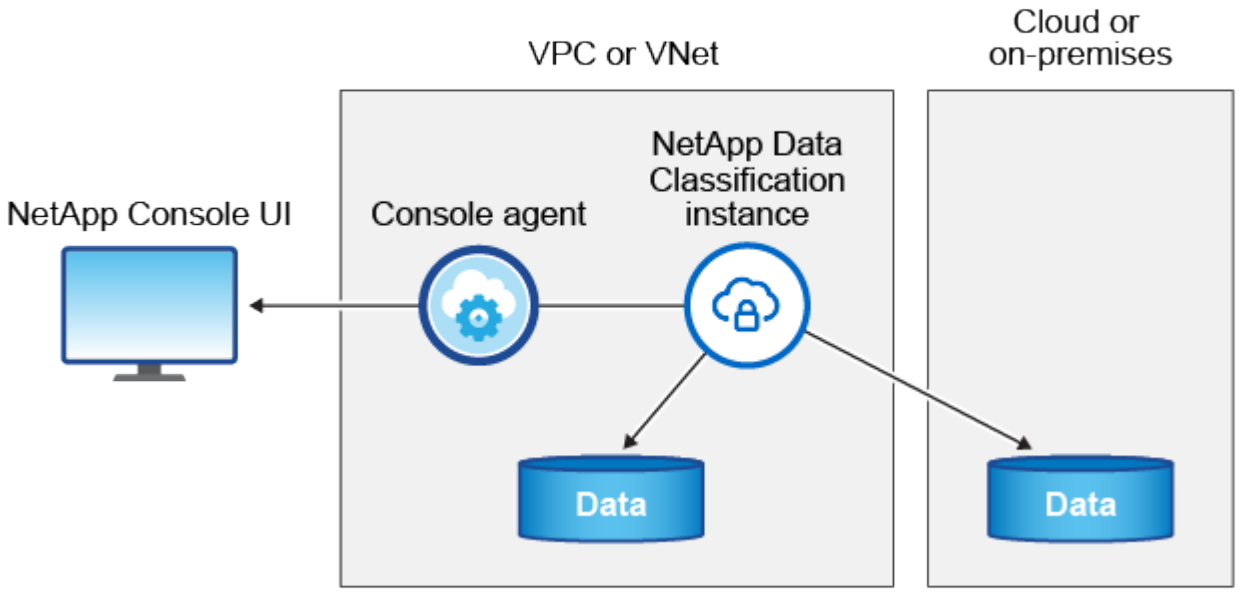
Coûts de transfert de données

Les coûts de transfert de données dépendent de votre configuration. Si l'instance de classification des données et la source de données se trouvent dans la même zone de disponibilité et la même région, il n'y a aucun coût de transfert de données. Mais si la source de données, comme un système Cloud Volumes ONTAP , se trouve dans une zone de disponibilité ou une région *différente*, les frais de transfert de données vous seront facturés par votre fournisseur de cloud. Consultez ces liens pour plus de détails :

- "AWS : Tarifs d'Amazon Elastic Compute Cloud (Amazon EC2)"
- "Microsoft Azure : Détails des tarifs de la bande passante"
- "Google Cloud : tarifs du service de transfert de stockage"

L'instance de classification des données

Lorsque vous déployez la classification des données dans le cloud, la console déploie l'instance dans le même sous-réseau que l'agent de la console. "En savoir plus sur l'agent de console."



Notez ce qui suit à propos de l'instance par défaut :

- Dans AWS, la classification des données s'exécute sur un "instance m6i.4xlarge" avec un disque GP2 de 500 Gio. L'image du système d'exploitation est Amazon Linux 2. Lorsqu'il est déployé dans AWS, vous pouvez choisir une taille d'instance plus petite si vous analysez une petite quantité de données.
- Dans Azure, la classification des données s'exécute sur un"VM Standard_D16s_v3" avec un disque de 500 Gio. L'image du système d'exploitation est Ubuntu 22.04.
- Dans GCP, la classification des données s'exécute sur un"n2-standard-16 VM" avec un disque persistant standard de 500 Gio. L'image du système d'exploitation est Ubuntu 22.04.
- Dans les régions où l'instance par défaut n'est pas disponible, la classification des données s'exécute sur une instance alternative. "Voir les types d'instances alternatifs" .
- L'instance est nommée *CloudCompliance* avec un hachage généré (UUID) concaténé. Par exemple : *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- Une seule instance de classification des données est déployée par agent de console.

Vous pouvez également déployer la classification des données sur un hôte Linux dans vos locaux ou sur un hôte chez votre fournisseur de cloud préféré. Le logiciel fonctionne exactement de la même manière, quelle que soit la méthode d'installation choisie. Les mises à niveau du logiciel de classification des données sont automatisées tant que l'instance dispose d'un accès Internet.

L'instance doit rester en cours d'exécution en permanence, car la classification des données analyse en permanence les données.

Déployer sur différents types d'instances

Consultez les spécifications suivantes pour les types d'instances :

Taille du système	Spécifications	Limites
Très grand	32 processeurs, 128 Go de RAM, 1 To de SSD	Peut numériser jusqu'à 500 millions de fichiers.
Grand (par défaut)	16 processeurs, 64 Go de RAM, 500 Go de SSD	Peut numériser jusqu'à 250 millions de fichiers.

Lors du déploiement de la classification des données dans Azure ou GCP, envoyez un e-mail à ng-contact-data-sense@netapp.com pour obtenir de l'aide si vous souhaitez utiliser un type d'instance plus petit.

Comment fonctionne l'analyse de classification des données

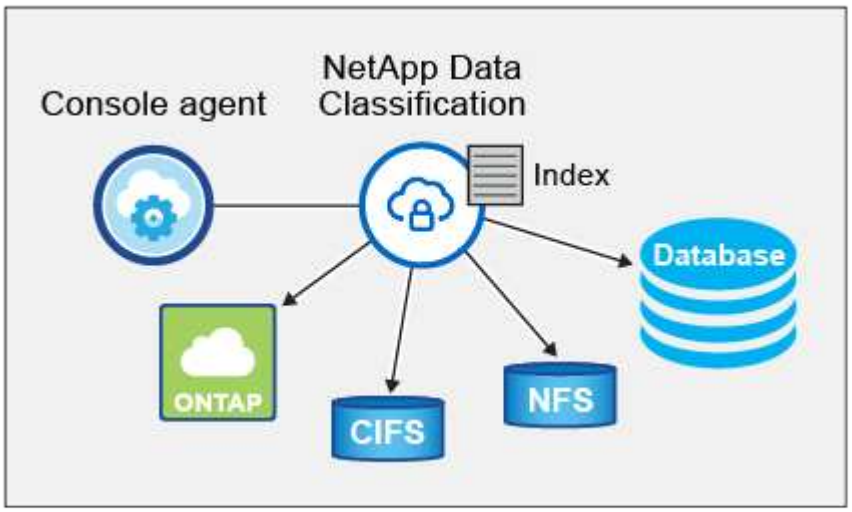
À un niveau élevé, l'analyse de classification des données fonctionne comme ceci :

1. Vous déployez une instance de classification des données dans la console.
2. Vous activez le mappage de haut niveau (appelé analyses *Mapping uniquement*) ou l'analyse de niveau profond (appelée analyses *Map & Classify*) sur une ou plusieurs sources de données.
3. La classification des données analyse les données à l'aide d'un processus d'apprentissage de l'IA.
4. Vous utilisez les tableaux de bord et les outils de reporting fournis pour vous aider dans vos efforts de conformité et de gouvernance.

Une fois que vous avez activé la classification des données et sélectionné les référentiels que vous souhaitez analyser (il s'agit des volumes, des schémas de base de données ou d'autres données utilisateur), l'analyse

des données commence immédiatement pour identifier les données personnelles et sensibles. Dans la plupart des cas, vous devez vous concentrer sur l'analyse des données de production en direct plutôt que sur les sauvegardes, les miroirs ou les sites de reprise après sinistre. Ensuite, la classification des données cartographie vos données organisationnelles, catégorise chaque fichier et identifie et extrait les entités et les modèles prédéfinis dans les données. Le résultat de l'analyse est un index des informations personnelles, des informations personnelles sensibles, des catégories de données et des types de fichiers.

La classification des données se connecte aux données comme n'importe quel autre client en montant des volumes NFS et CIFS. Les volumes NFS sont automatiquement accessibles en lecture seule, tandis que vous devez fournir les informations d'identification Active Directory pour analyser les volumes CIFS.



Après l'analyse initiale, Data Classification analyse en continu vos données de manière circulaire pour détecter les modifications incrémentielles. C'est pourquoi il est important de maintenir l'instance en cours d'exécution.

Vous pouvez activer et désactiver les analyses au niveau du volume ou au niveau du schéma de base de données.

> La classification des données n'impose pas de limite à la quantité de données qu'elle peut analyser. Chaque agent de console prend en charge l'analyse et l'affichage de 500 Tio de données. Pour scanner plus de 500 Tio de données,"installer un autre agent de console" alors"déployer une autre instance de classification des données" . + L'interface utilisateur de la console affiche les données d'un seul connecteur. Pour obtenir des conseils sur l'affichage des données de plusieurs agents de console, consultez"Travailler avec plusieurs agents de console" .

Quelle est la différence entre les analyses de cartographie et de classification

Vous pouvez effectuer deux types d'analyses dans la classification des données :

- **Les analyses de cartographie uniquement** fournissent uniquement un aperçu de haut niveau de vos données et sont effectuées sur des sources de données sélectionnées. Les analyses de cartographie uniquement prennent moins de temps que les analyses de cartographie et de classification, car elles n'accèdent pas aux fichiers pour voir les données qu'ils contiennent. Vous souhaiterez peut-être procéder ainsi dans un premier temps pour identifier les domaines de recherche, puis effectuer une analyse de cartographie et de classification sur ces domaines.
- **Les analyses de cartographie et de classification** fournissent une analyse approfondie de vos données.

Pour plus de détails sur les différences entre les analyses de cartographie et de classification, voir"Quelle est

la différence entre les analyses de cartographie et de classification ?" .

Informations catégorisées par la classification des données

La classification des données collecte, indexe et attribue des catégories aux données suivantes :

- **Métadonnées standard** sur les fichiers : le type de fichier, sa taille, ses dates de création et de modification, etc.
- **Données personnelles** : Informations personnelles identifiables (PII) telles que les adresses e-mail, les numéros d'identification ou les numéros de carte de crédit, que Data Classification identifie à l'aide de mots, de chaînes et de modèles spécifiques dans les fichiers. "En savoir plus sur les données personnelles" .
- **Données personnelles sensibles** : Types particuliers d'informations personnelles sensibles (IPS), telles que les données de santé, l'origine ethnique ou les opinions politiques, telles que définies par le Règlement général sur la protection des données (RGPD) et d'autres réglementations sur la confidentialité. "En savoir plus sur les données personnelles sensibles" .
- **Catégories** : La classification des données prend les données numérisées et les divise en différents types de catégories. Les catégories sont des sujets basés sur l'analyse par l'IA du contenu et des métadonnées de chaque fichier. "En savoir plus sur les catégories".
- **Reconnaissance d'entité de nom** : la classification des données utilise l'IA pour extraire les noms naturels des personnes à partir de documents. "En savoir plus sur la réponse aux demandes d'accès aux données des personnes concernées" .

Présentation du réseau

Data Classification déploie un serveur unique, ou cluster, où vous le souhaitez : dans le cloud ou sur site. Les serveurs se connectent via des protocoles standard aux sources de données et indexent les résultats dans un cluster Elasticsearch, qui est également déployé sur les mêmes serveurs. Cela permet la prise en charge des environnements multicloud, cross-cloud, cloud privé et sur site.

La console déploie l'instance de classification des données avec un groupe de sécurité qui active les connexions HTTP entrantes à partir de l'agent de la console.

Lorsque vous utilisez la console en mode SaaS, la connexion à la console est effectuée via HTTPS et les données privées envoyées entre votre navigateur et l'instance de classification des données sont sécurisées par un cryptage de bout en bout à l'aide de TLS 1.2, ce qui signifie que NetApp et des tiers ne peuvent pas les lire.

Les règles sortantes sont complètement ouvertes. Un accès Internet est nécessaire pour installer et mettre à niveau le logiciel de classification des données et pour envoyer des mesures d'utilisation.

Si vous avez des exigences réseau strictes,"en savoir plus sur les points de terminaison contactés par la classification des données" .

Accéder à la NetApp Data Classification

Vous pouvez accéder à la NetApp Data Classification via la NetApp Console.

Pour vous connecter à la console, vous pouvez utiliser vos informations d'identification du site de support NetApp ou vous inscrire pour une connexion à la NetApp Console à l'aide de votre e-mail et d'un mot de passe. "En savoir plus sur la connexion à la console" .

Des tâches spécifiques nécessitent des rôles d'utilisateur de console spécifiques. "En savoir plus sur les rôles d'accès à la console pour tous les services" .

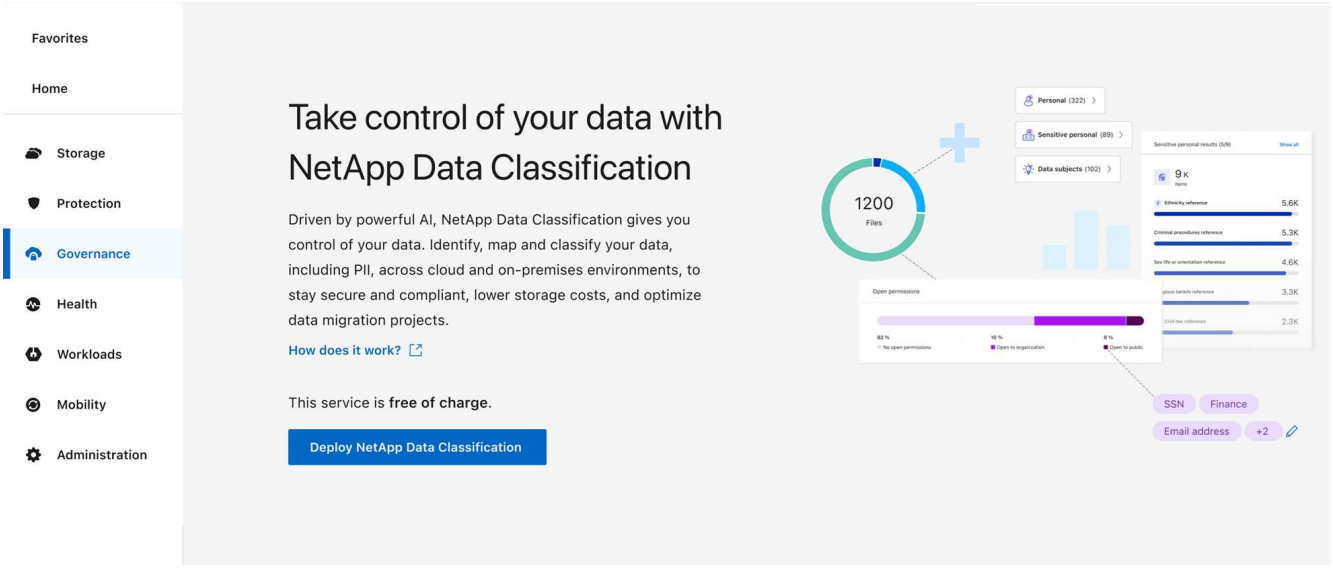
Avant de commencer

- "Vous devez ajouter un agent de console."
- "Comprendre quel style de déploiement de classification des données convient à votre charge de travail."

Étapes

1. Dans un navigateur Web, accédez à l'"Console" .
2. Connectez-vous à la console.
3. Depuis la page principale de la NetApp Console, sélectionnez **Gouvernance** > **Classification des données**.
4. Si c'est la première fois que vous accédez à la classification des données, la page de destination apparaît.

   Sélectionnez **Déployer la classification sur site ou dans le cloud** pour commencer à déployer votre instance de classification. Pour plus d'informations, voir"Quel déploiement de classification des données devez-vous utiliser ?"



   Sinon, le tableau de bord de classification des données s'affiche.

Déployer la classification des données

Quel déploiement de NetApp Data Classification devez-vous utiliser ?

Vous pouvez déployer NetApp Data Classification de différentes manières. Découvrez quelle méthode répond à vos besoins.

La classification des données peut être déployée des manières suivantes :

- "Déployer dans le cloud à l'aide de la console" . La console déploie l'instance de classification des données dans le même réseau de fournisseur de cloud que l'agent de la console.
- "Installer sur un hôte Linux avec accès Internet" . Installez Data Classification sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud, disposant d'un accès Internet. Ce type d'installation peut être

une bonne option si vous préférez analyser les systèmes ONTAP sur site à l'aide d'une instance de classification des données également située sur site, bien que cela ne soit pas une exigence.

- "Installer sur un hôte Linux dans un site sur site sans accès Internet", également connu sous le nom de *mode privé*. Ce type d'installation, qui utilise un script d'installation, n'a aucune connectivité à la couche SaaS de la console.

> Le mode privé BlueXP (interface BlueXP héritée) est généralement utilisé avec des environnements locaux qui n'ont pas de connexion Internet et avec des régions cloud sécurisées, notamment AWS Secret Cloud, AWS Top Secret Cloud et Azure IL6. NetApp continue de prendre en charge ces environnements avec l'interface BlueXP héritée. Pour la documentation du mode privé dans l'ancienne interface BlueXP , voir"Documentation PDF pour le mode privé BlueXP" .

L'installation sur un hôte Linux avec accès Internet et l'installation sur site sur un hôte Linux sans accès Internet utilisent un script d'installation. Le script commence par vérifier si le système et l'environnement répondent aux prérequis. Si les conditions préalables sont remplies, l'installation démarre. Si vous souhaitez vérifier les conditions préalables indépendamment de l'exécution de l'installation de la classification des données, vous pouvez télécharger un progiciel distinct qui teste uniquement les conditions préalables.

"Vérifiez que votre hôte Linux est prêt à installer la classification des données" .

Déployer la NetApp Data Classification dans le cloud à l'aide de la NetApp Console

Vous pouvez déployer NetApp Data Classification dans le cloud avec la NetApp Console. La console déploie l'instance de classification des données dans le même réseau de fournisseur de cloud que l'agent de la console.

Notez que vous pouvez également"installer Data Classification sur un hôte Linux disposant d'un accès Internet" . Ce type d'installation peut être une bonne option si vous préférez analyser les systèmes ONTAP sur site à l'aide d'une instance de classification des données également située sur site, mais ce n'est pas une exigence. Le logiciel fonctionne exactement de la même manière, quelle que soit la méthode d'installation choisie.

Démarrage rapide

Commencez rapidement en suivant ces étapes ou faites défiler les sections restantes pour obtenir tous les détails.

Créer un agent de console

Si vous n'avez pas encore d'agent de console, créez-en un. Voir "création d'un agent de console dans AWS" , "création d'un agent de console dans Azure" , ou "création d'un agent de console dans GCP" .

Vous pouvez également "installer l'agent de console sur site" sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud.

Prérequis

Assurez-vous que votre environnement peut répondre aux prérequis. Cela inclut l'accès Internet sortant pour l'instance, la connectivité entre l'agent de console et la classification des données via le port 443, et bien plus encore. << Prérequis,Voir la liste complète>>.

3 Déployer la classification des données

Lancez l'assistant d'installation pour déployer l'instance de classification des données dans le cloud.

Créer un agent de console

Si vous ne disposez pas encore d'un agent de console, créez un agent de console chez votre fournisseur de cloud. Voir "création d'un agent de console dans AWS" ou "création d'un agent de console dans Azure" , ou "création d'un agent de console dans GCP" . Dans la plupart des cas, vous aurez probablement configuré un agent de console avant de tenter d'activer la classification des données, car la plupart "Les fonctionnalités de la console nécessitent un agent de console" , mais il y a des cas où vous devrez en créer un maintenant.

Il existe certains scénarios dans lesquels vous devez utiliser un agent de console déployé chez un fournisseur de cloud spécifique :

- Lors de l'analyse des données dans Cloud Volumes ONTAP dans AWS ou Amazon FSx pour les compartiments ONTAP , vous utilisez un agent de console dans AWS.
- Lors de l'analyse des données dans Cloud Volumes ONTAP dans Azure ou dans Azure NetApp Files, vous utilisez un agent de console dans Azure.
  - Pour Azure NetApp Files, il doit être déployé dans la même région que les volumes que vous souhaitez analyser.
- Lors de l'analyse des données dans Cloud Volumes ONTAP dans GCP, vous utilisez un agent de console dans GCP.

Les systèmes ONTAP sur site, les partages de fichiers NetApp et les bases de données peuvent être analysés lors de l'utilisation de l'un de ces agents de console cloud.

Notez que vous pouvez également "installer l'agent de console sur site" sur un hôte Linux dans votre réseau ou dans le cloud. Certains utilisateurs prévoyant d'installer la classification des données sur site peuvent également choisir d'installer l'agent de console sur site.

Comme vous pouvez le constater, il peut y avoir des situations où vous devrez utiliser "plusieurs agents de console" .

La classification des données n'impose pas de limite à la quantité de données qu'elle peut analyser. Chaque agent de console prend en charge l'analyse et l'affichage de 500 Tio de données. Pour scanner plus de 500 Tio de données,"installer un autre agent de console" alors"déployer une autre instance de classification des données" . + L'interface utilisateur de la console affiche les données d'un seul connecteur. Pour obtenir des conseils sur l'affichage des données de plusieurs agents de console, consultez"Travailler avec plusieurs agents de console" .

Soutien gouvernemental régional

La classification des données est prise en charge lorsque l'agent de console est déployé dans une région gouvernementale (AWS GovCloud, Azure Gov ou Azure DoD). Lorsqu'elle est déployée de cette manière, la classification des données présente les restrictions suivantes :

"En savoir plus sur le déploiement de l'agent de console dans une région gouvernementale" .

Prérequis

Passez en revue les conditions préalables suivantes pour vous assurer que vous disposez d'une configuration prise en charge avant de déployer la classification des données dans le cloud. Lorsque vous déployez la classification des données dans le cloud, elle est située dans le même sous-réseau que l'agent de la console.

Activer l'accès Internet sortant à partir de la classification des données

La classification des données nécessite un accès Internet sortant. Si votre réseau virtuel ou physique utilise un serveur proxy pour l'accès à Internet, assurez-vous que l'instance de classification des données dispose d'un accès Internet sortant pour contacter les points de terminaison suivants. Le proxy doit être non transparent. Les proxys transparents ne sont actuellement pas pris en charge.

Consultez le tableau approprié ci-dessous selon que vous déployez la classification des données dans AWS, Azure ou GCP.

Points de terminaison requis pour AWS

Points de terminaison	But
\ https://api.console.netapp.com	Communication avec le service Console, qui inclut les comptes NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Communication avec le site Web de la console pour l'authentification centralisée des utilisateurs.
\ 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/	Fournit un accès aux images logicielles, aux manifestes et aux modèles.
\ https://kinesis.us-east-1.amazonaws.com	Permet à NetApp de diffuser des données à partir des enregistrements d'audit.
\ 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	Permet à la classification des données d'accéder et de télécharger des manifestes et des modèles, et d'envoyer des journaux et des métriques.

Points de terminaison requis pour Azure

Points de terminaison	But
\ https://api.console.netapp.com	Communication avec le service Console, qui inclut les comptes NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Communication avec le site Web de la console pour l'authentification centralisée des utilisateurs.
\ 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/	Fournit un accès aux images logicielles, aux manifestes, aux modèles et permet d'envoyer des journaux et des métriques.
\ https://support.compliance.api.console.netapp.com/	Permet à NetApp de diffuser des données à partir des enregistrements d'audit.

Points de terminaison requis pour GCP

Points de terminaison	But
\ https://api.console.netapp.com	Communication avec le service Console, qui inclut les comptes NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Communication avec le site Web de la console pour l'authentification centralisée des utilisateurs.

Points de terminaison	But
\ 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/	Fournit un accès aux images logicielles, aux manifestes, aux modèles et permet d'envoyer des journaux et des métriques.
\ https://support.compliance.api.console.netapp.com/	Permet à NetApp de diffuser des données à partir des enregistrements d'audit.

Assurez-vous que la classification des données dispose des autorisations requises

Assurez-vous que Data Classification dispose des autorisations nécessaires pour déployer des ressources et créer des groupes de sécurité pour l'instance Data Classification.

- "Autorisations Google Cloud"
- "Autorisations AWS"
- "Autorisations Azure"

Assurez-vous que l'agent de la console peut accéder à la classification des données

Assurez la connectivité entre l'agent de console et l'instance de classification des données. Le groupe de sécurité de l'agent de console doit autoriser le trafic entrant et sortant sur le port 443 vers et depuis l'instance de classification des données. Cette connexion permet le déploiement de l'instance de classification des données et vous permet d'afficher les informations dans les onglets Conformité et Gouvernance. La classification des données est prise en charge dans les régions gouvernementales dans AWS et Azure.

Des règles de groupe de sécurité entrantes et sortantes supplémentaires sont requises pour les déploiements AWS et AWS GovCloud. Voir "Règles pour l'agent de console dans AWS" pour plus de détails.

Des règles de groupe de sécurité entrantes et sortantes supplémentaires sont requises pour les déploiements Azure et Azure Government. Voir "Règles pour l'agent de console dans Azure" pour plus de détails.

Assurez-vous de pouvoir maintenir la classification des données en cours d'exécution

L'instance de classification des données doit rester active pour analyser en continu vos données.

Assurer la connectivité du navigateur Web à la classification des données

Une fois la classification des données activée, assurez-vous que les utilisateurs accèdent à l'interface de la console à partir d'un hôte disposant d'une connexion à l'instance de classification des données.

L'instance de classification des données utilise une adresse IP privée pour garantir que les données indexées ne sont pas accessibles sur Internet. Par conséquent, le navigateur Web que vous utilisez pour accéder à la console doit disposer d'une connexion à cette adresse IP privée. Cette connexion peut provenir d'une connexion directe à votre fournisseur de cloud (par exemple, un VPN) ou d'un hôte situé sur le même réseau que l'instance de classification des données.

Vérifiez vos limites de vCPU

Assurez-vous que la limite vCPU de votre fournisseur de cloud permet le déploiement d'une instance avec le nombre de cœurs nécessaire. Vous devrez vérifier la limite de vCPU pour la famille d'instances concernée dans la région où la console s'exécute. "Voir les types d'instances requis" .

Consultez les liens suivants pour plus de détails sur les limites du vCPU :

- "Documentation AWS : quotas de service Amazon EC2"
- "Documentation Azure : Quotas de processeurs virtuels pour machines virtuelles"
- "Documentation Google Cloud : quotas de ressources"

Déployer la classification des données dans le cloud

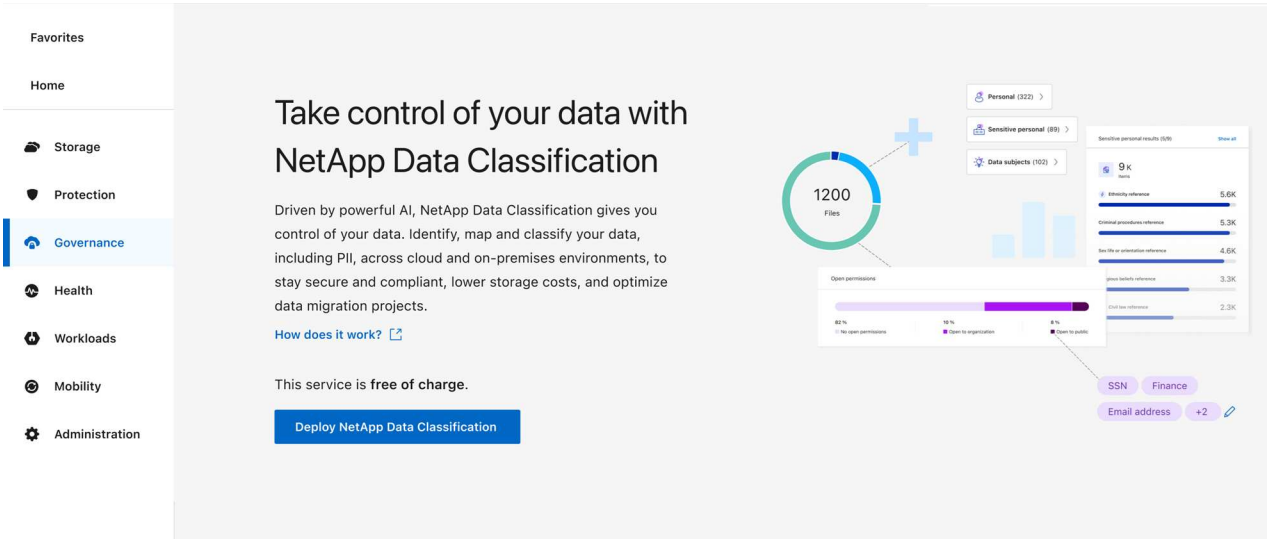
Suivez ces étapes pour déployer une instance de classification des données dans le cloud. L'agent de console déploiera l'instance dans le cloud, puis installera le logiciel de classification des données sur cette instance.

Dans les régions où le type d'instance par défaut n'est pas disponible, la classification des données s'exécute sur un"type d'instance alternatif" .

Déployer dans AWS

Étapes

1. Depuis la page principale de la classification des données, sélectionnez **Déployer la classification sur site ou dans le cloud**.

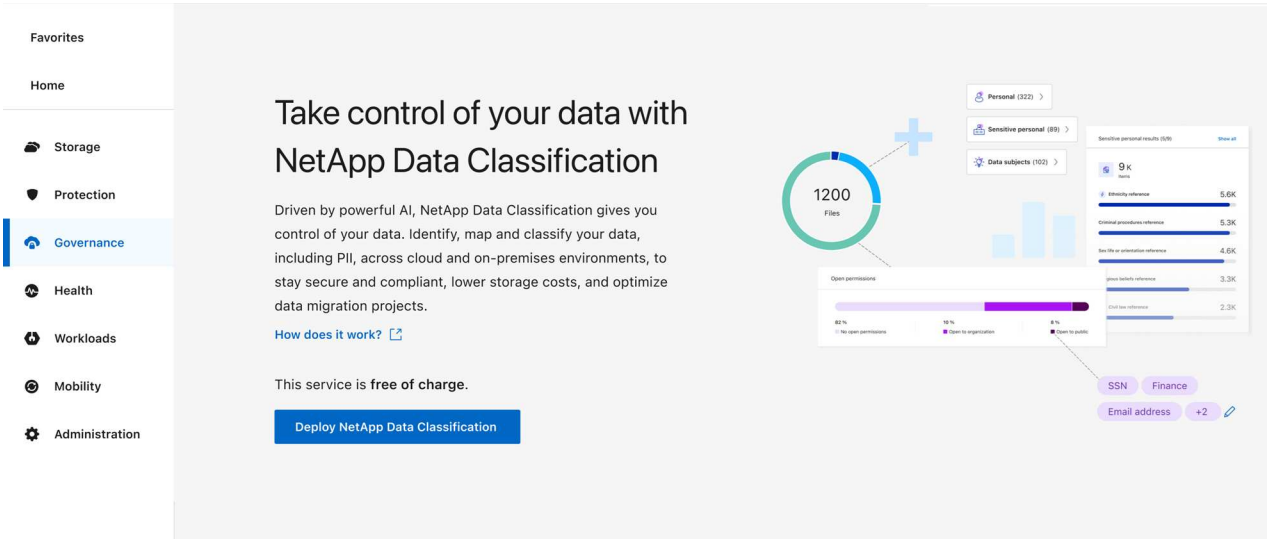


2. Depuis la page *Installation*, sélectionnez **Déployer > Déployer** pour utiliser la taille d'instance « Grande » et démarrer l'assistant de déploiement cloud.
3. L'assistant affiche la progression au fur et à mesure des étapes de déploiement. Lorsque des entrées sont requises ou si vous rencontrez des problèmes, vous êtes invité à le faire.
4. Une fois l'instance déployée et la classification des données installée, sélectionnez **Continuer vers la configuration** pour accéder à la page *Configuration*.

Déployer dans Azure

Étapes

1. Depuis la page principale de la classification des données, sélectionnez **Déployer la classification sur site ou dans le cloud**.



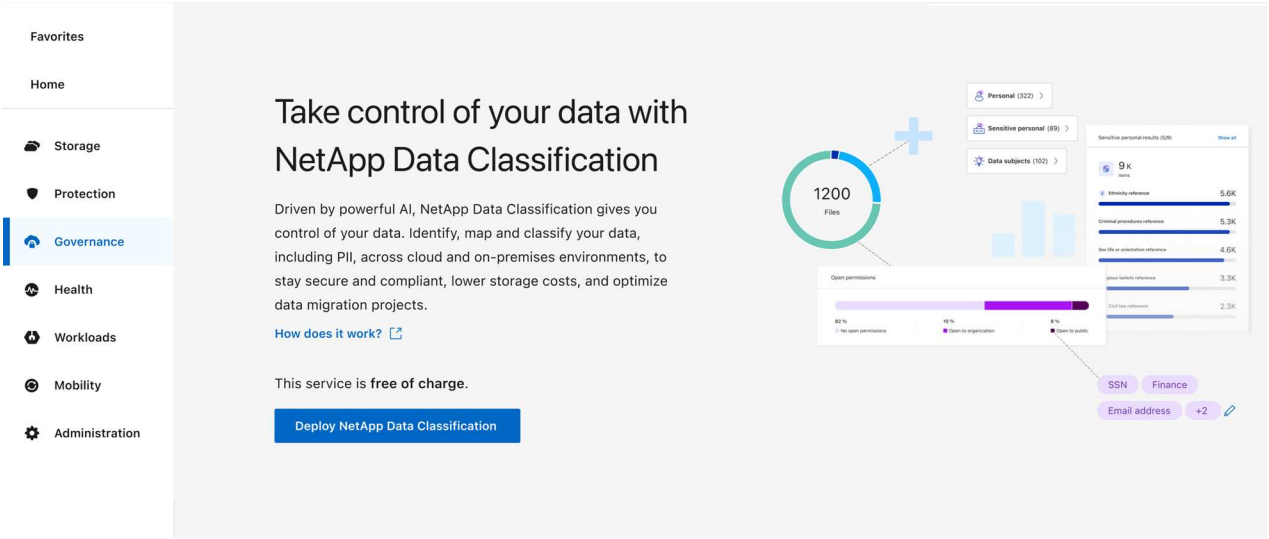
2. Sélectionnez **Déployer** pour démarrer l'assistant de déploiement cloud.

3. L'assistant affiche la progression au fur et à mesure des étapes de déploiement. Il s'arrêtera et demandera une saisie s'il rencontre des problèmes.
4. Une fois l'instance déployée et la classification des données installée, sélectionnez **Continuer vers la configuration** pour accéder à la page *Configuration*.

Déployer dans Google Cloud

Étapes

1. Depuis la page principale de la classification des données, sélectionnez **Gouvernance > Classification**.
2. Sélectionnez **Déployer la classification sur site ou dans le cloud**.



3. Sélectionnez **Déployer** pour démarrer l'assistant de déploiement cloud.
4. L'assistant affiche la progression au fur et à mesure des étapes de déploiement. Il s'arrêtera et demandera une saisie s'il rencontre des problèmes.
5. Une fois l'instance déployée et la classification des données installée, sélectionnez **Continuer vers la configuration** pour accéder à la page *Configuration*.

Résultat

La console déploie l'instance de classification des données dans votre fournisseur de cloud.

Les mises à niveau de l'agent de console et du logiciel de classification des données sont automatisées tant que les instances disposent d'une connectivité Internet.

Et ensuite?

Depuis la page Configuration, vous pouvez sélectionner les sources de données que vous souhaitez analyser.

Installer NetApp Data Classification sur un hôte disposant d'un accès Internet

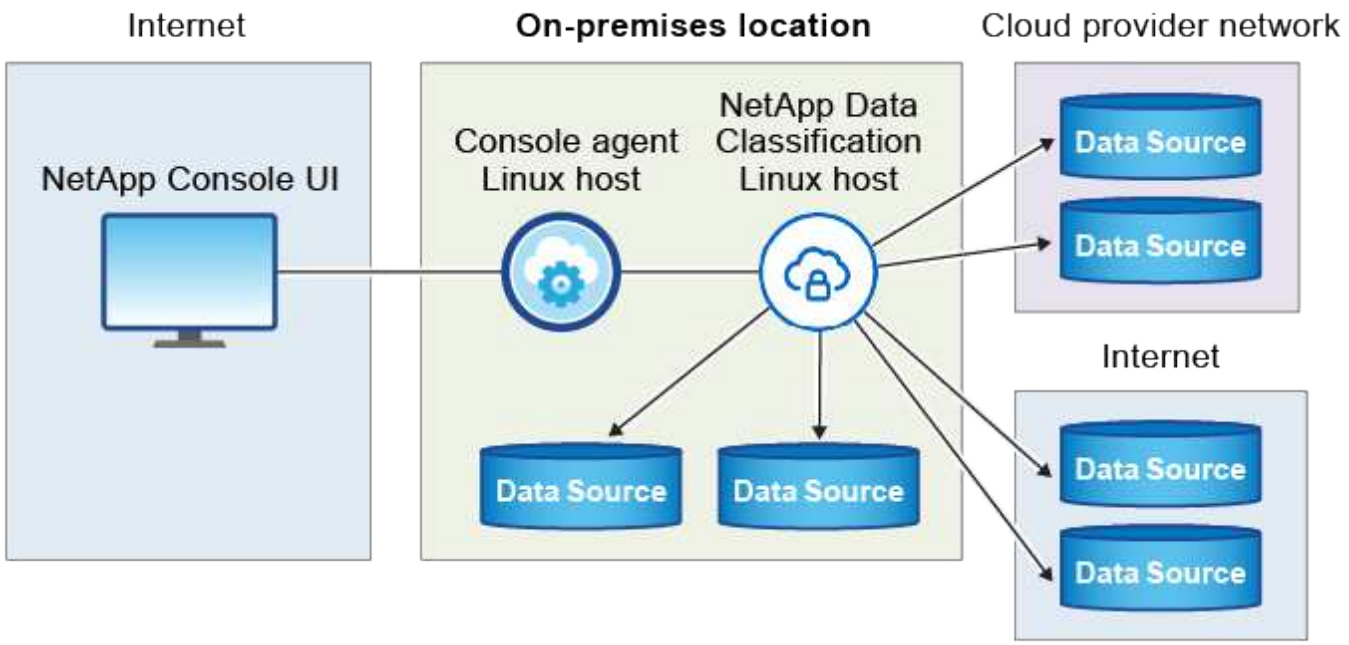
Pour déployer NetApp Data Classification sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud disposant d'un accès Internet, vous devez déployer l'hôte Linux manuellement sur votre réseau ou dans le cloud.

L'installation sur site est une bonne option si vous préférez analyser les systèmes ONTAP sur site à l'aide

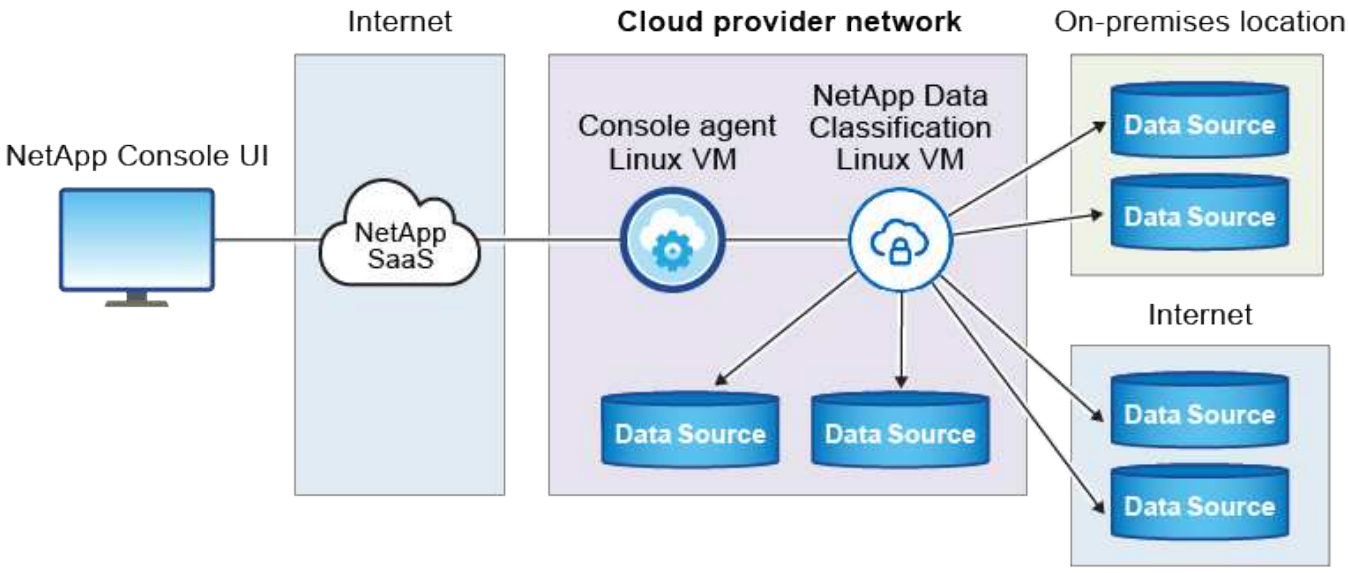
d'une instance de classification des données également située sur site. Ce n'est pas une exigence. Le logiciel fonctionne de la même manière quelle que soit la méthode d'installation choisie.

Le script d'installation de la classification des données commence par vérifier si le système et l'environnement répondent aux prérequis requis. Si toutes les conditions préalables sont remplies, l'installation démarre. Si vous souhaitez vérifier les conditions préalables indépendamment de l'exécution de l'installation de la classification des données, vous pouvez télécharger un progiciel distinct qui teste uniquement les conditions préalables. "Découvrez comment vérifier si votre hôte Linux est prêt à installer la classification des données" .

L'installation typique sur un hôte Linux *dans vos locaux* comporte les composants et connexions suivants.



L'installation typique sur un hôte Linux *dans le cloud* comporte les composants et connexions suivants.



Démarrage rapide

Commencez rapidement en suivant ces étapes ou faites défiler les sections restantes pour obtenir tous les

détails.

1 Créer un agent de console

Si vous n'avez pas encore d'agent de console, "déployer l'agent de console sur site" sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud.

Vous pouvez également créer un agent de console avec votre fournisseur de cloud. Voir "création d'un agent de console dans AWS" , "création d'un agent de console dans Azure" , ou "création d'un agent de console dans GCP" .

2 Réviser les prérequis

Assurez-vous que votre environnement peut répondre aux prérequis. Cela inclut l'accès Internet sortant pour l'instance, la connectivité entre l'agent de console et la classification des données via le port 443, et bien plus encore. Voir la liste complète .

Vous avez également besoin d'un système Linux qui répond auxexigences suivantes .

3 Téléchargez et déployez la classification des données

Téléchargez le logiciel Cloud Data Classification à partir du site de support NetApp et copiez le fichier d'installation sur l'hôte Linux que vous prévoyez d'utiliser. Lancez ensuite l'assistant d'installation et suivez les invites pour déployer l'instance de classification des données.

Créer un agent de console

Un agent de console est requis avant de pouvoir installer et utiliser la classification des données. Dans la plupart des cas, vous aurez probablement configuré un agent de console avant de tenter d'activer la classification des données, car la plupart "Les fonctionnalités de la console nécessitent un agent de console" , mais il y a des cas où vous devrez en créer un maintenant.

Pour en créer un dans votre environnement de fournisseur de cloud, consultez "création d'un agent de console dans AWS" , "création d'un agent de console dans Azure" , ou "création d'un agent de console dans GCP" .

Il existe certains scénarios dans lesquels vous devez utiliser un agent de console déployé chez un fournisseur de cloud spécifique :

- Lors de l'analyse des données dans Cloud Volumes ONTAP dans AWS ou Amazon FSx for ONTAP, vous utilisez un agent de console dans AWS.
- Lors de l'analyse des données dans Cloud Volumes ONTAP dans Azure ou dans Azure NetApp Files, vous utilisez un agent de console dans Azure.

  Pour Azure NetApp Files, il doit être déployé dans la même région que les volumes que vous souhaitez analyser.

- Lors de l'analyse des données dans Cloud Volumes ONTAP dans GCP, vous utilisez un agent de console dans GCP.

Les systèmes ONTAP sur site, les partages de fichiers NetApp et les comptes de base de données peuvent être analysés à l'aide de l'un de ces agents de console cloud.

Notez que vous pouvez également "déployer l'agent de console sur site" sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud. Certains utilisateurs prévoyant d'installer la classification des données sur site peuvent également choisir d'installer l'agent de console sur site.

Vous aurez besoin de l'adresse IP ou du nom d'hôte du système d'agent de la console lors de l'installation de la classification des données. Vous disposerez de ces informations si vous avez installé l'agent Console dans vos locaux. Si l'agent de la console est déployé dans le cloud, vous pouvez trouver ces informations depuis la console : sélectionnez l'icône Aide puis **Support** puis **Agent de la console**.

Préparer le système hôte Linux

Le logiciel de classification des données doit s'exécuter sur un hôte qui répond aux exigences spécifiques du système d'exploitation, aux exigences de RAM, aux exigences logicielles, etc. L'hôte Linux peut être dans votre réseau ou dans le cloud.

Assurez-vous de pouvoir maintenir la classification des données en cours d'exécution. La machine de classification des données doit rester allumée pour analyser en continu vos données.

- La classification des données n'est pas prise en charge sur un hôte partagé avec d'autres applications : l'hôte doit être un hôte dédié.
- Lors de la création du système hôte dans vos locaux, vous pouvez choisir parmi ces tailles de système en fonction de la taille de l'ensemble de données que vous prévoyez d'analyser pour la classification des données.

Taille du système	processeur	RAM (la mémoire d'échange doit être désactivée)	Disque
Extra Large	32 processeurs	128 Go de RAM	• 1 Tio SSD sur /, ou 100 Gio disponibles sur /opt • 895 Gio disponibles sur /var/lib/docker • 5 Gio sur /tmp • **Pour Podman, 30 Go sur /var/tmp**
Grand	16 processeurs	64 Go de RAM	• 500 Gio SSD sur /, ou 100 Gio disponibles sur /opt • 400 Gio disponibles sur /var/lib/docker ou pour Podman /var/lib/containers • 5 Gio sur /tmp • **Pour Podman, 30 Go sur /var/tmp**

- Lors du déploiement d'une instance de calcul dans le cloud pour votre installation de classification des données, il est recommandé d'utiliser un système qui répond aux exigences système « Large » ci-dessus :
  - **Type d'instance Amazon Elastic Compute Cloud (Amazon EC2)** : « m6i.4xlarge ». "Voir d'autres types d'instances AWS" .
  - **Taille de la machine virtuelle Azure** : « Standard_D16s_v3 ». "Voir d'autres types d'instances Azure"

.

- **Type de machine GCP** : « n2-standard-16 ». "Voir les types d'instances GCP supplémentaires" .

- **Autorisations de dossier UNIX** : Les autorisations UNIX minimales suivantes sont requises :

Dossier	Autorisations minimales
/tmp	`rwxrwxrwt`
/opter	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/système	`rwxr-xr-x`

- **Système opérateur**:
  - Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Docker :
    - Red Hat Enterprise Linux versions 7.8 et 7.9
    - Ubuntu 22.04 (nécessite la version 1.23 ou supérieure de Data Classification)
    - Ubuntu 24.04 (nécessite la version 1.23 ou supérieure de Data Classification)
  - Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version 1.30 ou supérieure de Data Classification :
    - Red Hat Enterprise Linux versions 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 et 9.6.
  - Les extensions vectorielles avancées (AVX2) doivent être activées sur le système hôte.
- **Red Hat Subscription Management** : L'hôte doit être enregistré auprès de Red Hat Subscription Management. S'il n'est pas enregistré, le système ne peut pas accéder aux référentiels pour mettre à jour les logiciels tiers requis lors de l'installation.
- **Logiciel supplémentaire** : Vous devez installer le logiciel suivant sur l'hôte avant d'installer Data Classification :
  - Selon le système d'exploitation que vous utilisez, vous devrez installer l'un des moteurs de conteneur :
    - Docker Engine version 19.3.1 ou supérieure. "Voir les instructions d'installation" .
    - Podman version 4 ou supérieure. Pour installer Podman, entrez(`sudo yum install podman netavark -y`).
- Version Python 3.6 ou supérieure. "Voir les instructions d'installation" .
  - **Considérations NTP** : NetApp recommande de configurer le système de classification des données pour utiliser un service NTP (Network Time Protocol). L'heure doit être synchronisée entre le système de classification des données et le système d'agent de la console.
- **Considérations relatives au pare-feu** : Si vous envisagez d'utiliser `firewalld` , nous vous recommandons de l'activer avant d'installer Data Classification. Exécutez les commandes suivantes pour configurer `firewalld` afin qu'il soit compatible avec la classification des données :

```
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
```

Si vous prévoyez d'utiliser des hôtes de classification de données supplémentaires comme nœuds de scanner, ajoutez ces règles à votre système principal à ce stade :

```
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
```

Notez que vous devez redémarrer Docker ou Podman chaque fois que vous activez ou mettez à jour `firewalld` paramètres.

L'adresse IP du système hôte de classification des données ne peut pas être modifiée après l'installation.

Activer l'accès Internet sortant à partir de la classification des données

La classification des données nécessite un accès Internet sortant. Si votre réseau virtuel ou physique utilise un serveur proxy pour l'accès à Internet, assurez-vous que l'instance de classification des données dispose d'un accès Internet sortant pour contacter les points de terminaison suivants.

Points de terminaison	But
\ https://api.console.netapp.com	Communication avec la console, qui inclut les comptes NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Communication avec le site Web de la console pour l'authentification centralisée des utilisateurs.
\ 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/	Fournit un accès aux images logicielles, aux manifestes, aux modèles et permet d'envoyer des journaux et des métriques.
\ https://support.compliance.api.console.netapp.com/	Permet à NetApp de diffuser des données à partir des enregistrements d'audit.
\ https://github.com/docker \ https://download.docker.com	Fournit des packages prérequis pour l'installation de Docker.

Points de terminaison	But
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Fournit des packages prérequis pour l'installation d'Ubuntu.

Vérifiez que tous les ports requis sont activés

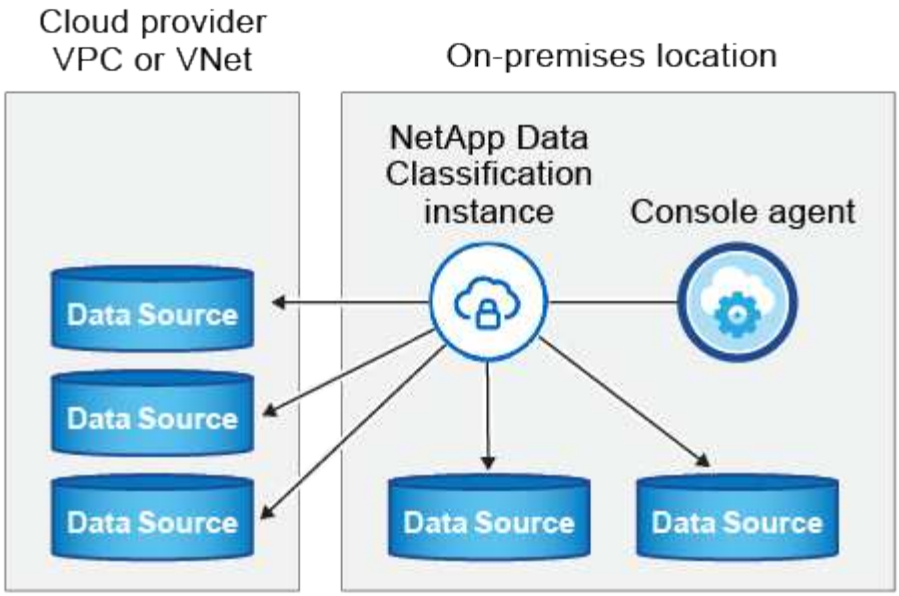
Vous devez vous assurer que tous les ports requis sont ouverts pour la communication entre l'agent de console, la classification des données, Active Directory et vos sources de données.

Type de connexion	Ports	Description
Agent de console <> Classification des données	8080 (TCP), 443 (TCP) et 80. 9000	Les règles de pare-feu ou de routage de l'agent de console doivent autoriser le trafic entrant et sortant sur le port 443 vers et depuis l'instance de classification des données. Assurez-vous que le port 8080 est ouvert afin de pouvoir voir la progression de l'installation dans la console. Si un pare-feu est utilisé sur l'hôte Linux, le port 9000 est requis pour les processus internes au sein d'un serveur Ubuntu.
Agent de console <> cluster ONTAP (NAS)	443 (TCP)	La console découvre les clusters ONTAP à l'aide de HTTPS. Si vous utilisez des stratégies de pare-feu personnalisées, elles doivent répondre aux exigences suivantes : • L'hôte de l'agent de console doit autoriser l'accès HTTPS sortant via le port 443. Si l'agent de la console est dans le cloud, toutes les communications sortantes sont autorisées par le pare-feu prédéfini ou les règles de routage. • Le cluster ONTAP doit autoriser l'accès HTTPS entrant via le port 443. La stratégie de pare-feu « mgmt » par défaut autorise l'accès HTTPS entrant à partir de toutes les adresses IP. Si vous avez modifié cette politique par défaut ou si vous avez créé votre propre politique de pare-feu, vous devez associer le protocole HTTPS à cette politique et activer l'accès à partir de l'hôte de l'agent de la console.

Type de connexion	Ports	Description
Classification des données <> cluster ONTAP	• Pour NFS - 111 (TCP\UDP) et 2049 (TCP\UDP) • Pour CIFS - 139 (TCP\UDP) et 445 (TCP\UDP)	La classification des données nécessite une connexion réseau à chaque sous-réseau Cloud Volumes ONTAP ou à un système ONTAP sur site. Les pare-feu ou les règles de routage pour Cloud Volumes ONTAP doivent autoriser les connexions entrantes à partir de l'instance de classification des données. Assurez-vous que ces ports sont ouverts à l'instance de classification des données : • Pour NFS - 111 et 2049 • Pour CIFS - 139 et 445 Les stratégies d'exportation de volume NFS doivent autoriser l'accès à partir de l'instance de classification des données.
Classification des données <> Active Directory	389 (TCP et UDP), 636 (TCP), 3268 (TCP) et 3269 (TCP)	Vous devez déjà avoir un Active Directory configuré pour les utilisateurs de votre entreprise. De plus, la classification des données nécessite des informations d'identification Active Directory pour analyser les volumes CIFS. Vous devez disposer des informations pour Active Directory : • Adresse IP du serveur DNS ou plusieurs adresses IP • Nom d'utilisateur et mot de passe pour le serveur • Nom de domaine (nom Active Directory) • Que vous utilisiez ou non un LDAP sécurisé (LDAPS) • Port du serveur LDAP (généralement 389 pour LDAP et 636 pour LDAP sécurisé)

Installer la classification des données sur l'hôte Linux

Pour les configurations typiques, vous installerez le logiciel sur un seul système hôte. Voir ces étapes ici .



VoirPréparation du système hôte Linux etRévision des prérequis pour obtenir la liste complète des exigences avant de déployer la classification des données.

Les mises à niveau du logiciel de classification des données sont automatisées tant que l'instance dispose d'une connexion Internet.

> La classification des données ne peut actuellement pas analyser les compartiments S3, Azure NetApp Files ou FSx pour ONTAP lorsque le logiciel est installé sur site. Dans ces cas, vous devrez déployer un agent de console distinct et une instance de classification des données dans le cloud et "basculer entre les connecteurs" pour vos différentes sources de données.

Installation sur un seul hôte pour les configurations typiques

Passez en revue les exigences et suivez ces étapes lors de l'installation du logiciel de classification des données sur un seul hôte local.

"Regardez cette vidéo"pour voir comment installer Data Classification.

Notez que toutes les activités d'installation sont enregistrées lors de l'installation de Data Classification. Si vous rencontrez des problèmes lors de l'installation, vous pouvez afficher le contenu du journal d'audit d'installation. Il est écrit à `/opt/netapp/install_logs/` .

Avant de commencer

- Vérifiez que votre système Linux répond auxexigences de l'hôte .
- Vérifiez que le système dispose des deux packages logiciels prérequis installés (Docker Engine ou Podman et Python 3).
- Assurez-vous que vous disposez des privilèges root sur le système Linux.
- Si vous utilisez un proxy pour accéder à Internet :
  - Vous aurez besoin des informations du serveur proxy (adresse IP ou nom d'hôte, port de connexion, schéma de connexion : https ou http, nom d'utilisateur et mot de passe).
  - Si le proxy effectue une interception TLS, vous devez connaître le chemin sur le système Linux de classification des données où les certificats CA TLS sont stockés.

- Le proxy doit être non transparent. La classification des données ne prend actuellement pas en charge les proxys transparents.
- L'utilisateur doit être un utilisateur local. Les utilisateurs de domaine ne sont pas pris en charge.

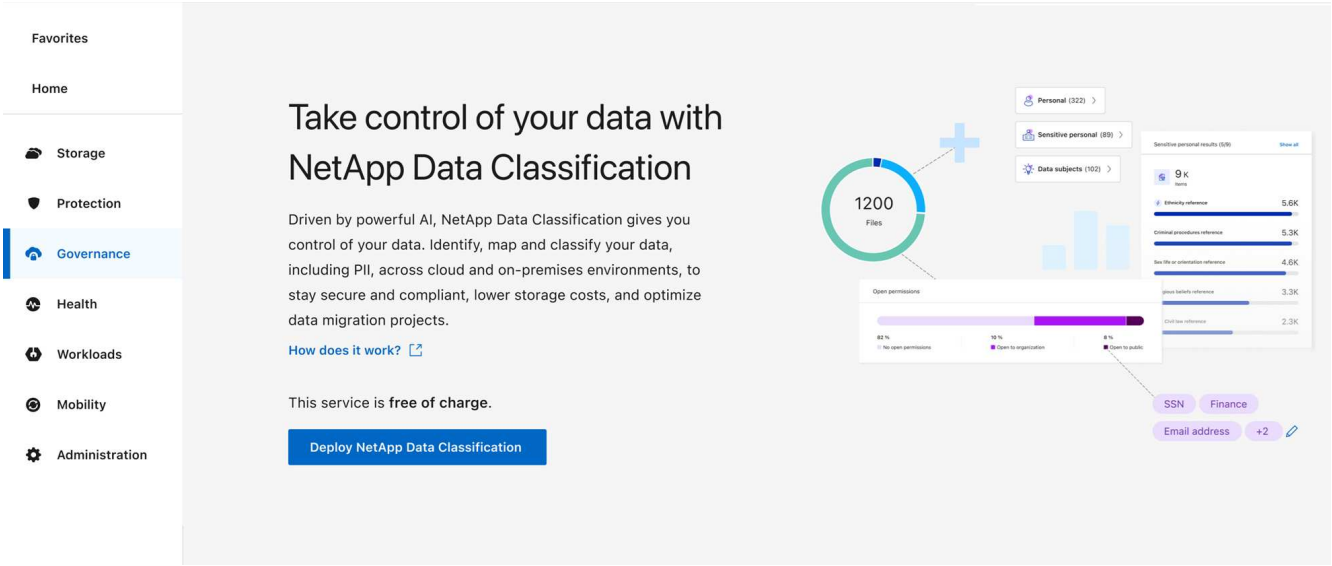
- Vérifiez que votre environnement hors ligne répond aux exigences requisesautorisations et connectivité .

Étapes

1. Téléchargez le logiciel de classification des données à partir du "Site de support NetApp" . Le fichier que vous devez sélectionner est nommé **DATASENSE-INSTALLER-<version>.tar.gz**.
2. Copiez le fichier d'installation sur l'hôte Linux que vous prévoyez d'utiliser (en utilisant `scp` ou une autre méthode).
3. Décompressez le fichier d'installation sur la machine hôte, par exemple :

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

4. Dans la console, sélectionnez **Gouvernance > Classification**.
5. Sélectionnez **Déployer la classification sur site ou dans le cloud**.



6. Selon que vous installez Data Classification sur une instance que vous avez préparée dans le cloud ou sur une instance que vous avez préparée dans vos locaux, sélectionnez l'option **Déployer** appropriée pour démarrer l'installation de Data Classification.
7. La boîte de dialogue *Déployer la classification des données sur site* s'affiche. Copiez la commande fournie (par exemple : `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`) et collez-le dans un fichier texte pour pouvoir l'utiliser plus tard. Sélectionnez ensuite **Fermer** pour fermer la boîte de dialogue.
8. Sur la machine hôte, entrez la commande que vous avez copiée, puis suivez une série d'invites, ou vous pouvez fournir la commande complète, y compris tous les paramètres requis, comme arguments de ligne de commande.

   Notez que le programme d'installation effectue une pré-vérification pour s'assurer que votre système et vos exigences réseau sont en place pour une installation réussie. "Regardez cette vidéo" pour comprendre les messages et les implications du pré-contrôle.

Entrez les paramètres comme demandé :	Entrez la commande complète :
a. Collez la commande que vous avez copiée à l'étape 7 : `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` Si vous effectuez l'installation sur une instance cloud (pas dans vos locaux), ajoutez `--manual-cloud-install <cloud_provider>`. b. Saisissez l'adresse IP ou le nom d'hôte de la machine hôte de classification des données afin que le système d'agent de la console puisse y accéder. c. Saisissez l'adresse IP ou le nom d'hôte de la machine hôte de l'agent de console afin que le système de classification des données puisse y accéder. d. Saisissez les détails du proxy lorsque vous y êtes invité. Si votre agent de console utilise déjà un proxy, il n'est pas nécessaire de saisir à nouveau ces informations ici, car la classification des données utilisera automatiquement le proxy utilisé par l'agent de console.	Alternativement, vous pouvez créer la commande entière à l'avance, en fournissant les paramètres d'hôte et de proxy nécessaires : `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>`

Valeurs des variables :

- *account_id* = ID de compte NetApp
- *client_id* = ID client de l'agent de console (ajoutez le suffixe « clients » à l'ID client s'il n'est pas déjà présent)
- *user_token* = jeton d'accès utilisateur JWT
- *ds_host* = Adresse IP ou nom d'hôte du système Linux de classification des données.
- *cm_host* = Adresse IP ou nom d'hôte du système agent de la console.
- *cloud_provider* = Lors de l'installation sur une instance cloud, saisissez « AWS », « Azure » ou « Gcp » selon le fournisseur de cloud.
- *proxy_host* = IP ou nom d'hôte du serveur proxy si l'hôte est derrière un serveur proxy.
- *proxy_port* = Port de connexion au serveur proxy (par défaut 80).
- *proxy_scheme* = Schéma de connexion : https ou http (par défaut http).
- *proxy_user* = Utilisateur authentifié pour se connecter au serveur proxy, si une authentification de base est requise. L'utilisateur doit être un utilisateur local - les utilisateurs de domaine ne sont pas pris en charge.
- *proxy_password* = Mot de passe pour le nom d'utilisateur que vous avez spécifié.
- *ca_cert_dir* = Chemin sur le système Linux de classification des données contenant des ensembles de certificats CA TLS supplémentaires. Requis uniquement si le proxy effectue une interception TLS.

Résultat

Le programme d'installation de la classification des données installe les packages, enregistre l'installation et installe la classification des données. L'installation peut prendre 10 à 20 minutes.

S'il existe une connectivité via le port 8080 entre la machine hôte et l'instance de l'agent de la console, vous verrez la progression de l'installation dans l'onglet Classification des données de la console.

Et ensuite?

Depuis la page Configuration, vous pouvez sélectionner les sources de données que vous souhaitez analyser.

Installer NetApp Data Classification sur un hôte Linux sans accès Internet

L'installation de NetApp Data Classification sur un hôte Linux dans un site local qui n'a pas accès à Internet est appelée *mode privé*. Ce type d'installation, qui utilise un script d'installation, n'a aucune connectivité à la couche SaaS de la NetApp Console .

> Le mode privé BlueXP (interface BlueXP héritée) est généralement utilisé avec des environnements locaux qui n'ont pas de connexion Internet et avec des régions cloud sécurisées, notamment AWS Secret Cloud, AWS Top Secret Cloud et Azure IL6. NetApp continue de prendre en charge ces environnements avec l'interface BlueXP héritée. Pour la documentation du mode privé dans l'ancienne interface BlueXP , voir"Documentation PDF pour le mode privé BlueXP" .

Vérifiez que votre hôte Linux est prêt à installer NetApp Data Classification

Avant d'installer manuellement NetApp Data Classification sur un hôte Linux, exécutez éventuellement un script sur l'hôte pour vérifier que toutes les conditions préalables sont réunies pour l'installation de Data Classification. Vous pouvez exécuter ce script sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud. L'hôte peut être connecté à Internet ou résider sur un site qui n'a pas accès à Internet (un *site sombre*).

Il existe également un script de test prérequis qui fait partie du script d'installation de la classification des données. Le script décrit ici est spécifiquement conçu pour les utilisateurs qui souhaitent vérifier l'hôte Linux indépendamment de l'exécution du script d'installation de la classification des données.

Commencer

Vous effectuerez les tâches suivantes.

1. Vous pouvez également installer un agent de console si vous n'en avez pas déjà un installé. Vous pouvez exécuter le script de test sans avoir installé d'agent de console, mais le script vérifie la connectivité entre l'agent de console et la machine hôte de classification des données. Il est donc recommandé de disposer d'un agent de console.
2. Préparez la machine hôte et vérifiez qu'elle répond à toutes les exigences.
3. Activez l'accès Internet sortant à partir de la machine hôte de classification des données.
4. Vérifiez que tous les ports requis sont activés sur tous les systèmes.
5. Téléchargez et exécutez le script de test prérequis.

Créer un agent de console

Un agent de console est requis avant de pouvoir installer et utiliser la classification des données. Vous pouvez toutefois exécuter le script Prérequis sans agent de console.

Tu peux "installer l'agent de console sur site" sur un hôte Linux de votre réseau ou sur un hôte Linux dans le cloud. Certains utilisateurs prévoyant d'installer la classification des données sur site peuvent également choisir d'installer l'agent de console sur site.

Pour créer un agent de console dans votre environnement de fournisseur de cloud, consultez "création d'un agent de console dans AWS" , "création d'un agent de console dans Azure" , ou "création d'un agent de console dans GCP" .

Vous aurez besoin de l'adresse IP ou du nom d'hôte du système d'agent de la console lors de l'exécution du script Prérequis. Vous disposerez de ces informations si vous avez installé l'agent Console dans vos locaux. Si l'agent Console est déployé dans le cloud, vous pouvez trouver ces informations depuis la Console : sélectionnez l'icône Aide puis **Support** puis **Agent Console**.

Vérifier les exigences de l'hôte

Le logiciel de classification des données doit s'exécuter sur un hôte qui répond aux exigences spécifiques du système d'exploitation, aux exigences de RAM, aux exigences logicielles, etc.

- La classification des données n'est pas prise en charge sur un hôte partagé avec d'autres applications : l'hôte doit être un hôte dédié.
- Lors de la création du système hôte dans vos locaux, vous pouvez choisir parmi ces tailles de système en fonction de la taille de l'ensemble de données que vous prévoyez d'analyser pour la classification des données.

Taille du système	processeur	RAM (la mémoire d'échange doit être désactivée)	Disque
Extra Large	32 processeurs	128 Go de RAM	• 1 Tio SSD sur /, ou 100 Gio disponibles sur /opt • 895 Gio disponibles sur /var/lib/docker • 5 Gio sur /tmp • **Pour Podman, 30 Go sur /var/tmp**
Grand	16 processeurs	64 Go de RAM	• 500 Gio SSD sur /, ou 100 Gio disponibles sur /opt • 400 Gio disponibles sur /var/lib/docker ou pour Podman /var/lib/containers • 5 Gio sur /tmp • **Pour Podman, 30 Go sur /var/tmp**

- Lors du déploiement d'une instance de calcul dans le cloud pour votre installation de classification des

données, il est recommandé d'utiliser un système qui répond aux exigences système « Large » ci-dessus :

  - **Type d'instance Amazon Elastic Compute Cloud (Amazon EC2)** : « m6i.4xlarge ». "Voir d'autres types d'instances AWS" .
  - **Taille de la machine virtuelle Azure** : « Standard_D16s_v3 ». "Voir d'autres types d'instances Azure" .
  - **Type de machine GCP** : « n2-standard-16 ». "Voir les types d'instances GCP supplémentaires" .

- **Autorisations de dossier UNIX** : Les autorisations UNIX minimales suivantes sont requises :

Dossier	Autorisations minimales
/tmp	`rwxrwxrwt`
/opter	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/système	`rwxr-xr-x`

- **Système opérateur**:
  - Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Docker :
    - Red Hat Enterprise Linux versions 7.8 et 7.9
    - Ubuntu 22.04 (nécessite la version 1.23 ou supérieure de Data Classification)
    - Ubuntu 24.04 (nécessite la version 1.23 ou supérieure de Data Classification)
  - Les systèmes d'exploitation suivants nécessitent l'utilisation du moteur de conteneur Podman et nécessitent la version 1.30 ou supérieure de Data Classification :
    - Red Hat Enterprise Linux versions 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 et 9.6.
  - Les extensions vectorielles avancées (AVX2) doivent être activées sur le système hôte.
- **Red Hat Subscription Management** : L'hôte doit être enregistré auprès de Red Hat Subscription Management. S'il n'est pas enregistré, le système ne peut pas accéder aux référentiels pour mettre à jour les logiciels tiers requis lors de l'installation.
- **Logiciel supplémentaire** : Vous devez installer le logiciel suivant sur l'hôte avant d'installer Data Classification :
  - Selon le système d'exploitation que vous utilisez, vous devrez installer l'un des moteurs de conteneur :
    - Docker Engine version 19.3.1 ou supérieure. "Voir les instructions d'installation" .
    - Podman version 4 ou supérieure. Pour installer Podman, entrez(`sudo yum install podman netavark -y`).
- Version Python 3.6 ou supérieure. "Voir les instructions d'installation" .
  - **Considérations NTP** : NetApp recommande de configurer le système de classification des données pour utiliser un service NTP (Network Time Protocol). L'heure doit être synchronisée entre le système de classification des données et le système d'agent de la console.
- **Considérations relatives au pare-feu** : Si vous envisagez d'utiliser `firewalld` , nous vous recommandons de l'activer avant d'installer Data Classification. Exécutez les commandes suivantes pour configurer `firewalld` afin qu'il soit compatible avec la classification des données :

```
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
```

Si vous prévoyez d'utiliser des hôtes de classification de données supplémentaires comme nœuds de scanner (dans un modèle distribué), ajoutez ces règles à votre système principal à ce stade :

```
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
```

Notez que vous devez redémarrer Docker ou Podman chaque fois que vous activez ou mettez à jour `firewalld` paramètres.

Activer l'accès Internet sortant à partir de la classification des données

La classification des données nécessite un accès Internet sortant. Si votre réseau virtuel ou physique utilise un serveur proxy pour l'accès à Internet, assurez-vous que l'instance de classification des données dispose d'un accès Internet sortant pour contacter les points de terminaison suivants.

Cette section n'est pas requise pour les systèmes hôtes installés sur des sites sans connectivité Internet.

Points de terminaison	But
\ https://api.console.netapp.com	Communication avec le service Console, qui inclut les comptes NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Communication avec le site Web de la console pour l'authentification centralisée des utilisateurs.
\ 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/	Fournit un accès aux images logicielles, aux manifestes, aux modèles et permet d'envoyer des journaux et des métriques.
\ https://support.compliance.api.console.netapp.com/	Permet à NetApp de diffuser des données à partir des enregistrements d'audit.
\ https://github.com/docker \ https://download.docker.com	Fournit des packages prérequis pour l'installation de Docker.

Points de terminaison	But
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Fournit des packages prérequis pour l'installation d'Ubuntu.

Vérifiez que tous les ports requis sont activés

Vous devez vous assurer que tous les ports requis sont ouverts pour la communication entre l'agent de console, la classification des données, Active Directory et vos sources de données.

Type de connexion	Ports	Description
Agent de console <> Classification des données	8080 (TCP), 443 (TCP) et 80. 9000	Les règles de pare-feu ou de routage de l'agent de console doivent autoriser le trafic entrant et sortant sur le port 443 vers et depuis l'instance de classification des données. Assurez-vous que le port 8080 est ouvert afin de pouvoir voir la progression de l'installation dans la console. Si un pare-feu est utilisé sur l'hôte Linux, le port 9000 est requis pour les processus internes au sein d'un serveur Ubuntu.
Agent de console <> cluster ONTAP (NAS)	443 (TCP)	La console découvre les clusters ONTAP à l'aide de HTTPS. Si vous utilisez des stratégies de pare-feu personnalisées, l'hôte de l'agent de console doit autoriser l'accès HTTPS sortant via le port 443. Si l'agent de la console est dans le cloud, toutes les communications sortantes sont autorisées par le pare-feu prédéfini ou les règles de routage.

Exécutez le script des prérequis de classification des données

Suivez ces étapes pour exécuter le script des prérequis de classification des données.

"Regardez cette vidéo"pour voir comment exécuter le script Prérequis et interpréter les résultats.

Avant de commencer

- Vérifiez que votre système Linux répond auxexigences de l'hôte .
- Vérifiez que le système dispose des deux packages logiciels prérequis installés (Docker Engine ou Podman et Python 3).
- Assurez-vous que vous disposez des privilèges root sur le système Linux.

Étapes

1. Téléchargez le script des prérequis de classification des données à partir du "Site de support NetApp" . Le fichier que vous devez sélectionner est nommé **standalone-pre-requisite-tester-<version>**.
2. Copiez le fichier sur l'hôte Linux que vous prévoyez d'utiliser (en utilisant `scp` ou une autre méthode).
3. Attribuer des autorisations pour exécuter le script.

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

4. Exécutez le script à l'aide de la commande suivante.

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

   Ajoutez l'option « --darksite » uniquement si vous exécutez le script sur un hôte qui n'a pas accès à Internet. Certains tests prérequis sont ignorés lorsque l'hôte n'est pas connecté à Internet.

5. Le script vous demande l'adresse IP de la machine hôte de classification des données.
   - Entrez l'adresse IP ou le nom d'hôte.
6. Le script vous demande si vous disposez d'un agent de console installé.
   - Entrez **N** si vous n'avez pas d'agent de console installé.
   - Entrez **Y** si vous avez un agent de console installé. Ensuite, entrez l'adresse IP ou le nom d'hôte de l'agent de la console afin que le script de test puisse tester cette connectivité.
7. Le script exécute une variété de tests sur le système et affiche les résultats au fur et à mesure de sa progression. Une fois terminé, il écrit un journal de la session dans un fichier nommé `prerequisites-test-<timestamp>.log` dans le répertoire `/opt/netapp/install_logs`.

Résultat

Si tous les tests prérequis se sont déroulés avec succès, vous pouvez installer Data Classification sur l'hôte lorsque vous êtes prêt.

Si des problèmes sont détectés, ils sont classés comme « Recommandé » ou « Obligatoire » pour être résolus. Les problèmes recommandés sont généralement des éléments qui ralentiraient l'exécution des tâches d'analyse et de catégorisation de la classification des données. Ces éléments n'ont pas besoin d'être corrigés, mais vous souhaiterez peut-être les corriger.

Si vous rencontrez des problèmes « obligatoires », vous devez les résoudre et exécuter à nouveau le script de test des prérequis.

Activer l'analyse sur vos sources de données

Analyser les sources de données avec la NetApp Data Classification

NetApp Data Classification analyse les données dans les référentiels (volumes, schémas de base de données ou autres données utilisateur) que vous sélectionnez pour identifier les données personnelles et sensibles. La classification des données cartographie ensuite vos données organisationnelles, catégorise chaque fichier et identifie des modèles prédéfinis dans les données. Le résultat de l'analyse est un index des informations personnelles, des informations personnelles sensibles, des catégories de données et des types de fichiers.

Après l'analyse initiale, Data Classification analyse en continu vos données de manière circulaire pour détecter les modifications incrémentielles. C'est pourquoi il est important de maintenir l'instance en cours d'exécution.

Vous pouvez activer et désactiver les analyses au niveau du volume ou au niveau du schéma de base de données.

Quelle est la différence entre les analyses de cartographie et de classification

Vous pouvez effectuer deux types d'analyses dans la classification des données :

- **Les analyses de cartographie uniquement** fournissent uniquement un aperçu de haut niveau de vos données et sont effectuées sur des sources de données sélectionnées. Les analyses de cartographie uniquement prennent moins de temps que les analyses de cartographie et de classification, car elles n'accèdent pas aux fichiers pour voir les données qu'ils contiennent. Vous souhaiterez peut-être procéder ainsi dans un premier temps pour identifier les domaines de recherche, puis effectuer une analyse de cartographie et de classification sur ces domaines.
- **Les analyses de cartographie et de classification** fournissent une analyse approfondie de vos données.

Le tableau ci-dessous montre certaines des différences :

Fonctionnalité	Cartographier et classer les scans	Analyses de cartographie uniquement
Vitesse de numérisation	Lent	Rapide
Tarifs	Gratuit	Gratuit
Capacité	Limité à 500 Tio*	Limité à 500 Tio*
Liste des types de fichiers et de la capacité utilisée	Oui	Oui
Nombre de fichiers et capacité utilisée	Oui	Oui
Âge et taille des fichiers	Oui	Oui
Capacité à exécuter un"Rapport de mappage des données"	Oui	Oui
Page d'enquête sur les données pour afficher les détails du fichier	Oui	Non
Rechercher des noms dans les fichiers	Oui	Non
Créer"requêtes enregistrées" qui fournissent des résultats de recherche personnalisés	Oui	Non
Possibilité d'exécuter d'autres rapports	Oui	Non
Possibilité de voir les métadonnées des fichiers**	Non	Oui

* La classification des données n'impose pas de limite à la quantité de données qu'elle peut analyser. Chaque agent de console prend en charge l'analyse et l'affichage de 500 Tio de données. Pour scanner plus de 500 Tio de données,"installer un autre agent de console" alors"déployer une autre instance de classification des données" . + L'interface utilisateur de la console affiche les données d'un seul connecteur. Pour obtenir des conseils sur l'affichage des données de plusieurs agents de console, consultez"Travailler avec plusieurs agents de console" .

** Les métadonnées suivantes sont extraites des fichiers lors des analyses de mappage :

- Système
- Type de système
- Référentiel de stockage
- Type de fichier

- Capacité utilisée
- Nombre de fichiers
- Taille du fichier
- Création de fichier
- Dernier accès au fichier
- Fichier modifié pour la dernière fois
- Heure de découverte du fichier
- Extraction des autorisations

Différences entre les tableaux de bord de gouvernance :

Fonctionnalité	Cartographier et classer	Carte
Données obsolètes	Oui	Oui
Données non commerciales	Oui	Oui
Fichiers dupliqués	Oui	Oui
Requêtes enregistrées prédéfinies	Oui	Non
Requêtes enregistrées par défaut	Oui	Oui
Rapport DDA	Oui	Oui
Rapport de cartographie	Oui	Oui
Détection du niveau de sensibilité	Oui	Non
Données sensibles avec des autorisations étendues	Oui	Non
Autorisations ouvertes	Oui	Oui
L'âge des données	Oui	Oui
Taille des données	Oui	Oui
Catégories	Oui	Non
Types de fichiers	Oui	Oui

Différences entre les tableaux de bord de conformité :

Fonctionnalité	Cartographier et classer	Carte
Informations personnelles	Oui	Non
Informations personnelles sensibles	Oui	Non
Rapport d'évaluation des risques liés à la vie privée	Oui	Non
Rapport HIPAA	Oui	Non
Rapport PCI DSS	Oui	Non

Différences entre les filtres d'investigation :

Fonctionnalité	Cartographier et classer	Carte
Requêtes enregistrées	Oui	Oui
Type de système	Oui	Oui
Système	Oui	Oui
Référentiel de stockage	Oui	Oui
Type de fichier	Oui	Oui
Taille du fichier	Oui	Oui
Temps de création	Oui	Oui
Temps découvert	Oui	Oui
Dernière modification	Oui	Oui
Dernier accès	Oui	Oui
Autorisations ouvertes	Oui	Oui
Chemin du répertoire de fichiers	Oui	Oui
Catégorie	Oui	Non
Niveau de sensibilité	Oui	Non
Nombre d'identifiants	Oui	Non
Données personnelles	Oui	Non
Données personnelles sensibles	Oui	Non
Personne concernée	Oui	Non
Doublons	Oui	Oui
Statut de classification	Oui	Le statut est toujours « Informations limitées »
Événement d'analyse d'analyse	Oui	Oui
Hachage de fichier	Oui	Oui
Nombre d'utilisateurs avec accès	Oui	Oui
Autorisations utilisateur/groupe	Oui	Oui
Propriétaire du fichier	Oui	Oui
Type de répertoire	Oui	Oui

À quelle vitesse Data Classification analyse-t-il les données ?

La vitesse d'analyse est affectée par la latence du réseau, la latence du disque, la bande passante du réseau, la taille de l'environnement et les tailles de distribution des fichiers.

- Lors de l'exécution d'analyses de cartographie uniquement, la classification des données peut analyser

entre 100 et 150 Tio de données par jour.

- Lors de l'exécution d'analyses de cartographie et de classification, Data Classification peut analyser entre 15 et 40 Tio de données par jour.

Analyser Amazon FSx pour les volumes ONTAP avec la NetApp Data Classification

Suivez quelques étapes pour commencer à analyser le volume Amazon FSx for ONTAP avec NetApp Data Classification.

Avant de commencer

- Vous avez besoin d'un agent de console actif dans AWS pour déployer et gérer la classification des données.
- Le groupe de sécurité que vous avez sélectionné lors de la création du système doit autoriser le trafic provenant de l'instance de classification des données. Vous pouvez trouver le groupe de sécurité associé à l'aide de l'ENI connecté au système de fichiers FSx for ONTAP et le modifier à l'aide de la console de gestion AWS.

  "Groupes de sécurité AWS pour les instances Linux"

  "Groupes de sécurité AWS pour les instances Windows"

  "Interfaces réseau élastiques AWS (ENI)"

- Assurez-vous que les ports suivants sont ouverts sur l'instance de classification des données :
  - Pour NFS – ports 111 et 2049.
  - Pour CIFS – ports 139 et 445.

Déployer l'instance de classification des données

"Déployer la classification des données"s'il n'y a pas déjà une instance déployée.

Vous devez déployer la classification des données sur le même réseau AWS que l'agent de console pour AWS et les volumes FSx que vous souhaitez analyser.

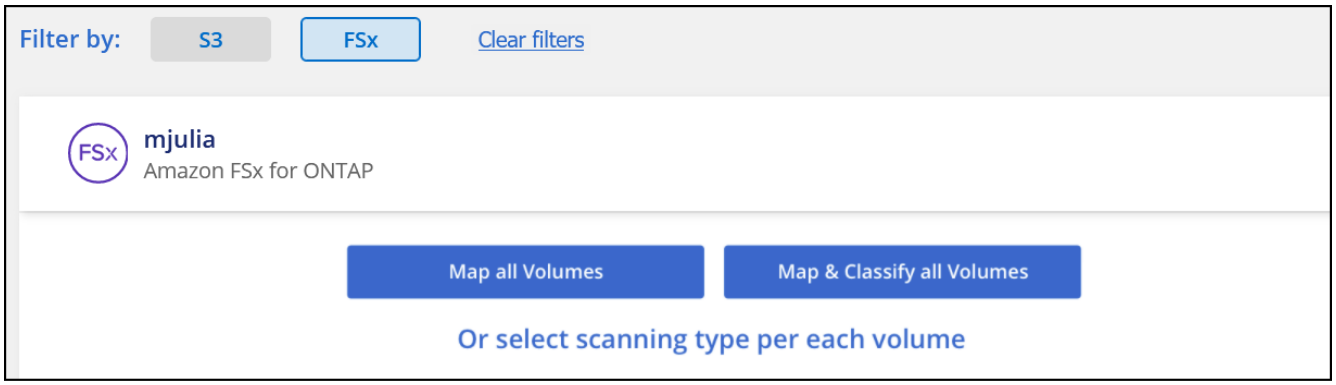
Remarque : le déploiement de la classification des données dans un emplacement local n'est actuellement pas pris en charge lors de l'analyse des volumes FSx.

Les mises à niveau du logiciel de classification des données sont automatisées tant que l'instance dispose d'une connexion Internet.

Activez la classification des données dans vos systèmes

Vous pouvez activer la classification des données pour les volumes FSx for ONTAP .

1. Depuis la NetApp Console, **Gouvernance > Classification**.
2. Dans le menu Classification des données, sélectionnez **Configuration**.



3. Sélectionnez la manière dont vous souhaitez analyser les volumes de chaque système. "En savoir plus sur les analyses de cartographie et de classification":
   - Pour mapper tous les volumes, sélectionnez **Mapper tous les volumes**.
   - Pour cartographier et classer tous les volumes, sélectionnez **Cartographier et classer tous les volumes**.
   - Pour personnaliser l'analyse de chaque volume, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**, puis choisissez les volumes que vous souhaitez mapper et/ou classer.
4. Dans la boîte de dialogue de confirmation, sélectionnez **Approuver** pour que la classification des données commence à analyser vos volumes.

Résultat

La classification des données commence à analyser les volumes que vous avez sélectionnés dans le système. Les résultats seront disponibles dans le tableau de bord de conformité dès que la classification des données aura terminé les analyses initiales. Le temps nécessaire dépend de la quantité de données : cela peut prendre quelques minutes ou quelques heures. Vous pouvez suivre la progression de l'analyse initiale en accédant au menu **Configuration** puis en sélectionnant la **Configuration système**. La progression de chaque analyse est affichée sous forme de barre de progression. Vous pouvez également survoler la barre de progression pour voir le nombre de fichiers analysés par rapport au nombre total de fichiers dans le volume.

- Par défaut, si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers de vos volumes car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**. La page résultante contient un paramètre que vous pouvez activer pour que la classification des données analyse les volumes quelles que soient les autorisations.
- La classification des données analyse un seul partage de fichiers sous un volume. Si vous avez plusieurs partages dans vos volumes, vous devrez analyser ces autres partages séparément en tant que groupe de partages. "Voir plus de détails sur cette limitation de classification des données" .

Vérifiez que la classification des données a accès aux volumes

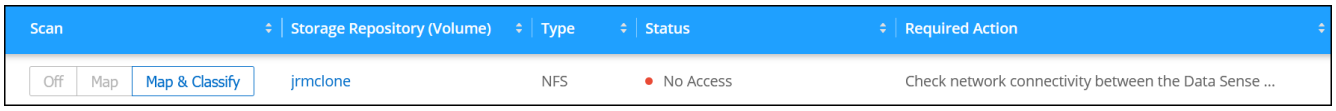
Assurez-vous que la classification des données peut accéder aux volumes en vérifiant votre réseau, vos groupes de sécurité et vos politiques d'exportation.

Vous devrez fournir à Data Classification les informations d'identification CIFS afin qu'il puisse accéder aux volumes CIFS.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Sur la page Configuration, sélectionnez **Afficher les détails** pour vérifier l'état et corriger les erreurs.

   Par exemple, l'image suivante montre un volume que Data Classification ne peut pas analyser en raison de problèmes de connectivité réseau entre l'instance Data Classification et le volume.



3. Assurez-vous qu'il existe une connexion réseau entre l'instance de classification des données et chaque réseau qui inclut des volumes pour FSx pour ONTAP.

   Pour FSx for ONTAP, la classification des données peut analyser les volumes uniquement dans la même région que la console.

4. Assurez-vous que les stratégies d'exportation de volume NFS incluent l'adresse IP de l'instance de classification des données afin qu'elle puisse accéder aux données sur chaque volume.
5. Si vous utilisez CIFS, fournissez à Data Classification les informations d'identification Active Directory afin qu'il puisse analyser les volumes CIFS.

   a. Dans le menu Classification des données, sélectionnez **Configuration**.
   b. Pour chaque système, sélectionnez **Modifier les informations d'identification CIFS** et saisissez le nom d'utilisateur et le mot de passe dont Data Classification a besoin pour accéder aux volumes CIFS sur le système.

      Les informations d'identification peuvent être en lecture seule, mais la fourniture d'informations d'identification d'administrateur garantit que la classification des données peut lire toutes les données nécessitant des autorisations élevées. Les informations d'identification sont stockées sur l'instance de classification des données.

      Si vous souhaitez vous assurer que les « dernières heures d'accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, il est recommandé que l'utilisateur dispose des autorisations d'écriture d'attributs dans CIFS ou des autorisations d'écriture dans NFS. Si possible, configurez l'utilisateur Active Directory en tant que membre d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

   Après avoir saisi les informations d'identification, vous devriez voir un message indiquant que tous les volumes CIFS ont été authentifiés avec succès.

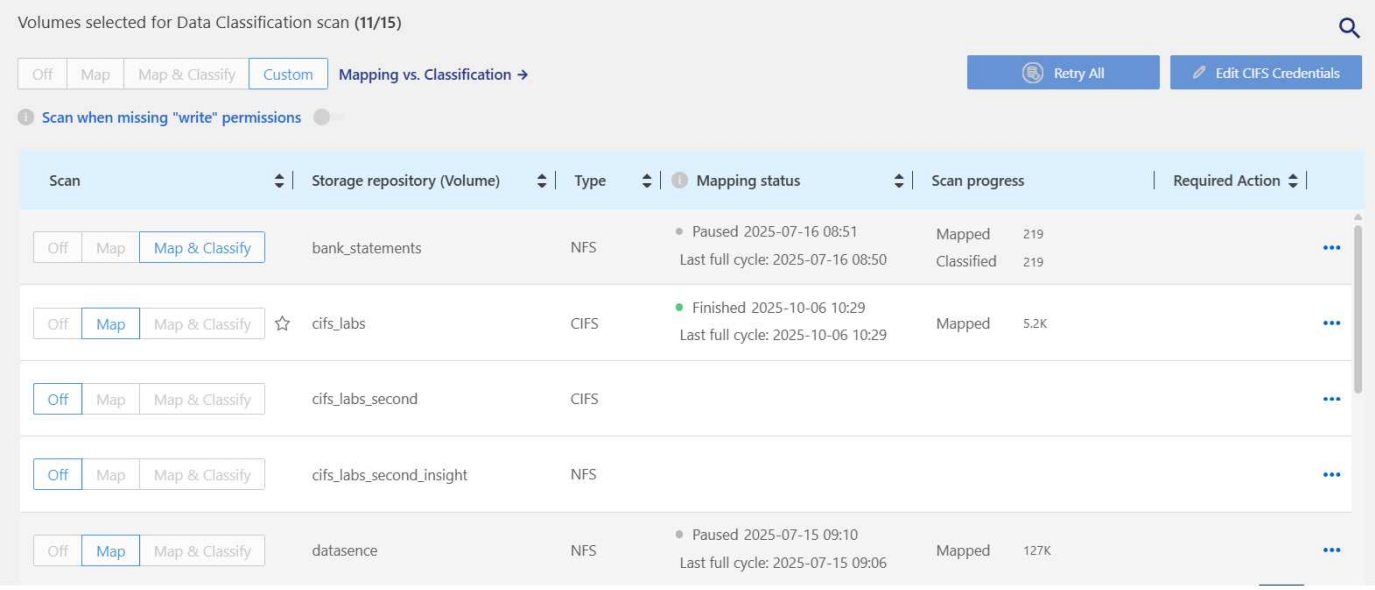
Activer et désactiver les analyses de conformité sur les volumes

Vous pouvez démarrer ou arrêter les analyses sur n'importe quel système à tout moment à partir de la page de configuration. Vous pouvez également passer d'analyses de cartographie uniquement à des analyses de cartographie et de classification, et vice-versa. Il est recommandé d'analyser tous les volumes d'un système.

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez sélectionné le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque cette option est définie sur **Personnalisé** ou **Désactivé** dans la zone d'en-tête, vous devrez activer le mappage et/ou l'analyse complète sur chaque nouveau volume que vous ajoutez au système.

Le commutateur en haut de la page pour **Analyser en cas d'absence d'autorisations « d'écriture »** est désactivé par défaut. Cela signifie que si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, activez l'interrupteur et tous les fichiers sont analysés quelles que soient les autorisations. "Apprendre encore plus".

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez défini le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque le paramètre pour tous les volumes est **Personnalisé** ou **Désactivé**, vous devez activer l'analyse manuellement pour chaque nouveau volume que vous ajoutez.



Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Choisissez un système, puis sélectionnez **Configuration**.
3. Pour activer ou désactiver les analyses pour tous les volumes, sélectionnez **Map**, **Map & Classify** ou **Off** dans l'en-tête au-dessus de tous les volumes.

   Pour activer ou désactiver les analyses de volumes individuels, recherchez les volumes dans la liste, puis sélectionnez **Map**, **Map & Classify** ou **Off** à côté du nom du volume.

Résultat

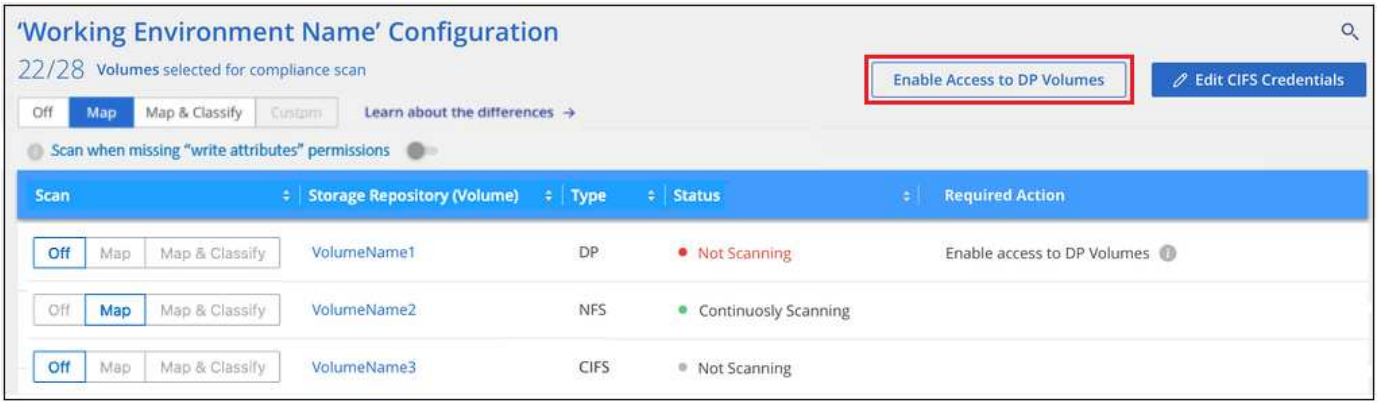
Lorsque vous activez l'analyse, la classification des données démarre l'analyse des volumes que vous avez sélectionnés dans le système. Les résultats commencent à apparaître dans le tableau de bord Conformité dès que la classification des données démarre l'analyse. Le temps d'exécution de l'analyse dépend de la quantité de données, allant de quelques minutes à quelques heures.

Analyser les volumes de protection des données

Par défaut, les volumes de protection des données (DP) ne sont pas analysés car ils ne sont pas exposés en externe et Data Classification ne peut pas y accéder. Il s'agit des volumes de destination pour les opérations SnapMirror à partir d'un système de fichiers FSx pour ONTAP .

Initialement, la liste des volumes identifie ces volumes comme *Type* **DP** avec le *Statut* **Pas d'analyse** et

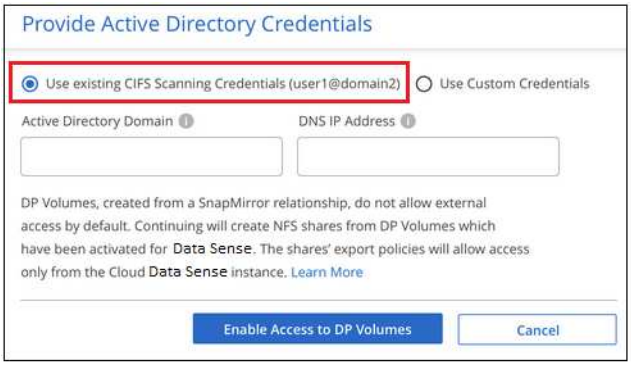
l'*Action requise* **Activer l'accès aux volumes DP**.

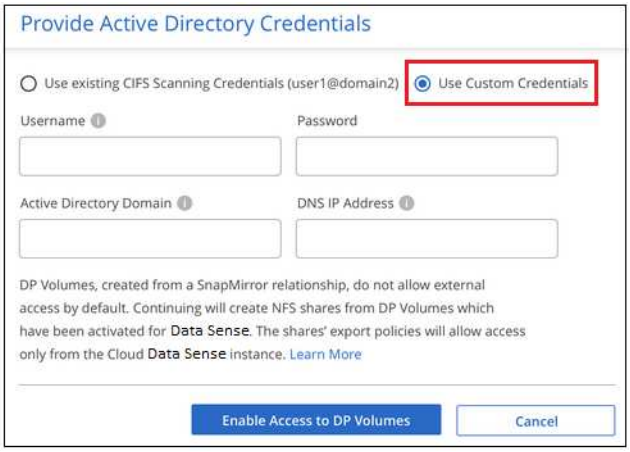


Étapes

Si vous souhaitez analyser ces volumes de protection des données :

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Sélectionnez **Activer l'accès aux volumes DP** en haut de la page.
3. Vérifiez le message de confirmation et sélectionnez à nouveau **Activer l'accès aux volumes DP**.
   - Les volumes initialement créés en tant que volumes NFS dans le système de fichiers source FSx pour ONTAP sont activés.
   - Les volumes initialement créés en tant que volumes CIFS dans le système de fichiers source FSx for ONTAP nécessitent que vous saisissiez les informations d'identification CIFS pour analyser ces volumes DP. Si vous avez déjà saisi les informations d'identification Active Directory pour que la classification des données puisse analyser les volumes CIFS, vous pouvez utiliser ces informations d'identification ou spécifier un autre ensemble d'informations d'identification d'administrateur.





4. Activez chaque volume DP que vous souhaitez analyser.

Résultat

Une fois activée, la classification des données crée un partage NFS à partir de chaque volume DP activé pour l'analyse. Les politiques d'exportation de partage autorisent uniquement l'accès à partir de l'instance de classification des données.

Si vous n'aviez aucun volume de protection des données CIFS lorsque vous avez initialement activé l'accès aux volumes DP, et que vous en avez ajouté ultérieurement, le bouton **Activer l'accès à CIFS DP** apparaît en haut de la page de configuration. Sélectionnez ce bouton et ajoutez les informations d'identification CIFS pour

activer l'accès à ces volumes CIFS DP.

> Les informations d'identification Active Directory sont enregistrées uniquement dans la machine virtuelle de stockage du premier volume DP CIFS. Par conséquent, tous les volumes DP sur cette SVM seront analysés. Tous les volumes résidant sur d'autres SVM n'auront pas les informations d'identification Active Directory enregistrées, de sorte que ces volumes DP ne seront pas analysés.

Analyser les volumes Azure NetApp Files avec la NetApp Data Classification

Suivez quelques étapes pour démarrer avec NetApp Data Classification pour Azure NetApp Files.

Découvrez le système Azure NetApp Files que vous souhaitez analyser

Si le système Azure NetApp Files que vous souhaitez analyser n'est pas déjà présent dans la NetApp Console en tant que système,"ajoutez-le dans la page Systèmes" .

Déployer l'instance de classification des données

"Déployer la classification des données"s'il n'y a pas déjà une instance déployée.

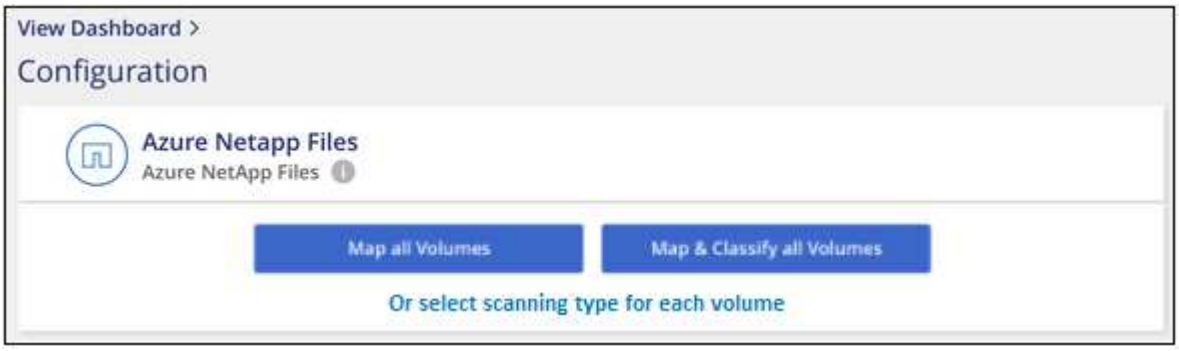
La classification des données doit être déployée dans le cloud lors de l'analyse des volumes Azure NetApp Files et doit être déployée dans la même région que les volumes que vous souhaitez analyser.

Remarque : le déploiement de la classification des données dans un emplacement local n'est actuellement pas pris en charge lors de l'analyse des volumes Azure NetApp Files .

Activez la classification des données dans vos systèmes

Vous pouvez activer la classification des données sur vos volumes Azure NetApp Files .

1. Dans le menu Classification des données, sélectionnez **Configuration**.



2. Sélectionnez la manière dont vous souhaitez analyser les volumes de chaque système. "En savoir plus sur les analyses de cartographie et de classification":
   - Pour mapper tous les volumes, sélectionnez **Mapper tous les volumes**.
   - Pour cartographier et classer tous les volumes, sélectionnez **Cartographier et classer tous les volumes**.
   - Pour personnaliser l'analyse de chaque volume, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**, puis choisissez les volumes que vous souhaitez mapper ou mapper et classer.

VoirActiver ou désactiver les analyses de conformité sur les volumes pour plus de détails.

3. Dans la boîte de dialogue de confirmation, sélectionnez **Approuver**.

Résultat

La classification des données commence à analyser les volumes que vous avez sélectionnés dans le système. Les résultats sont disponibles dans le tableau de bord Conformité dès que la classification des données termine les analyses initiales. Le temps nécessaire dépend de la quantité de données : cela peut prendre quelques minutes ou quelques heures. Vous pouvez suivre la progression de l'analyse initiale en accédant au menu **Configuration** puis en sélectionnant la **Configuration système**. La classification des données affiche une barre de progression pour chaque analyse. Vous pouvez survoler la barre de progression pour voir le nombre de fichiers analysés par rapport au nombre total de fichiers dans le volume.

- Par défaut, si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers de vos volumes car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**. La page résultante contient un paramètre que vous pouvez activer pour que la classification des données analyse les volumes quelles que soient les autorisations.
- La classification des données analyse un seul partage de fichiers sous un volume. Si vous avez plusieurs partages dans vos volumes, vous devrez analyser ces autres partages séparément en tant que groupe de partages. "En savoir plus sur cette limitation de classification des données" .

Vérifiez que la classification des données a accès aux volumes

Assurez-vous que la classification des données peut accéder aux volumes en vérifiant votre réseau, vos groupes de sécurité et vos politiques d'exportation. Vous devez fournir à Data Classification les informations d'identification CIFS afin qu'elle puisse accéder aux volumes CIFS.

Pour Azure NetApp Files, la classification des données ne peut analyser que les volumes dans la même région que la console.

Liste de contrôle

- Assurez-vous qu'il existe une connexion réseau entre l'instance de classification des données et chaque réseau qui inclut des volumes pour Azure NetApp Files.
- Assurez-vous que les ports suivants sont ouverts sur l'instance de classification des données :
  - Pour NFS – ports 111 et 2049.
  - Pour CIFS – ports 139 et 445.
- Assurez-vous que les stratégies d'exportation de volume NFS incluent l'adresse IP de l'instance de classification des données afin qu'elle puisse accéder aux données sur chaque volume.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.

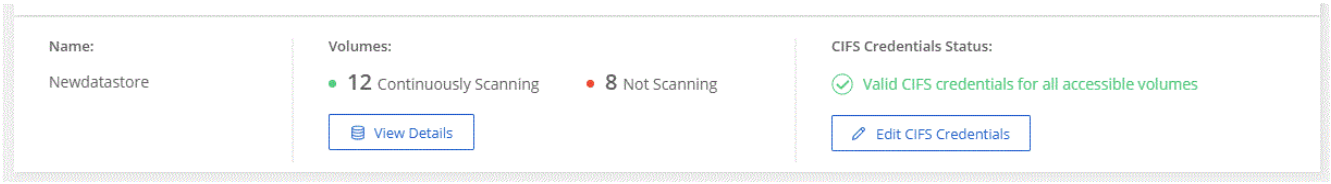
   a. Si vous utilisez CIFS (SMB), assurez-vous que les informations d'identification Active Directory sont correctes. Pour chaque système, sélectionnez **Modifier les informations d'identification CIFS**, puis saisissez le nom d'utilisateur et le mot de passe dont Data Classification a besoin pour accéder aux volumes CIFS sur le système.

      Les informations d'identification peuvent être en lecture seule ; la fourniture d'informations d'identification d'administrateur garantit que Data Classification peut lire toutes les données nécessitant

des autorisations élevées. Les informations d'identification sont stockées sur l'instance de classification des données.

Si vous souhaitez vous assurer que les « dernières heures d'accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, il est recommandé que l'utilisateur dispose des autorisations d'écriture d'attributs dans CIFS ou des autorisations d'écriture dans NFS. Si possible, configurez l'utilisateur Active Directory en tant que membre d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

Après avoir saisi les informations d'identification, vous devriez voir un message indiquant que tous les volumes CIFS ont été authentifiés avec succès.



2. Sur la page Configuration, sélectionnez **Afficher les détails** pour vérifier l'état de chaque volume CIFS et NFS. Si nécessaire, corrigez les erreurs telles que les problèmes de connectivité réseau.

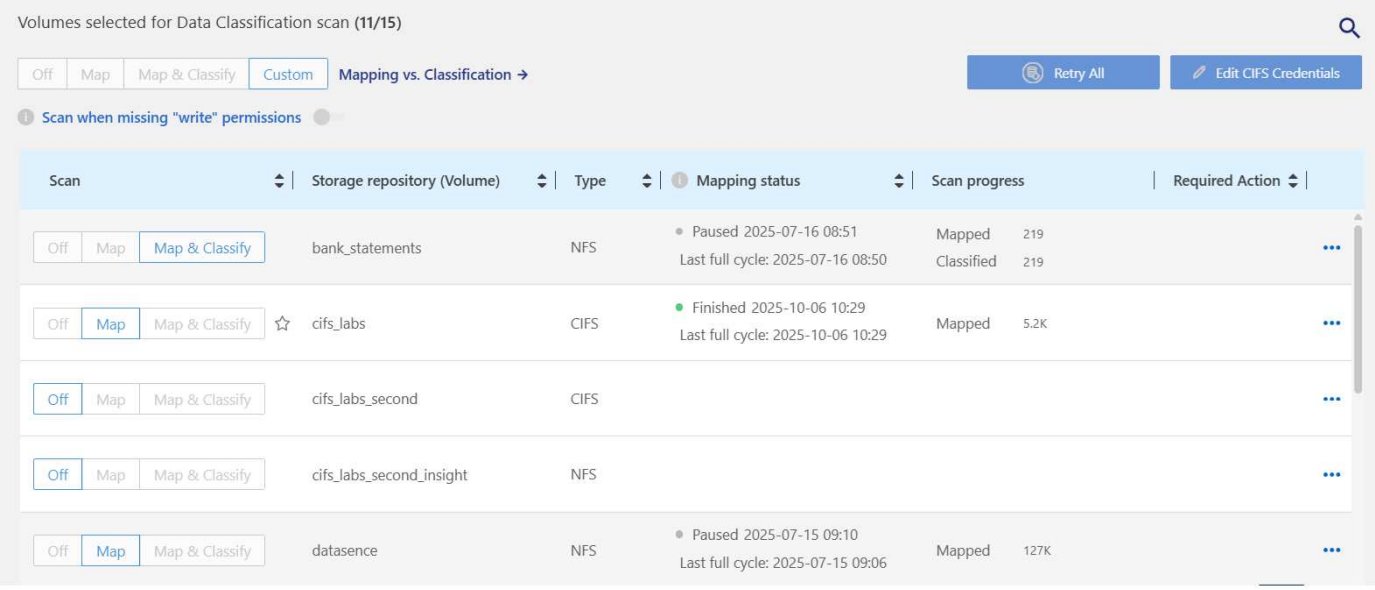
Activer ou désactiver les analyses de conformité sur les volumes

Vous pouvez démarrer ou arrêter les analyses sur n'importe quel système à tout moment à partir de la page de configuration. Vous pouvez également passer d'analyses de cartographie uniquement à des analyses de cartographie et de classification, et vice-versa. Il est recommandé d'analyser tous les volumes d'un système.

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez sélectionné le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque cette option est définie sur **Personnalisé** ou **Désactivé** dans la zone d'en-tête, vous devrez activer le mappage et/ou l'analyse complète sur chaque nouveau volume que vous ajoutez au système.

Le commutateur en haut de la page pour **Analyser en cas d'absence d'autorisations « d'écriture »** est désactivé par défaut. Cela signifie que si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, activez l'interrupteur et tous les fichiers sont analysés quelles que soient les autorisations. "Apprendre encore plus".

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez défini le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque le paramètre pour tous les volumes est **Personnalisé** ou **Désactivé**, vous devez activer l'analyse manuellement pour chaque nouveau volume que vous ajoutez.



Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Choisissez un système, puis sélectionnez **Configuration**.
3. Pour activer ou désactiver les analyses pour tous les volumes, sélectionnez **Map**, **Map & Classify** ou **Off** dans l'en-tête au-dessus de tous les volumes.

   Pour activer ou désactiver les analyses de volumes individuels, recherchez les volumes dans la liste, puis sélectionnez **Map**, **Map & Classify** ou **Off** à côté du nom du volume.

Résultat

Lorsque vous activez l'analyse, la classification des données démarre l'analyse des volumes que vous avez sélectionnés dans le système. Les résultats commencent à apparaître dans le tableau de bord Conformité dès que la classification des données démarre l'analyse. Le temps d'exécution de l'analyse dépend de la quantité de données, allant de quelques minutes à quelques heures.

Analysez les Cloud Volumes ONTAP et les volumes ONTAP sur site avec la NetApp Data Classification

Suivez quelques étapes pour commencer à analyser vos Cloud Volumes ONTAP et vos volumes ONTAP sur site à l'aide de NetApp Data Classification.

Prérequis

Avant d'activer la classification des données, assurez-vous que vous disposez d'une configuration prise en charge.

- Si vous numérisez des Cloud Volumes ONTAP et des systèmes ONTAP sur site accessibles via Internet, vous pouvez"déployer la classification des données dans le cloud" ou"dans un local équipé d'un accès Internet" .
- Si vous analysez des systèmes ONTAP sur site qui ont été installés sur un site sombre sans accès Internet, vous devez"déployer la classification des données dans le même emplacement sur site qui n'a pas d'accès Internet" . Cela nécessite que l'agent de console soit déployé dans le même emplacement sur site.

Vérifiez que la classification des données a accès aux volumes

Assurez-vous que la classification des données peut accéder aux volumes en vérifiant votre réseau, vos groupes de sécurité et vos politiques d'exportation. Vous devrez fournir à Data Classification les informations d'identification CIFS afin qu'il puisse accéder aux volumes CIFS.

Liste de contrôle

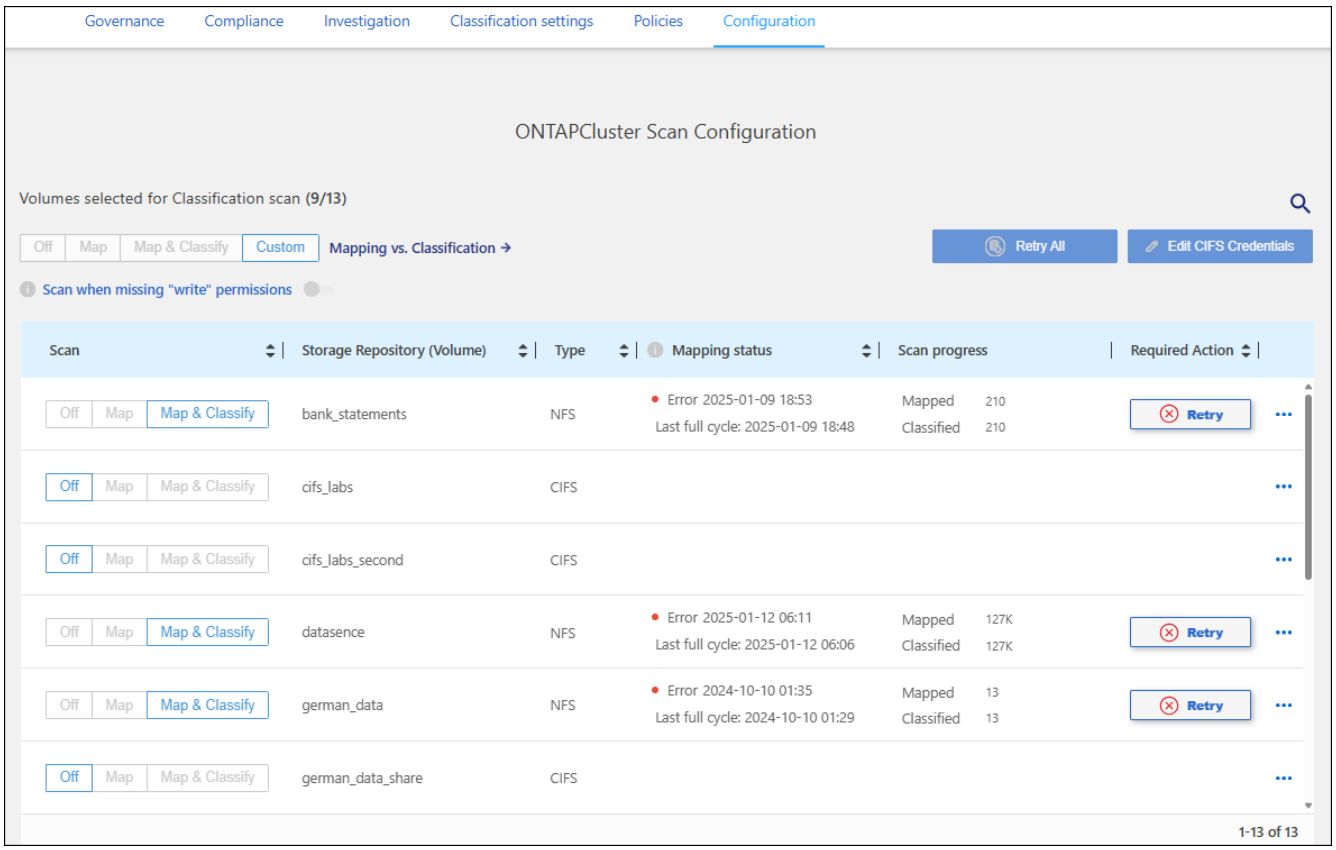
- Assurez-vous qu'il existe une connexion réseau entre l'instance de classification des données et chaque réseau qui inclut des volumes pour les clusters Cloud Volumes ONTAP ou ONTAP sur site.
- Assurez-vous que le groupe de sécurité pour Cloud Volumes ONTAP autorise le trafic entrant depuis l'instance de classification des données.

  Vous pouvez soit ouvrir le groupe de sécurité pour le trafic provenant de l'adresse IP de l'instance de classification des données, soit ouvrir le groupe de sécurité pour tout le trafic provenant de l'intérieur du réseau virtuel.

- Assurez-vous que les stratégies d'exportation de volume NFS incluent l'adresse IP de l'instance de classification des données afin qu'elle puisse accéder aux données sur chaque volume.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.



2. Si vous utilisez CIFS, fournissez à Data Classification les informations d'identification Active Directory afin qu'il puisse analyser les volumes CIFS. Pour chaque système, sélectionnez **Modifier les informations d'identification CIFS** et saisissez le nom d'utilisateur et le mot de passe dont Data Classification a besoin pour accéder aux volumes CIFS sur le système.

   Les informations d'identification peuvent être en lecture seule, mais la fourniture d'informations d'identification d'administrateur garantit que la classification des données peut lire toutes les données

nécessitant des autorisations élevées. Les informations d'identification sont stockées sur l'instance de classification des données.

Si vous souhaitez vous assurer que les « dernières heures d'accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, il est recommandé que l'utilisateur dispose des autorisations d'écriture d'attributs dans CIFS ou des autorisations d'écriture dans NFS. Si possible, configurez l'utilisateur Active Directory en tant que membre d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

Si vous avez correctement saisi les informations d'identification, un message confirme que tous les volumes CIFS ont été authentifiés avec succès.

3. Sur la page Configuration, sélectionnez **Configuration** pour vérifier l'état de chaque volume CIFS et NFS et corriger les erreurs éventuelles.

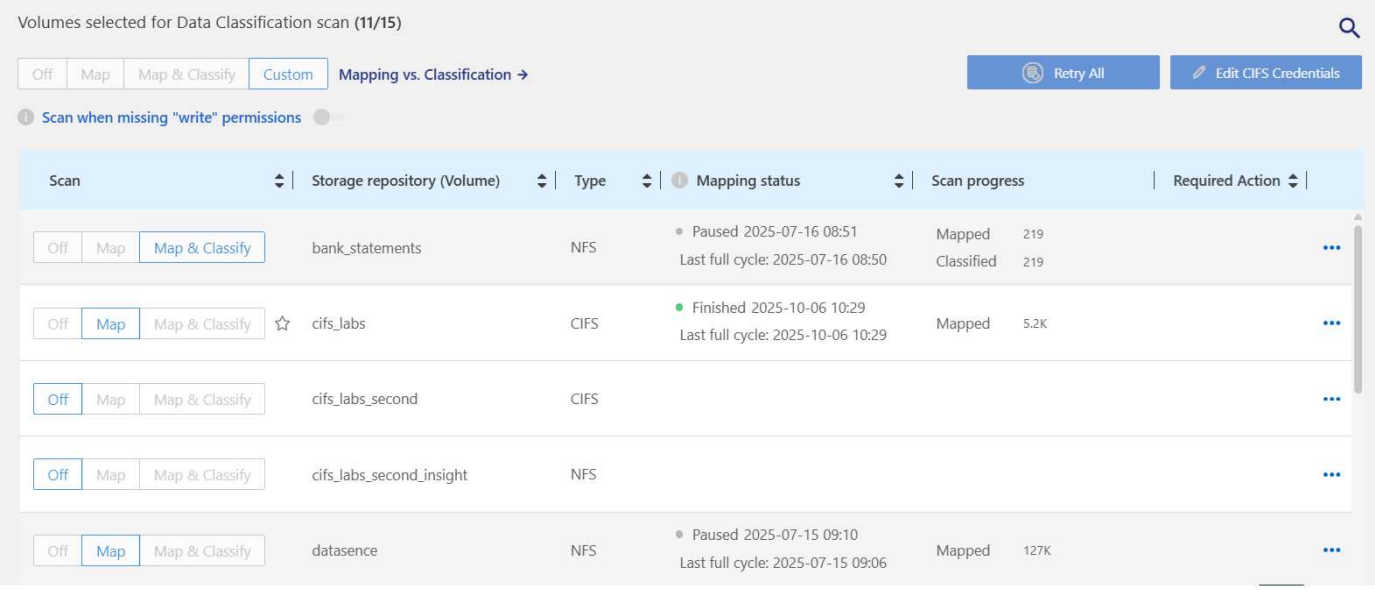
Activer ou désactiver les analyses de conformité sur les volumes

Vous pouvez démarrer ou arrêter les analyses sur n'importe quel système à tout moment à partir de la page de configuration. Vous pouvez également passer d'analyses de cartographie uniquement à des analyses de cartographie et de classification, et vice-versa. Il est recommandé d'analyser tous les volumes d'un système.

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez sélectionné le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque cette option est définie sur **Personnalisé** ou **Désactivé** dans la zone d'en-tête, vous devrez activer le mappage et/ou l'analyse complète sur chaque nouveau volume que vous ajoutez au système.

Le commutateur en haut de la page pour **Analyser en cas d'absence d'autorisations « d'écriture »** est désactivé par défaut. Cela signifie que si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, activez l'interrupteur et tous les fichiers sont analysés quelles que soient les autorisations. "Apprendre encore plus".

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez défini le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque le paramètre pour tous les volumes est **Personnalisé** ou **Désactivé**, vous devez activer l'analyse manuellement pour chaque nouveau volume que vous ajoutez.



Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Choisissez un système, puis sélectionnez **Configuration**.
3. Pour activer ou désactiver les analyses pour tous les volumes, sélectionnez **Map**, **Map & Classify** ou **Off** dans l'en-tête au-dessus de tous les volumes.

   Pour activer ou désactiver les analyses de volumes individuels, recherchez les volumes dans la liste, puis sélectionnez **Map**, **Map & Classify** ou **Off** à côté du nom du volume.

Résultat

Lorsque vous activez l'analyse, la classification des données démarre l'analyse des volumes que vous avez sélectionnés dans le système. Les résultats commencent à apparaître dans le tableau de bord Conformité dès que la classification des données démarre l'analyse. Le temps d'exécution de l'analyse dépend de la quantité de données, allant de quelques minutes à quelques heures.

La classification des données analyse un seul partage de fichiers sous un volume. Si vous avez plusieurs partages dans vos volumes, vous devrez analyser ces autres partages séparément en tant que groupe de partages. "Voir plus de détails sur cette limitation de classification des données" .

Analyser les schémas de base de données avec la NetApp Data Classification

Suivez quelques étapes pour commencer à analyser vos schémas de base de données avec NetApp Data Classification.

Réviser les prérequis

Passez en revue les conditions préalables suivantes pour vous assurer que vous disposez d'une configuration prise en charge avant d'activer la classification des données.

Bases de données prises en charge

La classification des données peut analyser les schémas des bases de données suivantes :

- Service de base de données relationnelle Amazon (Amazon RDS)
- MongoDB
- MySQL
- Oracle
- PostgreSQL
- SAP HANA
- Serveur SQL (MSSQL)

La fonctionnalité de collecte de statistiques **doit être activée** dans la base de données.

Exigences relatives à la base de données

Toute base de données connectée à l'instance de classification des données peut être analysée, quel que soit l'endroit où elle est hébergée. Vous avez juste besoin des informations suivantes pour vous connecter à la base de données :

- Adresse IP ou nom d'hôte
- Port
- Nom du service (uniquement pour l'accès aux bases de données Oracle)
- Informations d'identification permettant l'accès en lecture aux schémas

  Lors du choix d'un nom d'utilisateur et d'un mot de passe, il est important d'en choisir un qui dispose de toutes les autorisations de lecture sur tous les schémas et tables que vous souhaitez analyser. Nous vous recommandons de créer un utilisateur dédié au système de classification des données avec toutes les autorisations requises.

Pour MongoDB, un rôle d'administrateur en lecture seule est requis.

Déployer l'instance de classification des données

Déployez la classification des données s'il n'existe pas déjà d'instance déployée.

Si vous numérisez des schémas de bases de données accessibles sur Internet, vous pouvez"déployer la classification des données dans le cloud" ou"déployer la classification des données dans un emplacement sur site disposant d'un accès Internet" .

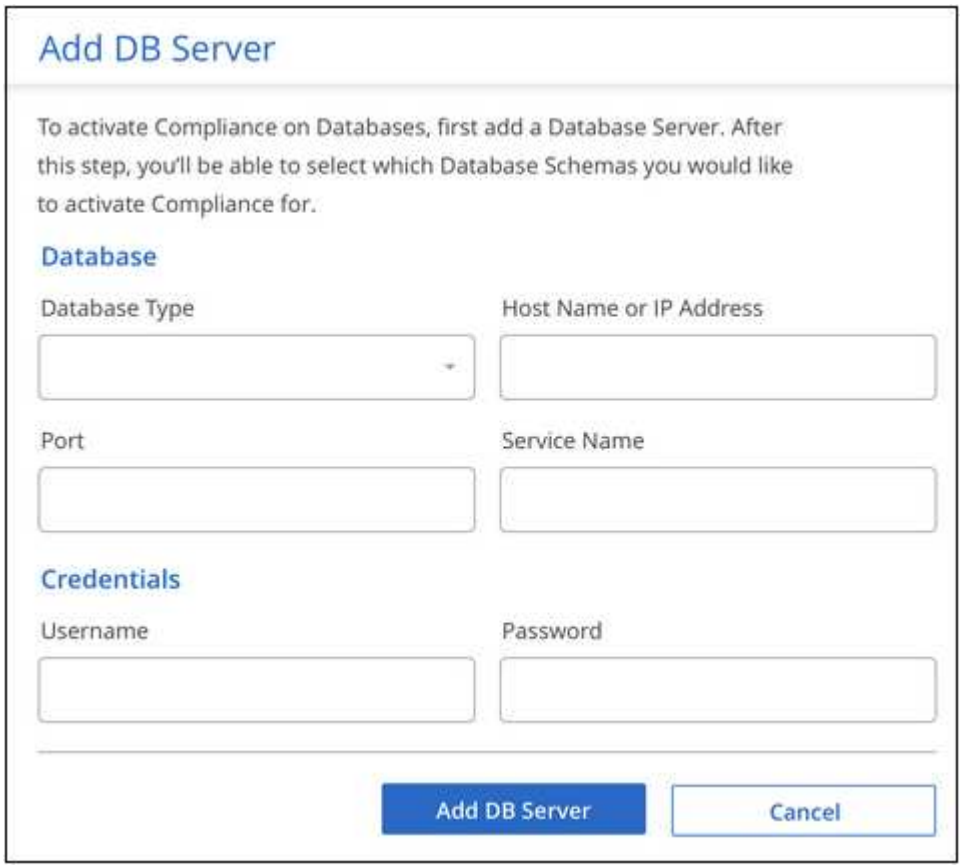
Si vous numérisez des schémas de base de données qui ont été installés sur un site sombre qui n'a pas d'accès Internet, vous devez"déployer la classification des données dans le même emplacement sur site qui n'a pas d'accès Internet" . Cela nécessite également que l'agent de console soit déployé dans le même emplacement sur site.

Ajouter le serveur de base de données

Ajoutez le serveur de base de données sur lequel résident les schémas.

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Depuis la page Configuration, sélectionnez **Ajouter un système** > **Ajouter un serveur de base de données**.

3. Saisissez les informations requises pour identifier le serveur de base de données.
   a. Sélectionnez le type de base de données.
   b. Entrez le port et le nom d'hôte ou l'adresse IP pour vous connecter à la base de données.
   c. Pour les bases de données Oracle, entrez le nom du service.
   d. Saisissez les informations d'identification pour que Data Classification puisse accéder au serveur.
   e. Sélectionnez **Ajouter un serveur de base de données**.



La base de données est ajoutée à la liste des systèmes.

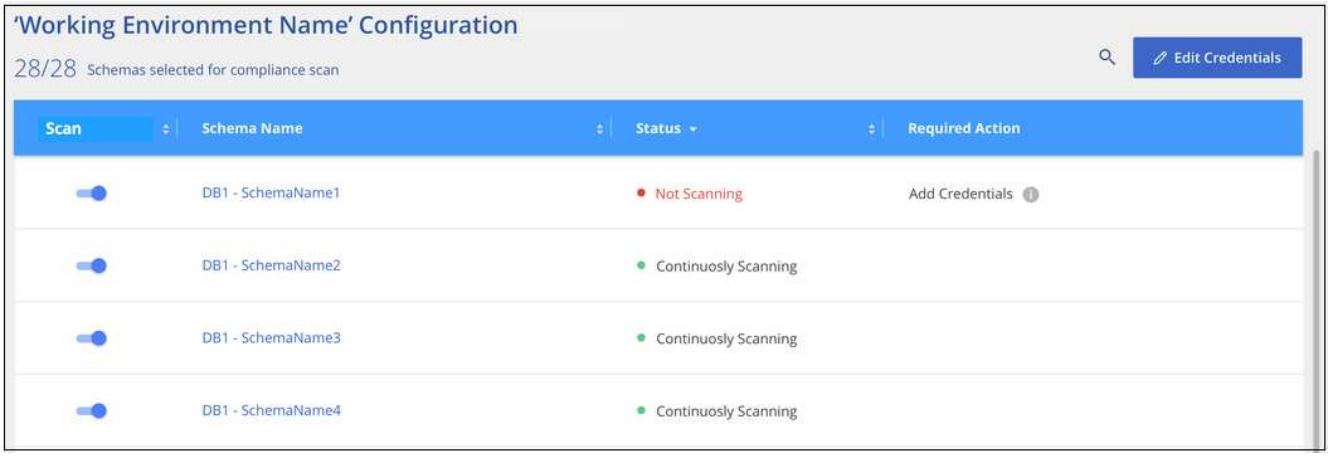
Activer et désactiver les analyses de conformité sur les schémas de base de données

Vous pouvez arrêter ou démarrer l'analyse complète de vos schémas à tout moment.

Il n'existe aucune option permettant de sélectionner des analyses de mappage uniquement pour les schémas de base de données.

1. Depuis la page Configuration, sélectionnez le bouton **Configuration** correspondant à la base de données que vous souhaitez configurer.

2. Sélectionnez les schémas que vous souhaitez analyser en déplaçant le curseur vers la droite.



Résultat

La classification des données commence à analyser les schémas de base de données que vous avez activés. Vous pouvez suivre la progression de l'analyse initiale en accédant au menu **Configuration** puis en sélectionnant la **Configuration système**. La progression de chaque analyse est affichée sous forme de barre de progression. Vous pouvez également survoler la barre de progression pour voir le nombre de fichiers analysés par rapport au nombre total de fichiers dans le volume. S'il y a des erreurs, elles apparaîtront dans la colonne Statut, à côté de l'action requise pour corriger l'erreur.

Data Classification analyse vos bases de données une fois par jour ; les bases de données ne sont pas analysées en continu comme les autres sources de données.

Analyser les Google Cloud NetApp Volumes avec la NetApp Data Classification

NetApp Data Classification prend en charge Google Cloud NetApp Volumes en tant que système. Découvrez comment analyser votre système Google Cloud NetApp Volumes .

Découvrez le système Google Cloud NetApp Volumes que vous souhaitez analyser

Si le système Google Cloud NetApp Volumes que vous souhaitez analyser n'est pas déjà présent dans la NetApp Console en tant que système,"ajoutez-le à la page Systèmes" .

Déployer l'instance de classification des données

"Déployer la classification des données"s'il n'y a pas déjà une instance déployée.

La classification des données doit être déployée dans le cloud lors de l'analyse des Google Cloud NetApp

Volumes et doit être déployée dans la même région que les volumes que vous souhaitez analyser.

Remarque : le déploiement de la classification des données dans un emplacement sur site n'est actuellement pas pris en charge lors de l'analyse des Google Cloud NetApp Volumes.

Activez la classification des données dans vos systèmes

Vous pouvez activer la classification des données sur votre système Google Cloud NetApp Volumes .

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Sélectionnez la manière dont vous souhaitez analyser les volumes de chaque système. "En savoir plus sur les analyses de cartographie et de classification":
   - Pour mapper tous les volumes, sélectionnez **Mapper tous les volumes**.
   - Pour cartographier et classer tous les volumes, sélectionnez **Cartographier et classer tous les volumes**.
   - Pour personnaliser l'analyse de chaque volume, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**, puis choisissez les volumes que vous souhaitez mapper et/ou classer.

     VoirActiver et désactiver les analyses de conformité sur les volumes pour plus de détails.
3. Dans la boîte de dialogue de confirmation, sélectionnez **Approuver**.

Résultat

La classification des données commence à analyser les volumes que vous avez sélectionnés dans le système. Les résultats sont disponibles dans le tableau de bord Conformité dès que la classification des données termine les analyses initiales. Le temps nécessaire dépend de la quantité de données : de quelques minutes à quelques heures. Vous pouvez suivre la progression de l'analyse initiale dans la section **Configuration système** du menu **Configuration**. La classification des données affiche une barre de progression pour chaque analyse. Vous pouvez également survoler la barre de progression pour voir le nombre de fichiers analysés par rapport au nombre total de fichiers dans le volume.

- Par défaut, si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers de vos volumes car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, sélectionnez **Ou sélectionnez le type d'analyse pour chaque volume**. La page résultante contient un paramètre que vous pouvez activer pour que la classification des données analyse les volumes quelles que soient les autorisations.
- La classification des données analyse un seul partage de fichiers sous un volume. Si vous avez plusieurs partages dans vos volumes, vous devez analyser ces autres partages séparément en tant que groupe de partages. "En savoir plus sur cette limitation de classification des données".

Vérifiez que la classification des données a accès aux volumes

Assurez-vous que la classification des données peut accéder aux volumes en vérifiant votre réseau, vos groupes de sécurité et vos politiques d'exportation. Pour les volumes CIFS, vous devez fournir une classification des données avec les informations d'identification CIFS.

Pour les Google Cloud NetApp Volumes, la classification des données ne peut analyser que les volumes situés dans la même région que la console.

Liste de contrôle

- Assurez-vous qu'il existe une connexion réseau entre l'instance de classification des données et chaque réseau qui inclut des volumes pour Google Cloud NetApp Volumes.
- Assurez-vous que les ports suivants sont ouverts sur l'instance de classification des données :
  - Pour NFS – ports 111 et 2049.
  - Pour CIFS – ports 139 et 445.
- Assurez-vous que les stratégies d'exportation de volume NFS incluent l'adresse IP de l'instance de classification des données afin qu'elle puisse accéder aux données sur chaque volume.

Étapes

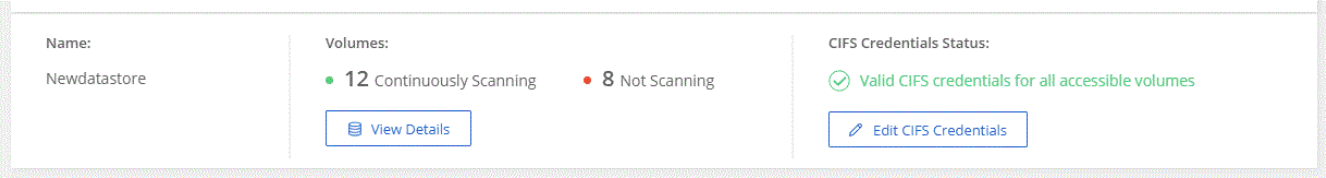
1. Dans le menu Classification des données, sélectionnez **Configuration**.

   a. Si vous utilisez CIFS (SMB), assurez-vous que les informations d'identification Active Directory sont correctes. Pour chaque système, sélectionnez **Modifier les informations d'identification CIFS**, puis saisissez le nom d'utilisateur et le mot de passe dont Data Classification a besoin pour accéder aux volumes CIFS sur le système.

      Les informations d'identification peuvent être en lecture seule, mais la fourniture d'informations d'identification d'administrateur garantit que la classification des données peut lire toutes les données nécessitant des autorisations élevées. Les informations d'identification sont stockées sur l'instance de classification des données.

      Si vous souhaitez vous assurer que les « dernières heures d'accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, il est recommandé que l'utilisateur dispose des autorisations d'écriture d'attributs dans CIFS ou des autorisations d'écriture dans NFS. Si possible, configurez l'utilisateur Active Directory en tant que membre d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

      Après avoir saisi les informations d'identification, vous devriez voir un message indiquant que tous les volumes CIFS ont été authentifiés avec succès.



2. Sur la page Configuration, sélectionnez **Afficher les détails** pour vérifier l'état de chaque volume CIFS et NFS et corriger les erreurs éventuelles.

Activer et désactiver les analyses de conformité sur les volumes

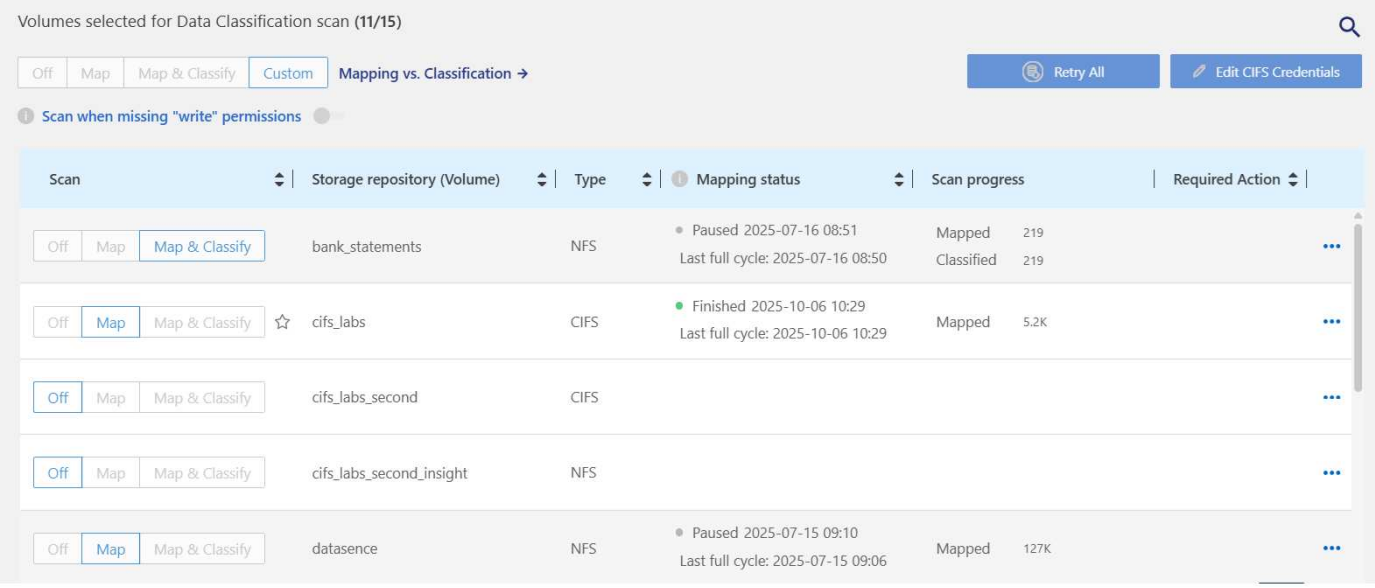
Vous pouvez démarrer ou arrêter les analyses sur n'importe quel système à tout moment à partir de la page de configuration. Vous pouvez également passer d'analyses de cartographie uniquement à des analyses de cartographie et de classification, et vice-versa. Il est recommandé d'analyser tous les volumes d'un système.

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez sélectionné le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque cette option est définie sur **Personnalisé** ou **Désactivé** dans la zone d'en-tête, vous devrez activer le mappage et/ou l'analyse complète sur chaque nouveau volume que vous ajoutez au système.

Le commutateur en haut de la page pour **Analyser en cas d'absence d'autorisations « d'écriture »** est

désactivé par défaut. Cela signifie que si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Si vous ne vous souciez pas de savoir si la dernière heure d'accès est réinitialisée, activez l'interrupteur et tous les fichiers sont analysés quelles que soient les autorisations. "Apprendre encore plus".

Les nouveaux volumes ajoutés au système sont automatiquement analysés uniquement lorsque vous avez défini le paramètre **Carte** ou **Carte et classification** dans la zone d'en-tête. Lorsque le paramètre pour tous les volumes est **Personnalisé** ou **Désactivé**, vous devez activer l'analyse manuellement pour chaque nouveau volume que vous ajoutez.



Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Choisissez un système, puis sélectionnez **Configuration**.
3. Pour activer ou désactiver les analyses pour tous les volumes, sélectionnez **Map**, **Map & Classify** ou **Off** dans l'en-tête au-dessus de tous les volumes.

   Pour activer ou désactiver les analyses de volumes individuels, recherchez les volumes dans la liste, puis sélectionnez **Map**, **Map & Classify** ou **Off** à côté du nom du volume.

Résultat

Lorsque vous activez l'analyse, la classification des données démarre l'analyse des volumes que vous avez sélectionnés dans le système. Les résultats commencent à apparaître dans le tableau de bord Conformité dès que la classification des données démarre l'analyse. Le temps d'exécution de l'analyse dépend de la quantité de données, allant de quelques minutes à quelques heures.

Analyser les partages de fichiers avec la NetApp Data Classification

Pour analyser les partages de fichiers, vous devez d'abord créer un groupe de partages de fichiers dans NetApp Data Classification. Les groupes de partages de fichiers sont destinés aux partages NFS ou CIFS (SMB) hébergés sur site ou dans le cloud.

L'analyse des données provenant de partages de fichiers non NetApp n'est pas prise en charge dans la version principale de la classification des données.

Prérequis

Passez en revue les conditions préalables suivantes pour vous assurer que vous disposez d'une configuration prise en charge avant d'activer la classification des données.

- Les actions peuvent être hébergées n'importe où, y compris dans le cloud ou sur site. Les partages CIFS des anciens systèmes de stockage NetApp 7-Mode peuvent être analysés en tant que partages de fichiers.
  - La classification des données ne peut pas extraire les autorisations ou la « dernière heure d'accès » des systèmes 7-Mode.
  - En raison d'un problème connu entre certaines versions Linux et les partages CIFS sur les systèmes 7-Mode, vous devez configurer le partage pour utiliser uniquement SMBv1 avec l'authentification NTLM activée.
- Une connectivité réseau est nécessaire entre l'instance de classification des données et les partages.
- Vous pouvez ajouter un partage DFS (Distributed File System) en tant que partage CIFS standard. Étant donné que la classification des données ne sait pas que le partage est basé sur plusieurs serveurs/volumes combinés en un seul partage CIFS, vous risquez de recevoir des erreurs d'autorisation ou de connectivité concernant le partage lorsque le message s'applique réellement uniquement à l'un des dossiers/partages situés sur un serveur/volume différent.
- Pour les partages CIFS (SMB), assurez-vous que vous disposez des informations d'identification Active Directory qui fournissent un accès en lecture aux partages. Les informations d'identification d'administrateur sont préférables au cas où la classification des données doit analyser des données nécessitant des autorisations élevées.

  Si vous souhaitez vous assurer que les « dernières heures d'accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, il est recommandé que l'utilisateur dispose des autorisations d'écriture d'attributs dans CIFS ou des autorisations d'écriture dans NFS. Si possible, configurez l'utilisateur Active Directory en tant que membre d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

- Tous les partages de fichiers CIFS d'un groupe doivent utiliser les mêmes informations d'identification Active Directory.
- Vous pouvez mélanger les partages NFS et CIFS (en utilisant Kerberos ou NTLM). Vous devez ajouter les actions au groupe séparément. Autrement dit, vous devez effectuer le processus deux fois, une fois par protocole.
  - Vous ne pouvez pas créer un groupe de partage de fichiers qui mélange les types d'authentification CIFS (Kerberos et NTLM).
- Si vous utilisez CIFS avec l'authentification Kerberos, assurez-vous que l'adresse IP fournie est accessible à la classification des données. Les partages de fichiers ne peuvent pas être ajoutés si l'adresse IP est inaccessible.

Créer un groupe de partage de fichiers

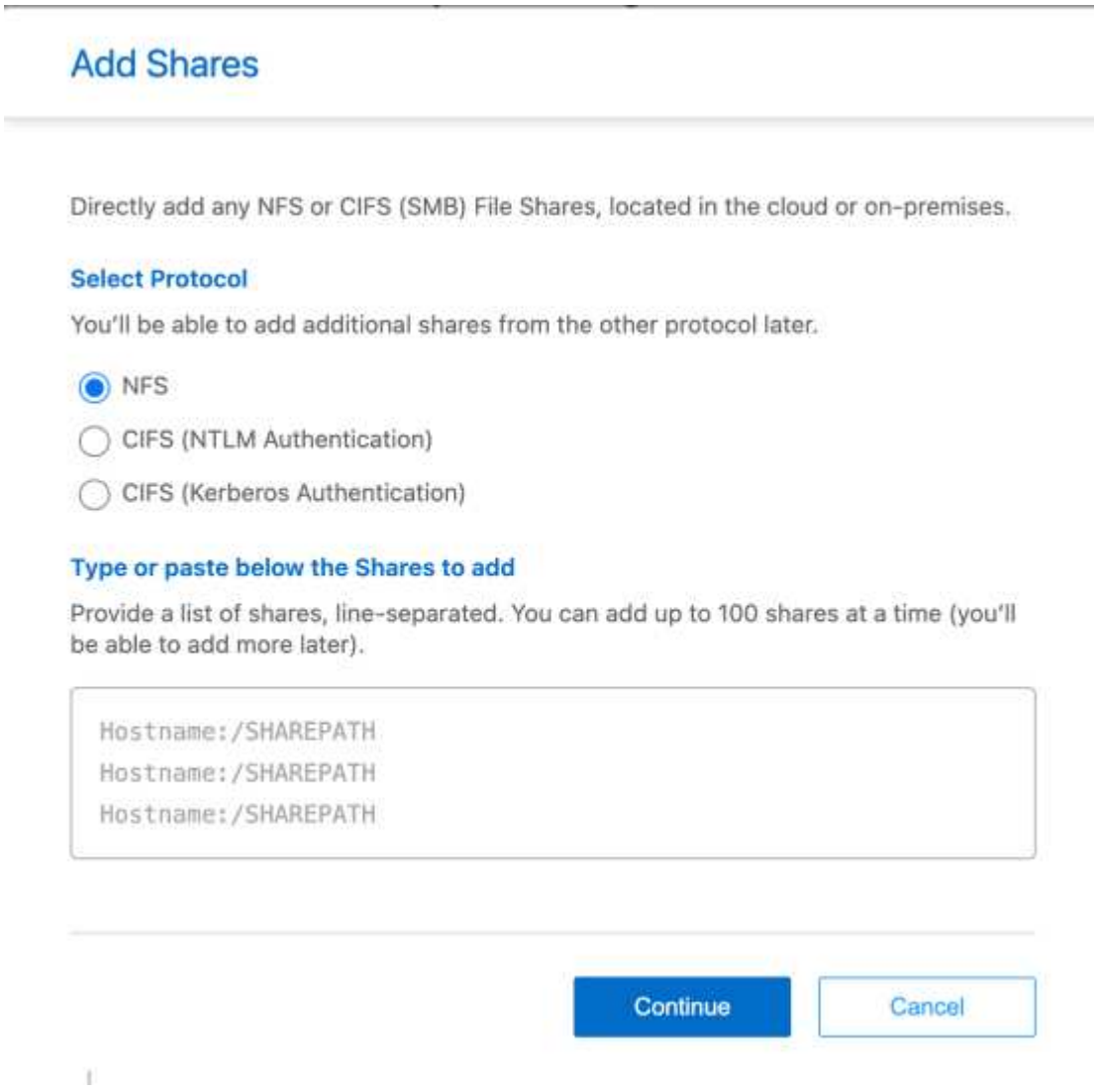
Lorsque vous ajoutez des partages de fichiers au groupe, vous devez utiliser le format `<host_name>:/<share_path>`.

Vous pouvez ajouter des partages de fichiers individuellement ou saisir une liste séparée par des lignes des

partages de fichiers que vous souhaitez analyser. Vous pouvez ajouter jusqu'à 100 actions à la fois.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Depuis la page Configuration, sélectionnez **Ajouter un système** > **Ajouter un groupe de partages de fichiers**.
3. Dans la boîte de dialogue Ajouter un groupe de partages de fichiers, saisissez le nom du groupe de partages, puis sélectionnez **Continuer**.
4. Sélectionnez le protocole pour les partages de fichiers que vous ajoutez.



   a. Si vous ajoutez des partages CIFS avec l'authentification NTLM, entrez les informations d'identification Active Directory pour accéder aux volumes CIFS. Bien que les informations d'identification en lecture seule soient prises en charge, il est recommandé de fournir un accès complet avec les informations d'identification d'administrateur. Sélectionnez **Enregistrer**.

5. Ajoutez les partages de fichiers que vous souhaitez analyser (un partage de fichiers par ligne). Sélectionnez ensuite **Continuer**.
6. Une boîte de dialogue de confirmation affiche le nombre de partages qui ont été ajoutés.

   Si la boîte de dialogue répertorie des partages qui n'ont pas pu être ajoutés, capturez ces informations afin

de pouvoir résoudre le problème. Si le problème concerne une convention de dénomination, vous pouvez rajouter le partage avec un nom corrigé.

7. Configurer l'analyse sur le volume :
   - Pour activer les analyses de mappage uniquement sur les partages de fichiers, sélectionnez **Map**.
   - Pour activer les analyses complètes sur les partages de fichiers, sélectionnez **Mappez et classez**.
   - Pour désactiver l'analyse sur les partages de fichiers, sélectionnez **Désactivé**.

> Le commutateur en haut de la page pour **Analyser lorsque les autorisations « attributs d'écriture » sont manquantes** est désactivé par défaut. Cela signifie que si Data Classification ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations d'écriture dans NFS, le système n'analysera pas les fichiers car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. + Si vous activez **Analyser en cas d'absence d'autorisations « attributs d'écriture »**, l'analyse réinitialise l'heure du dernier accès et analyse tous les fichiers, quelles que soient les autorisations. + Pour en savoir plus sur l'horodatage du dernier accès, voir"Métadonnées collectées à partir de sources de données dans la classification des données" .

Résultat

La classification des données commence à analyser les fichiers dans les partages de fichiers que vous avez ajoutés. Tu peuxSuivre la progression de la numérisation et afficher les résultats de l'analyse dans le **Tableau de bord**.

> Si l'analyse ne se termine pas correctement pour une configuration CIFS avec authentification Kerberos, vérifiez l'onglet **Configuration** pour détecter d'éventuelles erreurs.

Modifier un groupe de partage de fichiers

Après avoir créé un groupe de partages de fichiers, vous pouvez modifier le protocole CIFS ou ajouter et supprimer des partages de fichiers.

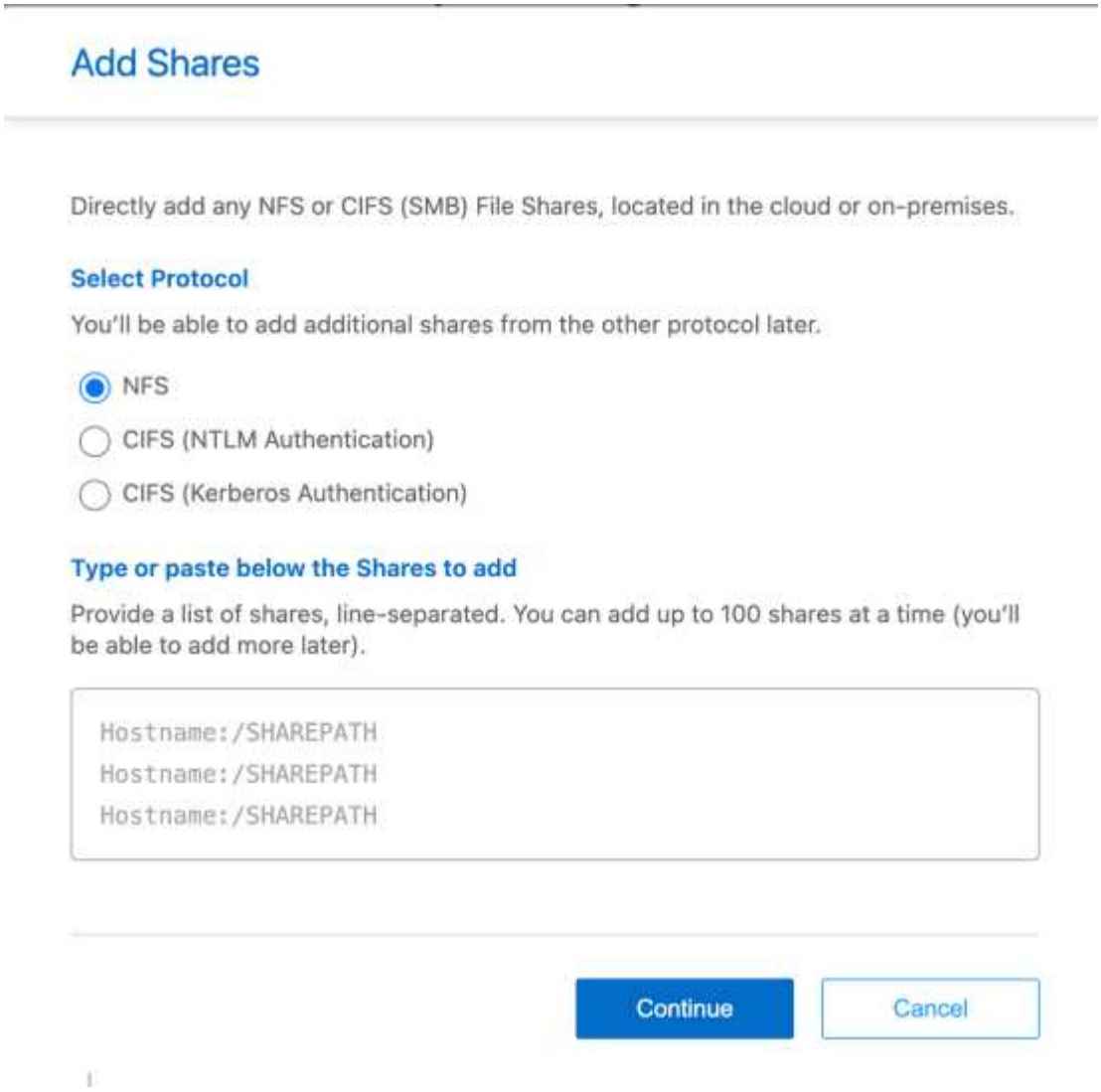
Modifier la configuration du protocole CIFS

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Depuis la page Configuration, sélectionnez le groupe de partages de fichiers que vous souhaitez modifier.
3. Sélectionnez **Modifier les informations d'identification CIFS**.

4. Choisissez la méthode d'authentification : **NTLM** ou **Kerberos**.
5. Saisissez le **Nom d'utilisateur** et le **Mot de passe** d'Active Directory.
6. Sélectionnez **Enregistrer** pour terminer le processus.

Ajouter des partages de fichiers aux analyses de conformité

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Depuis la page Configuration, sélectionnez le groupe de partages de fichiers que vous souhaitez modifier.
3. Sélectionnez **+ Ajouter des partages**.
4. Sélectionnez le protocole pour les partages de fichiers que vous ajoutez.



Si vous ajoutez des partages de fichiers à un protocole que vous avez déjà configuré, aucune modification n'est requise.

Si vous ajoutez des partages de fichiers avec un deuxième protocole, assurez-vous d'avoir correctement configuré l'authentification comme détaillé dans le"prérequis" .

5. Ajoutez les partages de fichiers que vous souhaitez analyser (un partage de fichiers par ligne) en utilisant le format `<host_name>:/<share_path>` .
6. Sélectionnez **Continuer** pour terminer l'ajout des partages de fichiers.

Supprimer un partage de fichiers des analyses de conformité

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Sélectionnez le système dont vous souhaitez supprimer les partages de fichiers.
3. Sélectionnez **Configuration**.
4. Depuis la page Configuration, sélectionnez les Actions ••• pour le partage de fichiers que vous souhaitez supprimer.
5. Dans le menu Actions, sélectionnez **Supprimer le partage**.

Suivre la progression de la numérisation

Vous pouvez suivre la progression de l'analyse initiale.

1. Sélectionnez le menu **Configuration**.
2. Sélectionnez la **Configuration système**.
3. Pour le référentiel de stockage, vérifiez la colonne Progression de l'analyse pour afficher son état.

Analyser les données StorageGRID avec la NetApp Data Classification

Suivez quelques étapes pour commencer à analyser les données dans StorageGRID directement avec NetApp Data Classification.

Examiner les exigences de StorageGRID

Passez en revue les conditions préalables suivantes pour vous assurer que vous disposez d'une configuration prise en charge avant d'activer la classification des données.

- Vous devez disposer de l'URL du point de terminaison pour vous connecter au service de stockage d'objets.
- Vous devez disposer de la clé d'accès et de la clé secrète de StorageGRID afin que la classification des données puisse accéder aux buckets.

Déployer l'instance de classification des données

Déployez la classification des données s'il n'existe pas déjà d'instance déployée.

Si vous numérisez des données de StorageGRID accessibles sur Internet, vous pouvez"déployer la classification des données dans le cloud" ou"déployer la classification des données dans un emplacement sur site disposant d'un accès Internet" .

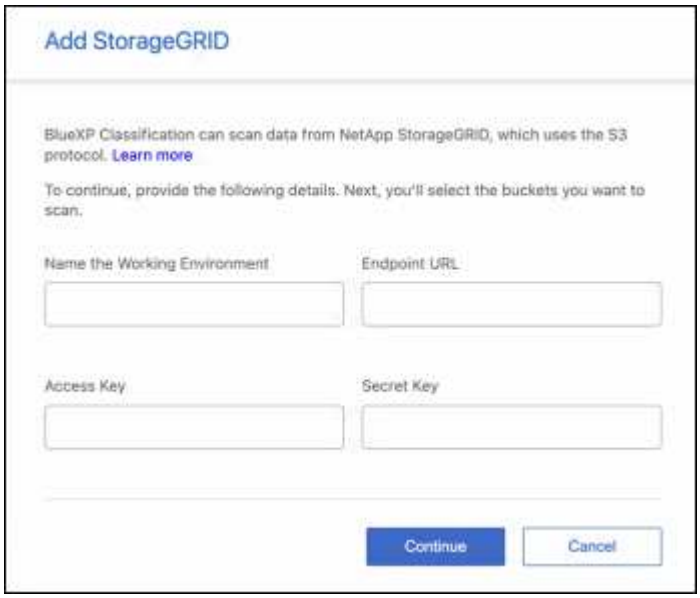
Si vous numérisez des données à partir de StorageGRID qui a été installé sur un site sombre qui n'a pas d'accès Internet, vous devez"déployer la classification des données dans le même emplacement sur site qui n'a pas d'accès Internet" . Cela nécessite également que l'agent de console soit déployé dans le même emplacement sur site.

Ajoutez le service StorageGRID à la classification des données

Ajoutez le service StorageGRID .

Étapes

1. Dans le menu Classification des données, sélectionnez l'option **Configuration**.
2. Depuis la page Configuration, sélectionnez **Ajouter un système** > **Ajouter StorageGRID**.
3. Dans la boîte de dialogue Ajouter un service StorageGRID , saisissez les détails du service StorageGRID et sélectionnez **Continuer**.
   a. Entrez le nom que vous souhaitez utiliser pour le système. Ce nom doit refléter le nom du service StorageGRID auquel vous vous connectez.
   b. Saisissez l'URL du point de terminaison pour accéder au service de stockage d'objets.
   c. Saisissez la clé d'accès et la clé secrète afin que la classification des données puisse accéder aux buckets dans StorageGRID.



Résultat

StorageGRID est ajouté à la liste des systèmes.

Activer et désactiver les analyses de conformité sur les buckets StorageGRID

Après avoir activé la classification des données sur StorageGRID, l'étape suivante consiste à configurer les buckets que vous souhaitez analyser. La classification des données découvre ces compartiments et les affiche dans le système que vous avez créé.

Étapes

1. Dans la page Configuration, recherchez le système StorageGRID .
2. Sur la mosaïque système StorageGRID , sélectionnez **Configuration**.
3. Effectuez l'une des étapes suivantes pour activer ou désactiver l'analyse :
   - Pour activer les analyses de mappage uniquement sur un bucket, sélectionnez **Carte**.
   - Pour activer les analyses complètes sur un bucket, sélectionnez **Cartographier et classer**.
   - Pour désactiver l'analyse sur un bucket, sélectionnez **Désactivé**.

Résultat

La classification des données commence à analyser les compartiments que vous avez activés. Vous pouvez suivre la progression de l'analyse initiale en accédant au menu **Configuration** puis en sélectionnant la **Configuration système**. La progression de chaque analyse est affichée sous forme de barre de progression. Vous pouvez également survoler la barre de progression pour voir le nombre de fichiers analysés par rapport au nombre total de fichiers dans le volume. S'il y a des erreurs, elles apparaîtront dans la colonne Statut, à côté de l'action requise pour corriger l'erreur.

Intégrez votre Active Directory à la NetApp Data Classification

Vous pouvez intégrer un Active Directory global à NetApp Data Classification pour améliorer les résultats rapportés par Data Classification sur les propriétaires de fichiers et sur les utilisateurs et groupes ayant accès à vos fichiers.

Lorsque vous configurez certaines sources de données (répertoriées ci-dessous), vous devez saisir les informations d'identification Active Directory pour que la classification des données analyse les volumes CIFS. Cette intégration fournit une classification des données avec les détails du propriétaire du fichier et des autorisations pour les données qui résident dans ces sources de données. L'Active Directory saisi pour ces sources de données peut différer des informations d'identification Active Directory globales que vous saisissez ici. La classification des données recherchera dans tous les annuaires Active Directory intégrés les détails des utilisateurs et des autorisations.

Cette intégration fournit des informations supplémentaires aux emplacements suivants dans la classification des données :

- Vous pouvez utiliser le « Propriétaire du fichier »"filtre" et voir les résultats dans les métadonnées du fichier dans le volet Investigation. Au lieu du propriétaire du fichier contenant le SID (Security IDentifier), il est renseigné avec le nom d'utilisateur réel.

  Vous pouvez également afficher plus de détails sur le propriétaire du fichier : nom du compte, adresse e-mail et nom du compte SAM, ou afficher les éléments appartenant à cet utilisateur.

- Tu peux voir"autorisations complètes du fichier" pour chaque fichier et répertoire lorsque vous cliquez sur le bouton « Afficher toutes les autorisations ».
- Dans le"Tableau de bord de gouvernance" , le panneau Autorisations d'ouverture affichera un niveau de détail plus élevé sur vos données.

Les SID des utilisateurs locaux et les SID des domaines inconnus ne sont pas traduits en nom d'utilisateur réel.

Sources de données prises en charge

Une intégration Active Directory avec la classification des données peut identifier les données à partir des sources de données suivantes :

- Systèmes ONTAP sur site
- Cloud Volumes ONTAP
- Azure NetApp Files
- FSx pour ONTAP

Connectez-vous à votre serveur Active Directory

Après avoir déployé la classification des données et activé l'analyse sur vos sources de données, vous pouvez intégrer la classification des données à votre Active Directory. Active Directory est accessible à l'aide d'une adresse IP de serveur DNS ou d'une adresse IP de serveur LDAP.

Les informations d'identification Active Directory peuvent être en lecture seule, mais la fourniture d'informations d'identification d'administrateur garantit que la classification des données peut lire toutes les données nécessitant des autorisations élevées. Les informations d'identification sont stockées sur l'instance de classification des données.

Pour les volumes/partages de fichiers CIFS, si vous souhaitez vous assurer que les « heures de dernier accès » de vos fichiers ne sont pas modifiées par les analyses de classification des données, l'utilisateur doit disposer de l'autorisation d'écriture des attributs. Si possible, nous vous recommandons de faire en sorte que l'utilisateur configuré par Active Directory fasse partie d'un groupe parent de l'organisation disposant d'autorisations sur tous les fichiers.

Exigences

- Vous devez déjà avoir un Active Directory configuré pour les utilisateurs de votre entreprise.
- Vous devez disposer des informations pour Active Directory :
  - Adresse IP du serveur DNS ou plusieurs adresses IP

    ou

  Adresse IP du serveur LDAP ou plusieurs adresses IP
  - Nom d'utilisateur et mot de passe pour accéder au serveur
  - Nom de domaine (nom Active Directory)
  - Que vous utilisiez ou non un LDAP sécurisé (LDAPS)
  - Port du serveur LDAP (généralement 389 pour LDAP et 636 pour LDAP sécurisé)
- Les ports suivants doivent être ouverts pour la communication sortante par l'instance de classification des données :

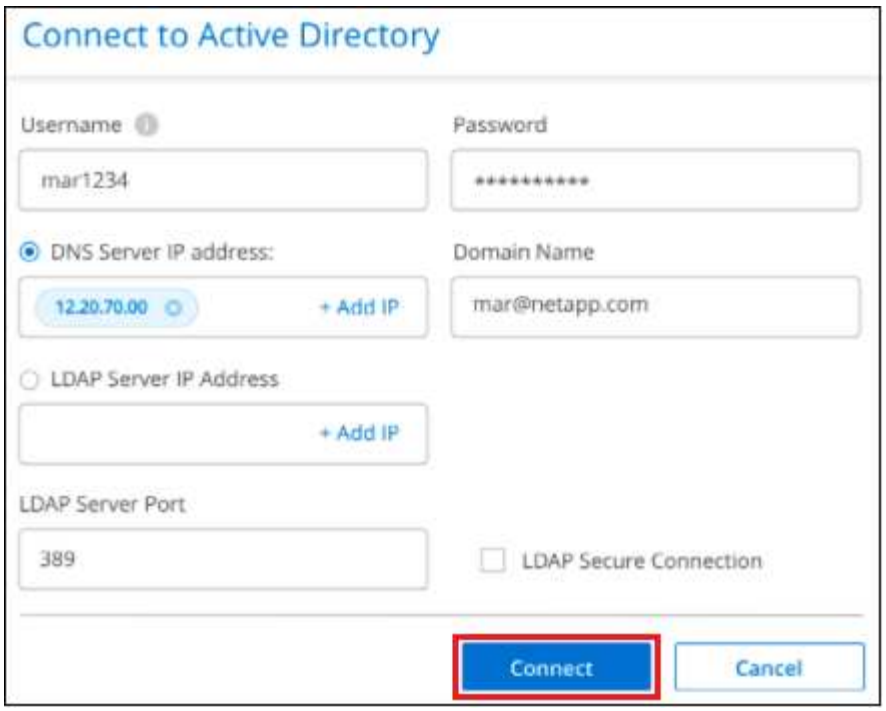
Protocole	Port	Destination	But
TCP et UDP	389	Active Directory	LDAP
TCP	636	Active Directory	LDAP sur SSL
TCP	3268	Active Directory	Catalogue mondial
TCP	3269	Active Directory	Catalogue global via SSL

Étapes

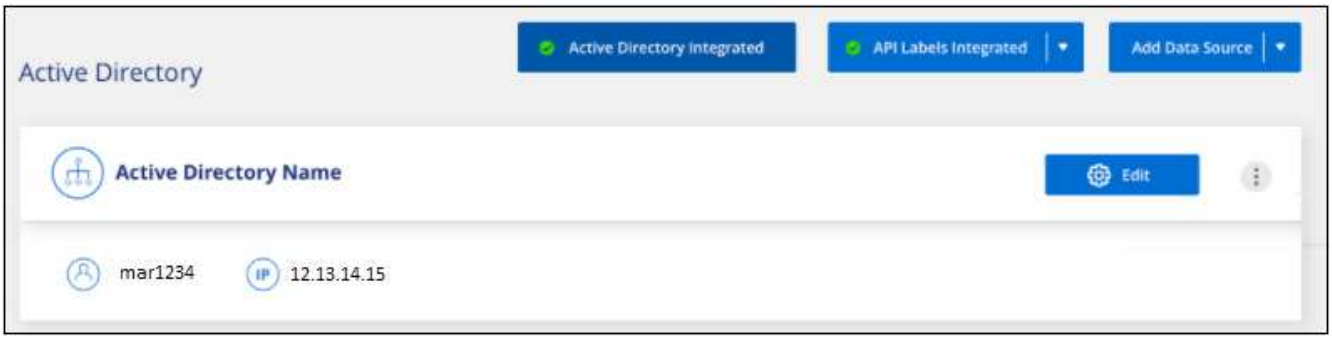
1. Depuis la page Configuration de la classification des données, cliquez sur **Ajouter Active Directory**.

2. Dans la boîte de dialogue Se connecter à Active Directory, entrez les détails d'Active Directory et cliquez sur **Connecter**.

   Vous pouvez ajouter plusieurs adresses IP, si nécessaire, en sélectionnant **Ajouter une IP**.



La classification des données s'intègre à Active Directory et une nouvelle section est ajoutée à la page de configuration.



Gérez votre intégration Active Directory

Si vous devez modifier des valeurs dans votre intégration Active Directory, cliquez sur le bouton **Modifier** et effectuez les modifications.

Vous pouvez également supprimer l'intégration en sélectionnant l'option bouton puis **Supprimer Active Directory**.

Classification des données d’utilisation

Affichez les détails de gouvernance sur les données stockées dans votre organisation avec NetApp Data Classification

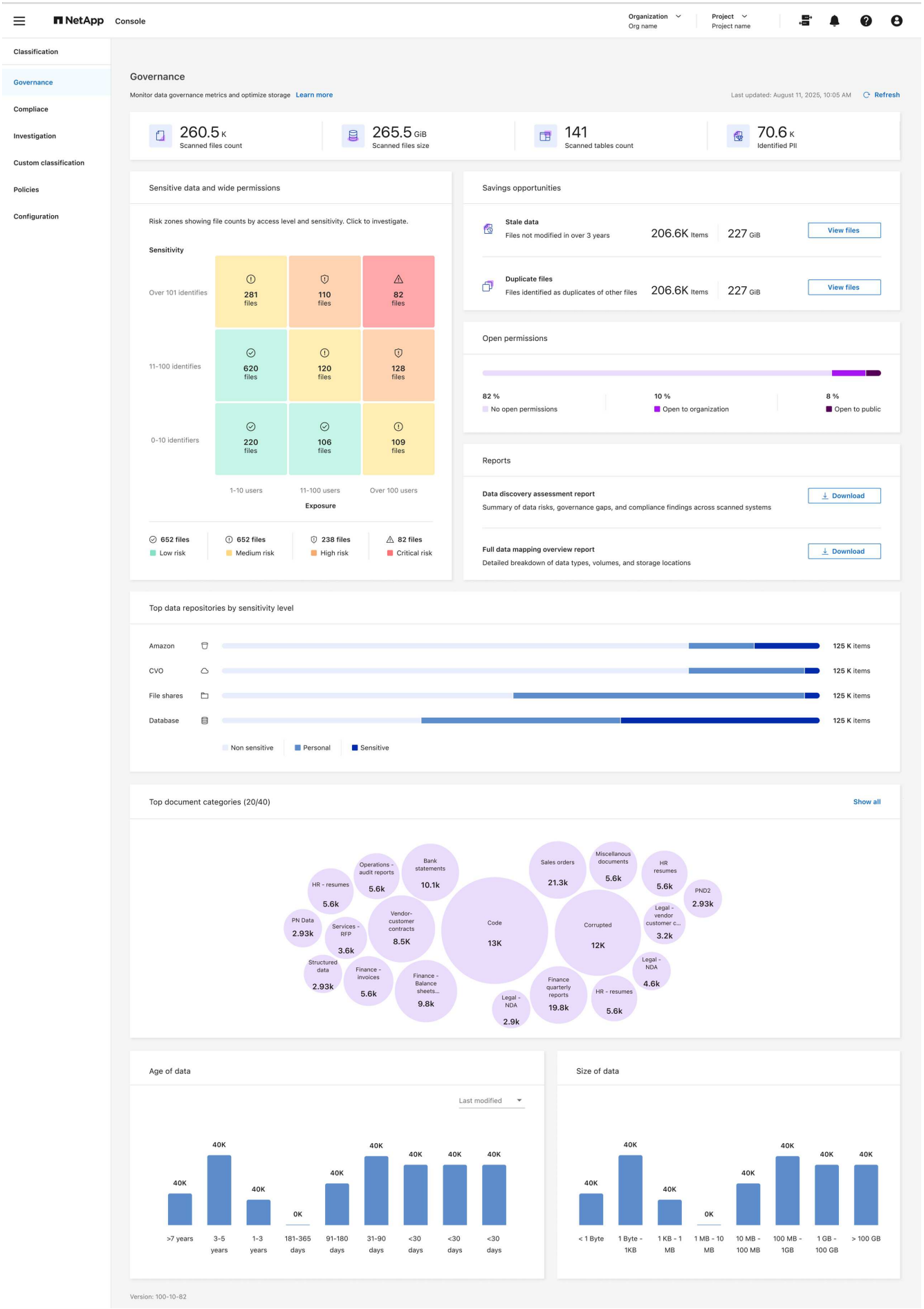
Maîtrisez les coûts liés aux données sur les ressources de stockage de votre organisation. La NetApp Data Classification identifie la quantité de données obsolètes, de fichiers en double et de fichiers très volumineux dans vos systèmes afin que vous puissiez décider si vous souhaitez supprimer ou hiérarchiser certains fichiers vers un stockage d’objets moins coûteux.

C’est ici que vous devriez commencer vos recherches. Depuis le tableau de bord de gouvernance, vous pouvez sélectionner un domaine pour une enquête plus approfondie.

De plus, si vous envisagez de migrer des données depuis des emplacements locaux vers le cloud, vous pouvez afficher la taille des données et si certaines d’entre elles contiennent des informations sensibles avant de les déplacer.

Consultez le tableau de bord de gouvernance

Le tableau de bord de gouvernance fournit des informations pour vous permettre d’augmenter l’efficacité et de contrôler les coûts liés aux données stockées sur vos ressources de stockage.


NetApp Console
Organization
Org name
Project
Project name
Classification
Governance
Compliace
Investigation
Custom classification
Policies
Configuration
Governance
Monitor data governance metrics and optimize storage Learn more
Last updated: August 11, 2025, 10:05 AM
Refresh
260.5 K
Scanned files count
265.5 GiB
Scanned files size
141
Scanned tables count
70.6 K
Identified PII
Sensitive data and wide permissions
Risk zones showing file counts by access level and sensitivity. Click to investigate.
Sensitivity
Over 101 identifies
11-100 identifies
0-10 identifiers
281 files
110 files
82 files
620 files
120 files
128 files
220 files
106 files
109 files
1-10 users
11-100 users
Over 100 users
Exposure
652 files
Low risk
652 files
Medium risk
238 files
High risk
82 files
Critical risk
Savings opportunities
Stale data
Files not modified in over 3 years
206.6K Items
227 GiB
View files
Duplicate files
Files identified as duplicates of other files
206.6K Items
227 GiB
View files
Open permissions
82 %
No open permissions
10 %
Open to organization
8 %
Open to public
Reports
Data discovery assessment report
Summary of data risks, governance gaps, and compliance findings across scanned systems
Download
Full data mapping overview report
Detailed breakdown of data types, volumes, and storage locations
Download
Top data repositories by sensitivity level
Amazon
CVO
File shares
Database
125 K items
Non sensitive
Personal
Sensitive
Top document categories (20/40)
Show all
Operations - audit reports 5.6k
Bank statements 10.1k
Sales orders 21.3k
Miscellanous documents 5.6k
HR resumes 5.6k
PND2 2.93k
HR - resumes 5.6k
PN Data 2.93k
Services - RFP 3.6k
Vendor-customer contracts 8.5K
Code 13K
Corrupted 12K
Legal - vendor customer c... 3.2k
Structured data 2.93k
Finance - invoices 5.6k
Finance - Balance sheets... 9.8k
Legal - NDA 2.9k
Finance quarterly reports 19.8k
HR - resumes 5.6k
Legal - NDA 4.6k
Age of data
Last modified
40K
>7 years
40K
3-5 years
40K
1-3 years
0K
181-365 days
40K
91-180 days
40K
31-90 days
40K
<30 days
40K
<30 days
40K
<30 days
Size of data
40K
< 1 Byte
40K
1 Byte - 1KB
40K
1 KB - 1 MB
0K
1 MB - 10 MB
40K
10 MB - 100 MB
40K
100 MB - 1GB
40K
1 GB - 100 GB
40K
> 100 GB
Version: 100-10-82

Étapes

1. Dans le menu de la NetApp Console , sélectionnez **Gouvernance > Classification**.
2. Sélectionnez **Gouvernance**.

   Le tableau de bord de gouvernance apparaît.

Examiner les possibilités d'économies

Le composant *Opportunités d'économie* affiche les données que vous pouvez supprimer ou hiérarchiser vers un stockage d'objets moins coûteux. Les données de *Saving Opportunities* sont mises à jour toutes les 2 heures. Vous pouvez également mettre à jour les données manuellement.

Étapes

1. Dans le menu Classification des données, sélectionnez **Gouvernance**.
2. Dans chaque mosaïque Opportunités d'économies du tableau de bord Gouvernance, sélectionnez **Optimiser le stockage** pour afficher les résultats filtrés dans la page Enquête. Pour découvrir les données que vous devriez supprimer ou transférer vers un stockage moins coûteux, étudiez les *Opportunités d'économie*.
   - **Données obsolètes** - Données qui ont été modifiées pour la dernière fois il y a plus de 3 ans.
   - **Fichiers en double** - Fichiers dupliqués à d'autres emplacements dans les sources de données que vous analysez. "Voir quels types de fichiers en double sont affichés" .

> Si l'une de vos sources de données implémente la hiérarchisation des données, les anciennes données qui résident déjà dans le stockage d'objets peuvent être identifiées dans la catégorie *Données obsolètes*.

Créer le rapport d'évaluation de la découverte de données

Le rapport d'évaluation de la découverte de données fournit une analyse de haut niveau de l'environnement analysé pour montrer les zones de préoccupation et les étapes de correction potentielles. Les résultats sont basés à la fois sur la cartographie et la classification de vos données. L'objectif de ce rapport est de sensibiliser à trois aspects importants de votre ensemble de données :

Fonctionnalité	Description
Préoccupations liées à la gouvernance des données	Une image détaillée de toutes les données que vous possédez et des domaines dans lesquels vous pouvez réduire la quantité de données pour économiser des coûts.
Expositions à la sécurité des données	Zones dans lesquelles vos données sont accessibles aux attaques internes ou externes en raison d'autorisations d'accès étendues.
Lacunes en matière de conformité des données	Où se trouvent vos informations personnelles ou sensibles, à des fins de sécurité et pour les DSAR (demandes d'accès aux données des personnes concernées).

Avec le rapport, vous pouvez effectuer les actions suivantes :

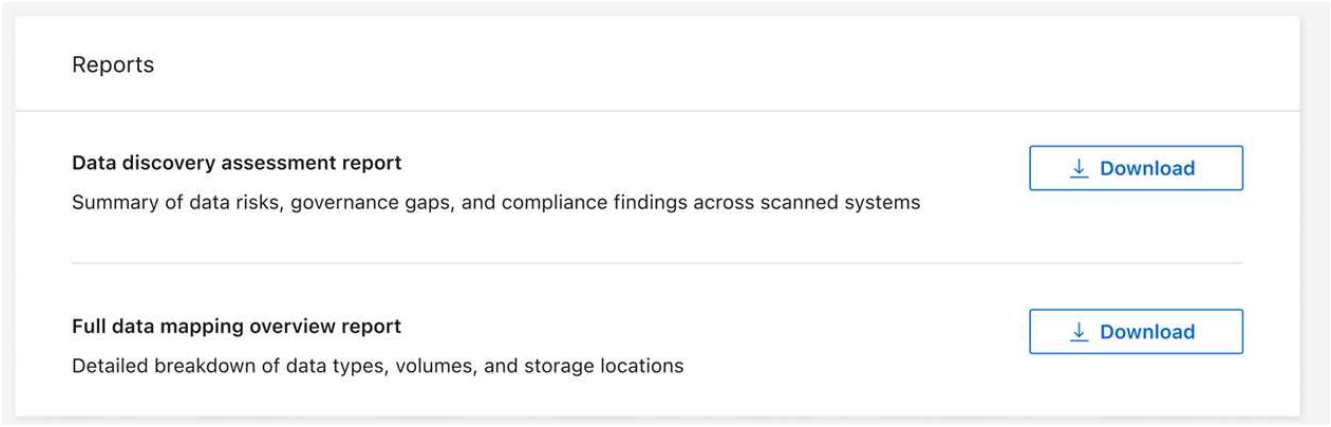
- Réduisez les coûts de stockage en modifiant votre politique de conservation, ou en déplaçant ou en supprimant certaines données (données obsolètes ou en double).
- Protégez vos données disposant d'autorisations étendues en révisant les politiques de gestion des

groupes globaux.

- Protégez vos données contenant des informations personnelles ou sensibles en déplaçant les PII vers des magasins de données plus sécurisés.

Étapes

1. Dans Classification des données, sélectionnez **Gouvernance**.
2. Dans la mosaïque des rapports, sélectionnez **Rapport d'évaluation de la découverte de données**.



Résultat

La classification des données génère un rapport PDF que vous pouvez consulter et partager.

Créer le rapport de synthèse du mappage des données

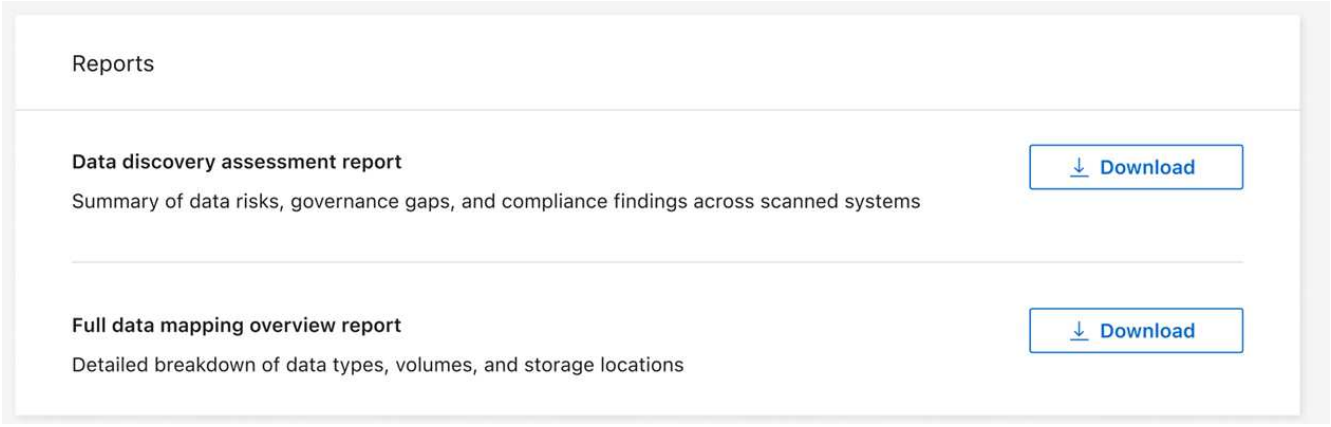
Le rapport d'aperçu du mappage des données fournit un aperçu des données stockées dans vos sources de données d'entreprise pour vous aider à prendre des décisions concernant les processus de migration, de sauvegarde, de sécurité et de conformité. Le rapport résume tous les systèmes et sources de données. Il fournit également une analyse pour chaque système.

Le rapport comprend les informations suivantes :

Catégorie	Description
Capacité d'utilisation	Pour tous les systèmes : répertorie le nombre de fichiers et la capacité utilisée pour chaque système. Pour les systèmes uniques : répertorie les fichiers qui utilisent le plus de capacité.
L'ère des données	Fournit trois tableaux et graphiques indiquant quand les fichiers ont été créés, modifiés pour la dernière fois ou consultés pour la dernière fois. Répertorie le nombre de fichiers et leur capacité utilisée, en fonction de certaines plages de dates.
Taille des données	Répertorie le nombre de fichiers qui existent dans certaines plages de taille dans vos systèmes.

Étapes

1. Dans Classification des données, sélectionnez **Gouvernance**.
2. Dans la mosaïque des rapports, sélectionnez **Rapport d'aperçu complet du mappage des données**.



Résultat

La classification des données génère un rapport PDF que vous pouvez consulter et envoyer à d'autres groupes selon vos besoins.

Si le rapport est supérieur à 1 Mo, le fichier PDF est conservé sur l'instance de classification des données et vous verrez un message contextuel indiquant l'emplacement exact. Lorsque Data Classification est installé sur une machine Linux dans vos locaux ou sur une machine Linux que vous avez déployée dans le cloud, vous pouvez accéder directement au fichier PDF. Lorsque la classification des données est déployée dans le cloud, vous devez autoriser avec SSH l'instance de classification des données pour télécharger le fichier PDF.

Consultez les principaux référentiels de données classés par sensibilité des données

La zone *Principaux référentiels de données par niveau de sensibilité* du rapport Présentation du mappage des données répertorie les quatre principaux référentiels de données (systèmes et sources de données) qui contiennent les éléments les plus sensibles. Le graphique à barres de chaque système est divisé en :

- Données non sensibles
- Données personnelles
- Données personnelles sensibles

Ces données sont actualisées toutes les deux heures et peuvent être actualisées manuellement.

Étapes

1. Pour voir le nombre total d'éléments dans chaque catégorie, positionnez votre curseur sur chaque section de la barre.
2. Pour filtrer les résultats qui apparaîtront sur la page Enquête, sélectionnez chaque zone dans la barre et approfondissez vos recherches.

Examiner les données sensibles et les autorisations étendues

La zone *Données sensibles et autorisations étendues* du tableau de bord de gouvernance affiche le nombre de fichiers contenant des données sensibles et disposant d'autorisations étendues. Le tableau présente les types d'autorisations suivants :

- Des autorisations les plus restrictives aux restrictions les plus permissives sur l'axe horizontal.
- Des données les moins sensibles aux données les plus sensibles sur l'axe vertical.

Étapes

1. Pour voir le nombre total de fichiers dans chaque catégorie, positionnez votre curseur sur chaque case.
2. Pour filtrer les résultats qui apparaîtront sur la page Enquête, sélectionnez une case et poursuivez vos recherches.

Examiner les données répertoriées par types d'autorisations ouvertes

La zone *Autorisations d'ouverture* du rapport Présentation du mappage des données affiche le pourcentage pour chaque type d'autorisations qui existent pour tous les fichiers en cours d'analyse. Le graphique montre les types d'autorisations suivants :

- Aucune autorisation d'ouverture
- Ouvert à l'organisation
- Ouvert au public
- Accès inconnu

Étapes

1. Pour voir le nombre total de fichiers dans chaque catégorie, positionnez votre curseur sur chaque case.
2. Pour filtrer les résultats qui apparaîtront sur la page Enquête, sélectionnez une case et poursuivez vos recherches.

Vérifiez l'âge et la taille des données

Vous pouvez examiner les éléments des graphiques *Âge* et *Taille* du rapport Présentation du mappage des données pour voir s'il existe des données que vous devez supprimer ou transférer vers un stockage d'objets moins coûteux.

Étapes

1. Dans le graphique Âge des données, pour voir les détails sur l'âge des données, placez votre curseur sur un point du graphique.
2. Pour filtrer par tranche d'âge ou de taille, sélectionnez cet âge ou cette taille.
    - **Graphique de l'âge des données** - Catégorise les données en fonction de l'heure à laquelle elles ont été créées, de la dernière fois où elles ont été consultées ou de la dernière fois où elles ont été modifiées.
    - **Taille du graphique de données** - Catégorise les données en fonction de leur taille.

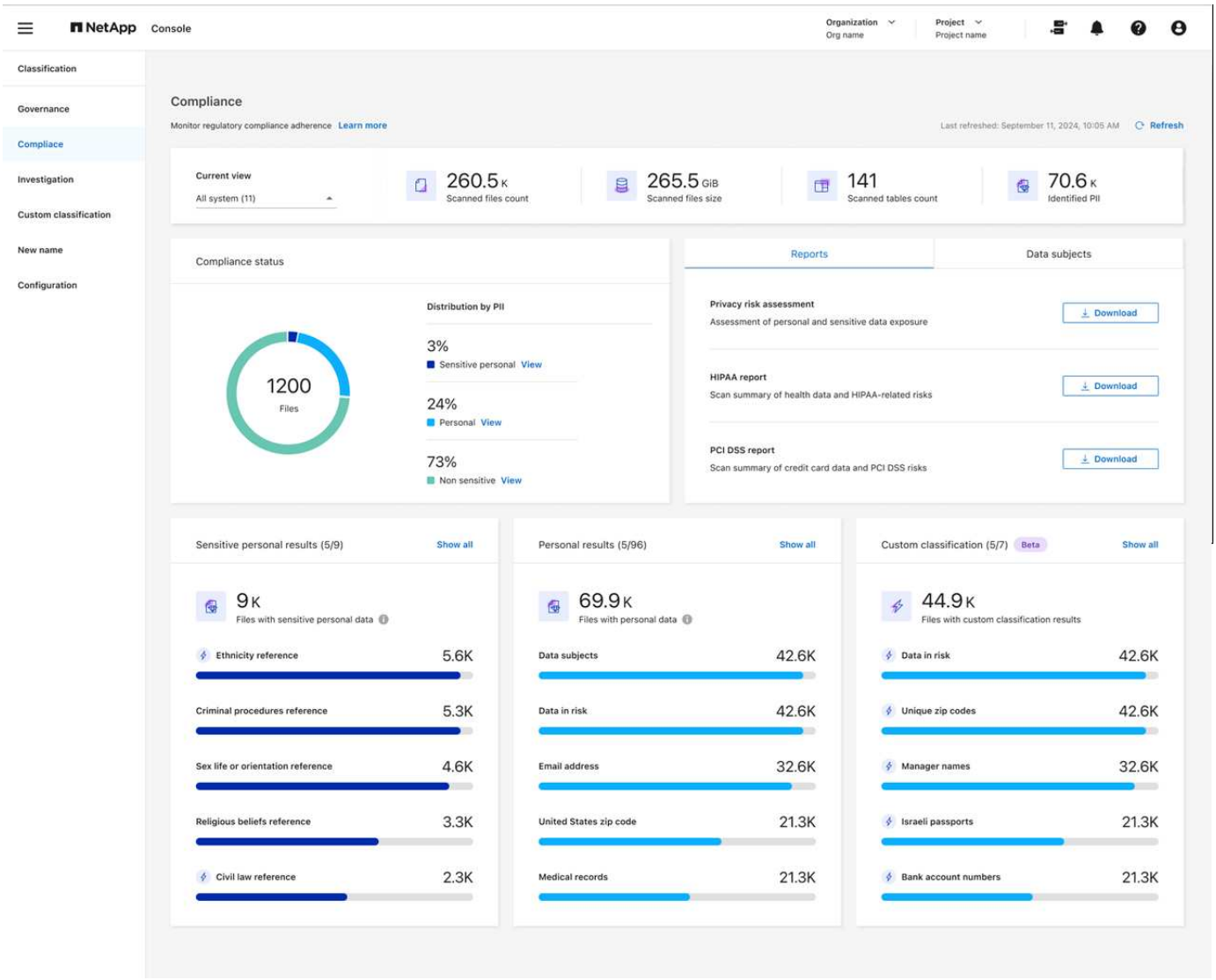
Si l'une de vos sources de données implémente la hiérarchisation des données, les anciennes données qui résident déjà dans le stockage d'objets peuvent être identifiées dans le graphique *Age of Data*.

Consultez les détails de conformité concernant les données privées stockées dans votre organisation avec NetApp Data Classification

Prenez le contrôle de vos données privées en consultant les détails sur les données personnelles (PII) et les données personnelles sensibles (SPII) de votre organisation. Vous pouvez également gagner en visibilité en examinant les catégories et les types de fichiers que NetApp Data Classification a trouvés dans vos données.

Les détails de conformité au niveau du fichier ne sont disponibles que si vous effectuez une analyse de classification complète. Les analyses de cartographie uniquement ne fournissent pas de détails au niveau du fichier.

Par défaut, le tableau de bord de classification des données affiche les données de conformité pour tous les systèmes et bases de données. Pour voir les données de certains systèmes uniquement, sélectionnez-les.



Vous pouvez filtrer les résultats de la page Enquête sur les données et télécharger un rapport des résultats sous forme de fichier CSV. Voir"Filtrage des données dans la page Enquête sur les données" pour plus de détails.

Afficher les fichiers contenant des données personnelles

La classification des données identifie automatiquement des mots, des chaînes et des modèles (Regex) spécifiques à l'intérieur des données. xref:./"Par exemple, les numéros de carte de crédit, les numéros de sécurité sociale, les numéros de compte bancaire, les mots de passe, etc." La classification des données identifie ce type d'informations dans des fichiers individuels, dans des fichiers au sein de répertoires (partages et dossiers) et dans des tables de base de données.

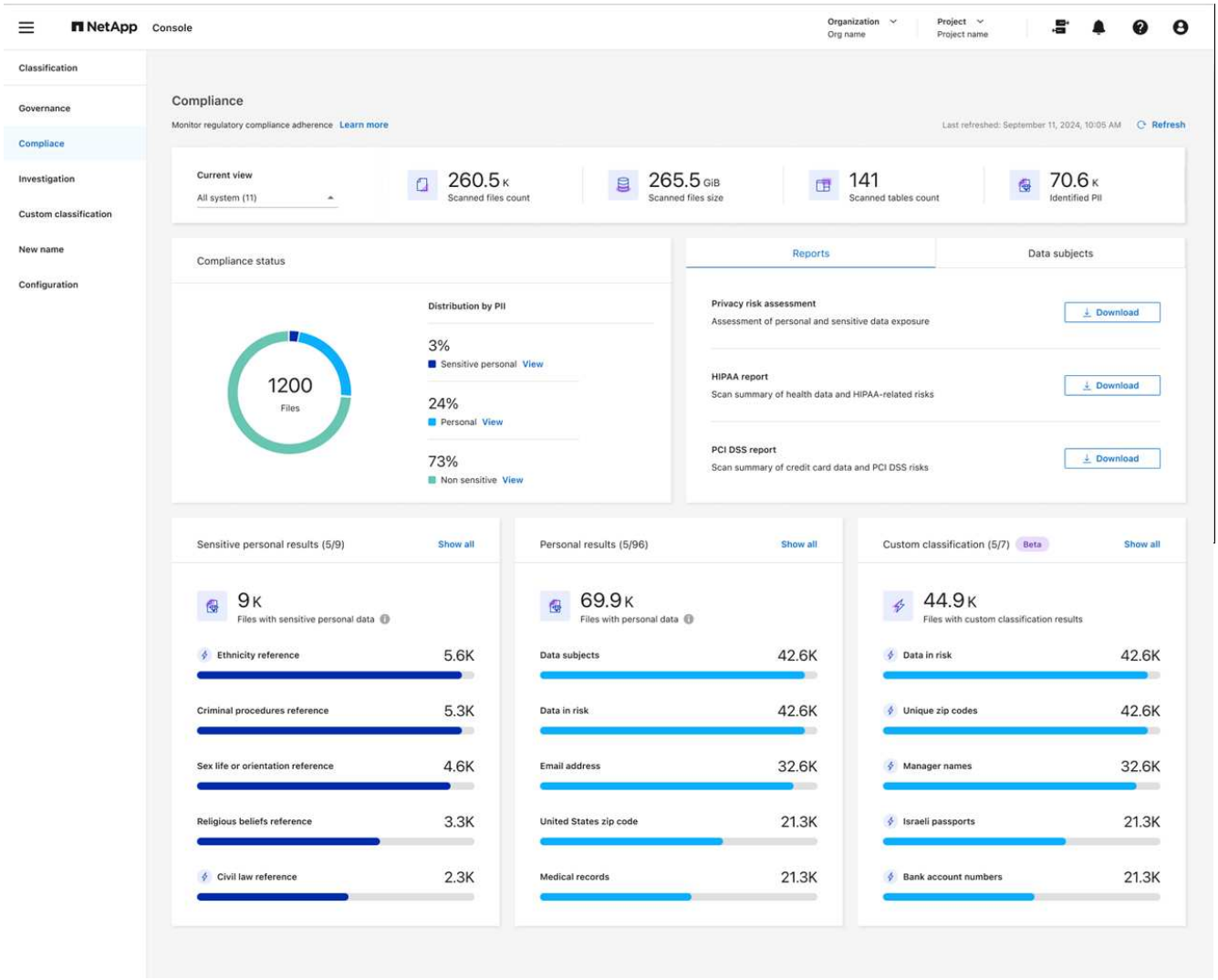
Vous pouvez également créer des termes de recherche personnalisés pour identifier les données personnelles

spécifiques à votre organisation. Pour plus d'informations, consultez la section "Créer une classification personnalisée" .

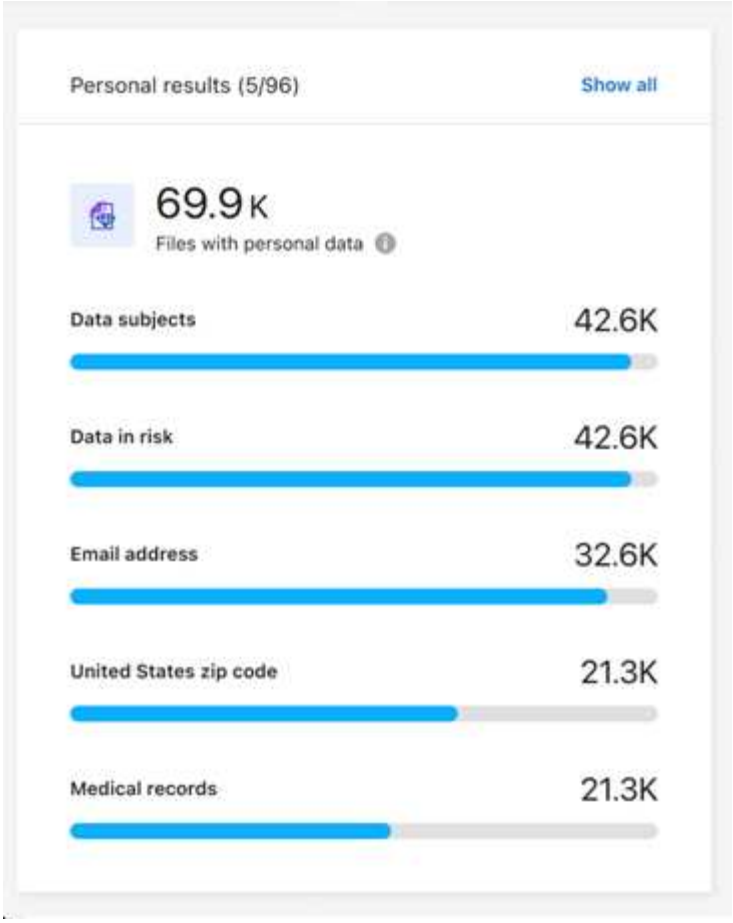
Pour certains types de données personnelles, la classification des données utilise la *validation de proximité* pour valider ses résultats. La validation s'effectue en recherchant un ou plusieurs mots-clés prédéfinis à proximité des données personnelles trouvées. Par exemple, la classification des données identifie un numéro de sécurité sociale américain (SSN) comme un SSN s'il voit un mot de proximité à côté de lui, par exemple, *SSN* ou *sécurité sociale*. "Le tableau des données personnelles" indique quand la classification des données utilise la validation de proximité.

Étapes

1. Dans le menu Classification des données, sélectionnez l'onglet **Conformité**.
2. Pour examiner les détails de toutes les données personnelles, sélectionnez l'icône à côté du pourcentage de données personnelles.

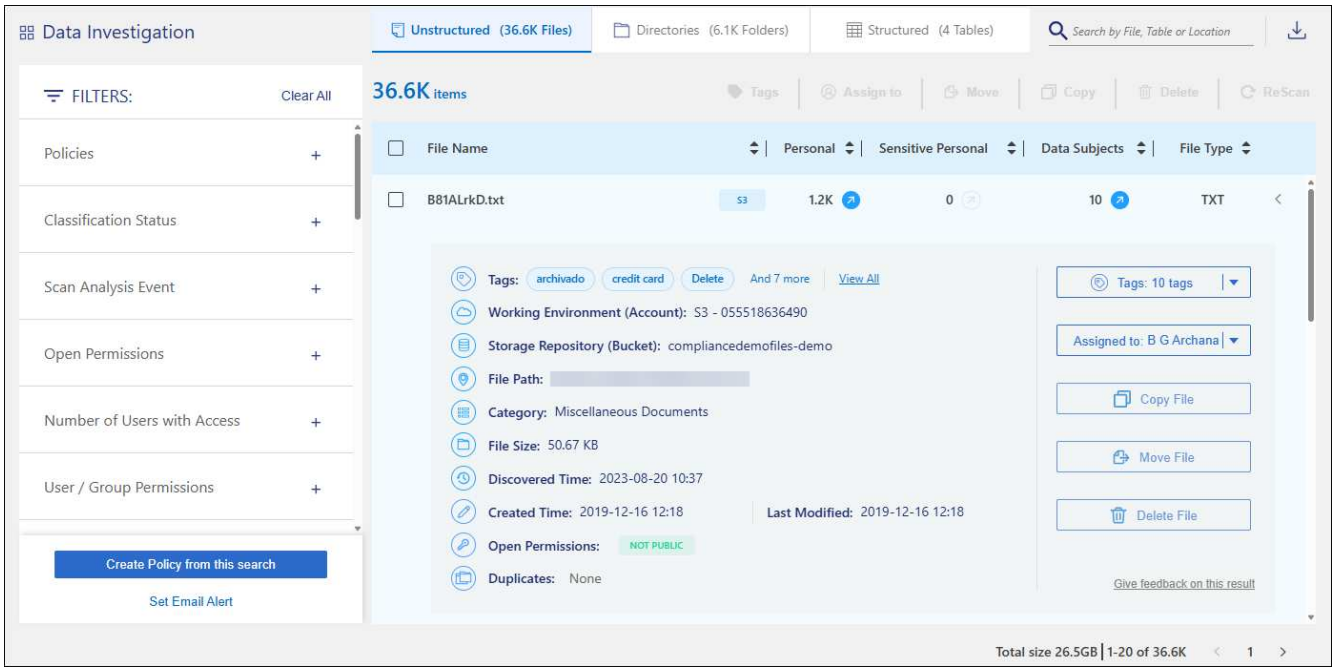


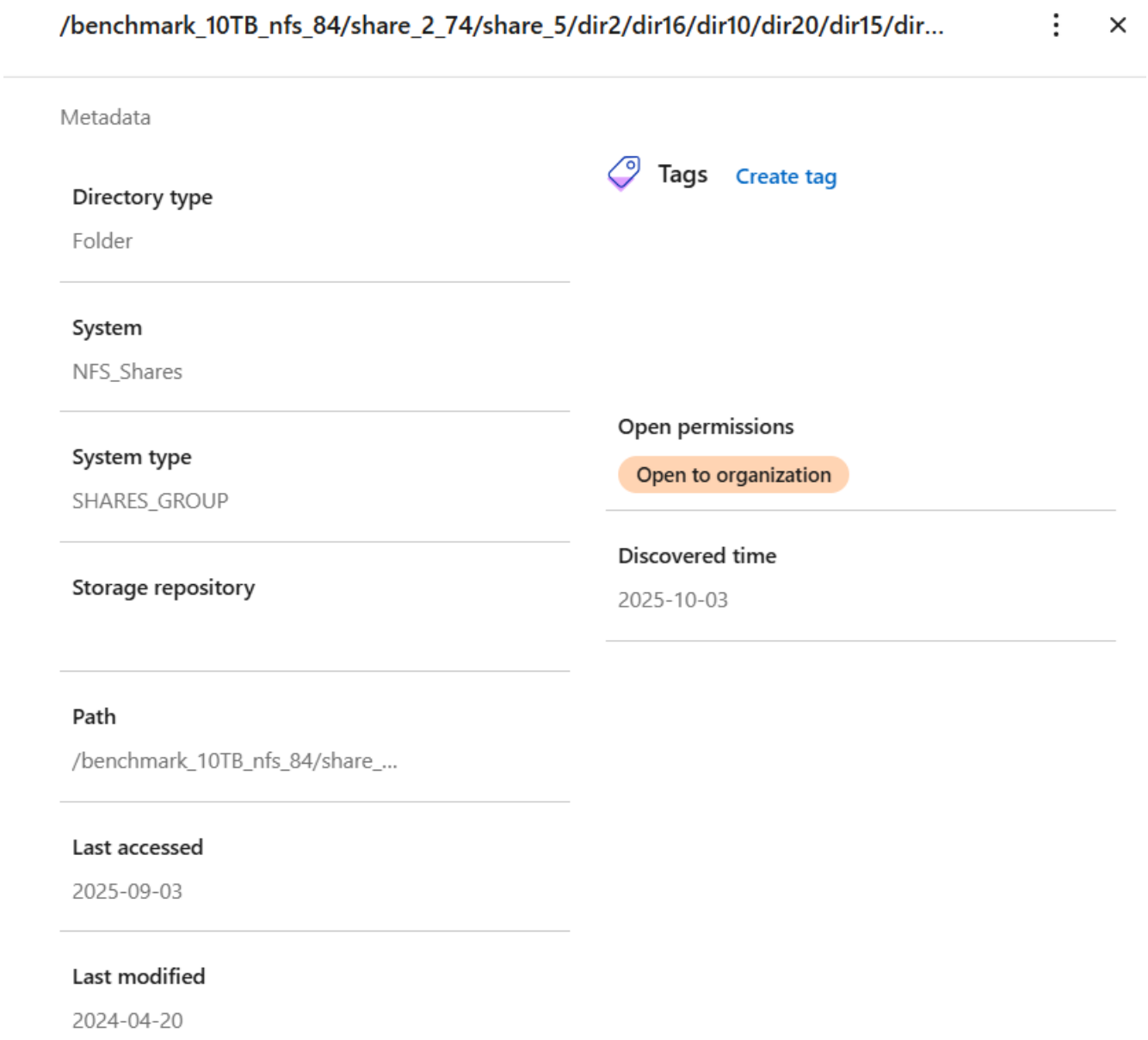
3. Pour examiner les détails d'un type spécifique de données personnelles, sélectionnez **Afficher tout**, puis sélectionnez l'icône en forme de flèche **Examiner les résultats** pour un type spécifique de données personnelles, par exemple, les adresses e-mail.



4. Examinez les données en recherchant, en triant, en développant les détails d'un fichier spécifique, en sélectionnant la flèche **Enquêter sur les résultats** pour voir les informations masquées ou en téléchargeant la liste des fichiers.

   L'image suivante montre les données personnelles trouvées dans un répertoire (partages et dossiers). Dans l'onglet **Structuré**, vous visualisez les données personnelles trouvées dans les bases de données. Dans l'onglet **Non structuré**, vous pouvez afficher les données au niveau du fichier.


Data Investigation
Unstructured (36.6K Files)
Directories (6.1K Folders)
Structured (4 Tables)
Search by File, Table or Location
FILTERS:
Clear All
Policies
Classification Status
Scan Analysis Event
Open Permissions
Number of Users with Access
User / Group Permissions
Create Policy from this search
Set Email Alert
36.6K items
Tags
Assign to
Move
Copy
Delete
ReScan
File Name
Personal
Sensitive Personal
Data Subjects
File Type
B81ALrkD.txt
S3
1.2K
0
10
TXT
Tags: archivado credit card Delete And 7 more View All
Working Environment (Account): S3 - 055518636490
Storage Repository (Bucket): compliancedemofiles-demo
File Path:
Category: Miscellaneous Documents
File Size: 50.67 KB
Discovered Time: 2023-08-20 10:37
Created Time: 2019-12-16 12:18
Last Modified: 2019-12-16 12:18
Open Permissions: NOT PUBLIC
Duplicates: None
Tags: 10 tags
Assigned to: B G Archana
Copy File
Move File
Delete File
Give feedback on this result
Total size 26.5GB | 1-20 of 36.6K
1



Afficher les fichiers contenant des données personnelles sensibles

La classification des données identifie automatiquement les types particuliers d'informations personnelles sensibles, telles que définies par les réglementations en matière de confidentialité telles que "articles 9 et 10 du RGPD" . Par exemple, des informations concernant la santé, l'origine ethnique ou l'orientation sexuelle d'une personne. "Voir la liste complète" . La classification des données identifie ce type d'informations dans des fichiers individuels, dans des fichiers au sein de répertoires (partages et dossiers) et dans des tables de base de données.

La classification des données utilise l'IA, le traitement du langage naturel (NLP), l'apprentissage automatique (ML) et l'informatique cognitive (CC) pour comprendre le sens du contenu qu'elle analyse afin d'extraire des entités et de les catégoriser en conséquence.

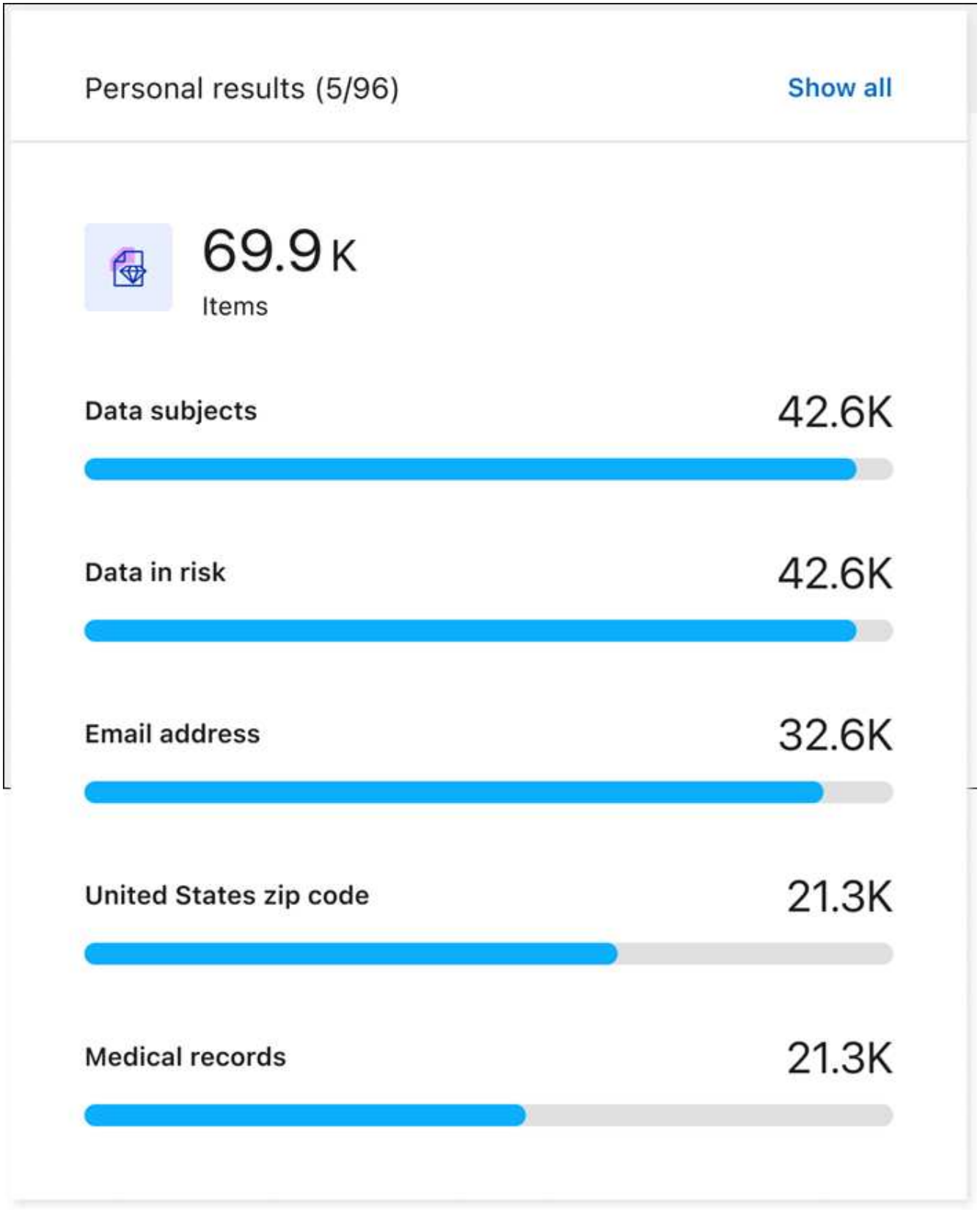
Par exemple, l'origine ethnique est une catégorie de données sensibles du RGPD. Grâce à ses capacités de PNL, la classification des données peut faire la différence entre une phrase qui dit « George est mexicain » (indiquant des données sensibles comme spécifié dans l'article 9 du RGPD), et « George mange de la

nourriture mexicaine ».

Seul l’anglais est pris en charge lors de la recherche de données personnelles sensibles. La prise en charge de davantage de langues sera ajoutée ultérieurement.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Pour examiner les détails de toutes les données personnelles sensibles, recherchez la carte **Résultats personnels sensibles**, puis sélectionnez **Afficher tout**.



.

3. Pour examiner les détails d'un type spécifique de données personnelles sensibles, sélectionnez **Afficher tout**, puis sélectionnez l'icône en forme de flèche **Enquêter sur les résultats** pour un type spécifique de données personnelles sensibles.
4. Examinez les données en recherchant, en triant, en développant les détails d'un fichier spécifique, en

cliquant sur **Enquêter sur les résultats** pour voir les informations masquées ou en téléchargeant la liste des fichiers.

Catégories de données privées dans la NetApp Data Classification

Il existe de nombreux types de données privées que NetApp Data Classification peut identifier dans vos volumes et bases de données.

La classification des données identifie deux types de données personnelles :

- **Informations personnelles identifiables (PII)**
- **Informations personnelles sensibles (SPII)**

Si vous avez besoin d'une classification des données pour identifier d'autres types de données privées, tels que des numéros d'identification nationaux supplémentaires ou des identifiants de soins de santé, contactez votre gestionnaire de compte.

Types de données personnelles

Les données personnelles, ou *informations personnellement identifiables* (PII), trouvées dans les fichiers peuvent être des données personnelles générales ou des identifiants nationaux. La troisième colonne du tableau ci-dessous indique si la classification des données utilise"validation de proximité" pour valider ses conclusions pour l'identifiant.

Les langues dans lesquelles ces éléments peuvent être reconnus sont identifiées dans le tableau.

Type	Identifiant	Validation de proximité ?	English	Allemand	Espagnol	Français	japonais
Général	Numéro de Carte de Crédit	Oui	✓	✓	✓		✓
	Personnes concernées	Non	✓	✓	✓		
	Adresse email	Non	✓	✓	✓		✓
	Numéro IBAN (numéro de compte bancaire international)	Non	✓	✓	✓		✓
	Adresse IP	Non	✓	✓	✓		✓
	Mot de passe	Oui	✓	✓	✓		✓

Type	Identifiant	Validation de proximité ?	English	Allemand	Espagnol	Français	japonais
Identifiants nationaux							

Type	Identifiant	Validation de proximité ?	English	Allemand	Espagnol	Français	japonais

Type	Identifiant	Validation de proximité ?	English	Allemand	Espagnol	Français	japonais

Type	Identifiant	Validation de proximité	English	Allemand	Espagnol	Français	japonais
	japonais (personnel et d'entreprise)						
	carte d'identité lettone	Oui	✓	✓	✓		
	carte d'identité lituanienne	Oui	✓	✓	✓		
	Carte d'identité luxembourgeoise	Oui	✓	✓	✓		
	Carte d'identité maltaise	Oui	✓	✓	✓		
	Numéro du Service national de santé (NHS)	Oui	✓	✓	✓		
	Compte bancaire néo-zélandais	Oui	✓	✓	✓		
	Permis de conduire néo-zélandais	Oui	✓	✓	✓		
	Numéro IRD de Nouvelle-Zélande (IDF)	Oui	✓	✓	✓		
	Numéro NHI (Indice national de santé) de Nouvelle-Zélande	Oui	✓	✓	✓		
	Numéro de passeport néo-zélandais	Oui	✓	✓	✓		
	Carte d'identité polonaise (PESEL)	Oui	✓	✓	✓		
	Numéro d'identification fiscale portugais (NIF)	Oui	✓	✓	✓		
	Carte d'identité roumaine (CNP)	Oui	✓	✓	✓		
	Carte d'identité nationale de Singapour (NRIC)	Oui	✓	✓	✓		
	Carte d'identité slovène (EMSO)	Oui	✓	✓	✓		
	Carte d'identité sud-africaine	Oui	✓	✓	✓		
	Numéro d'identification fiscale espagnol	Oui	✓	✓	✓		
	carte d'identité suédoise	Oui	✓	✓	✓		
	Carte d'identité britannique (NINO)	Oui	✓	✓	✓		
	Permis de conduire des États-Unis en Californie	Oui	✓	✓	✓		
	Permis de conduire de l'Indiana aux États-Unis	Oui	✓	✓	✓		
	Permis de conduire de l'État de New York aux États-Unis	Oui	✓	✓	✓		
	Permis de conduire des États-Unis au Texas	Oui	✓	✓	✓		
	Numéro de sécurité sociale aux États-Unis (SSN)	Oui	✓	✓	✓		

Types de données personnelles sensibles

La classification des données peut trouver les informations personnelles sensibles (SPII) suivantes dans les fichiers.

Les SPII suivants ne peuvent actuellement être reconnus qu'en anglais :

- **Référence aux procédures pénales** : Données concernant les condamnations pénales et les infractions

d’une personne physique.

- **Référence ethnique** : Données concernant l’origine raciale ou ethnique d’une personne physique.
- **Référence Santé** : Données concernant la santé d’une personne physique.
- **Codes médicaux ICD-9-CM** : Codes utilisés dans le secteur médical et de la santé.
- **Codes médicaux ICD-10-CM** : Codes utilisés dans le secteur médical et de la santé.
- **Référence aux croyances philosophiques** : Données concernant les croyances philosophiques d’une personne physique.
- **Référence aux opinions politiques** : Données concernant les opinions politiques d’une personne physique.
- **Référence aux croyances religieuses** : Données concernant les croyances religieuses d’une personne physique.
- **Référence relative à la vie sexuelle ou à l’orientation sexuelle** : Données concernant la vie sexuelle ou l’orientation sexuelle d’une personne physique.

Types de catégories

La classification des données catégorise vos données comme suit.

La plupart de ces catégories peuvent être reconnues en anglais, en allemand et en espagnol.

Catégorie	Type	English	Allemand	Espagnol
Finance	Bilans	✓	✓	✓
	Bons de commande	✓	✓	✓
	Factures	✓	✓	✓
	Rapports trimestriels	✓	✓	✓
HEURE	Vérifications des antécédents	✓		✓
	Plans de rémunération	✓	✓	✓
	Contrats de travail	✓		✓
	Avis des employés	✓		✓
	Santé	✓		✓
	CV	✓	✓	✓
Légal	Accords de confidentialité	✓	✓	✓
	Contrats fournisseur-client	✓	✓	✓
Commercialisation	Campagnes	✓	✓	✓
	Conférences	✓	✓	✓
Opérations	Rapports d’audit	✓	✓	✓
Ventes	Commandes de vente	✓	✓	

Catégorie	Type	English	Allemand	Espagnol
Services	RFI	✓		✓
	Demande de propositions	✓		✓
	TRUIE	✓	✓	✓
	Formation	✓	✓	✓
Support	Plaintes et contraventions	✓	✓	✓

Les métadonnées suivantes sont également catégorisées et identifiées dans les mêmes langues prises en charge :

- Données d'application
- Fichiers d'archives
- Audio
- Fil d'Ariane des données d'application métier de classification des données
- Fichiers CAO
- Code
- Corrompu
- Base de données et fichiers d'index
- Fichiers de conception
- Données de candidature par courrier électronique
- Crypté (fichiers avec un score d'entropie élevé)
- Exécutables
- Données d'application financière
- Données d'application de santé
- Images
- Journaux
- Documents divers
- Présentations diverses
- Feuilles de calcul diverses
- Divers « Inconnu »
- Fichiers protégés par mot de passe
- Données structurées
- Vidéos
- Fichiers de zéro octet

Types de fichiers

La classification des données analyse tous les fichiers pour obtenir des informations sur les catégories et les métadonnées et affiche tous les types de fichiers dans la section types de fichiers du tableau de bord. Lorsque la classification des données détecte des informations personnelles identifiables (PII) ou lorsqu'elle effectue

une recherche DSAR, seuls les formats de fichiers suivants sont pris en charge :

`.CSV, .DCM, .DOC, .DOCX, .JSON, .PDF, .PPTX, .RTF, .TXT, .XLS, .XLSX, Docs, Sheets, and Slides`

Exactitude des informations trouvées

NetApp ne peut pas garantir l'exactitude à 100 % des données personnelles et des données personnelles sensibles identifiées par Data Classification. Vous devez toujours valider les informations en examinant les données.

Sur la base de nos tests, le tableau ci-dessous montre l'exactitude des informations trouvées par Data Classification. Nous le décomposons par *précision* et *rappel* :

Précision
: La probabilité que ce que la classification des données trouve ait été correctement identifié. Par exemple, un taux de précision de 90 % pour les données personnelles signifie que 9 fichiers sur 10 identifiés comme contenant des informations personnelles contiennent réellement des informations personnelles. 1 fichier sur 10 serait un faux positif.

Rappel
: La probabilité que la classification des données trouve ce qu'elle devrait. Par exemple, un taux de rappel de 70 % pour les données personnelles signifie que la classification des données peut identifier 7 fichiers sur 10 qui contiennent réellement des informations personnelles dans votre organisation. La classification des données manquerait de 30 % des données et elles n'apparaîtraient pas dans le tableau de bord.

Nous améliorons constamment la précision de nos résultats. Ces améliorations seront automatiquement disponibles dans les futures versions de la classification des données.

Type	Précision	Rappel
Données personnelles - Généralités	90%-95%	60%-80%
Données personnelles - Identifiants de pays	30%-60%	40%-60%
Données personnelles sensibles	80%-95%	20%-30%
Catégories	90%-97%	60%-80%

Créer une classification personnalisée dans NetApp Data Classification

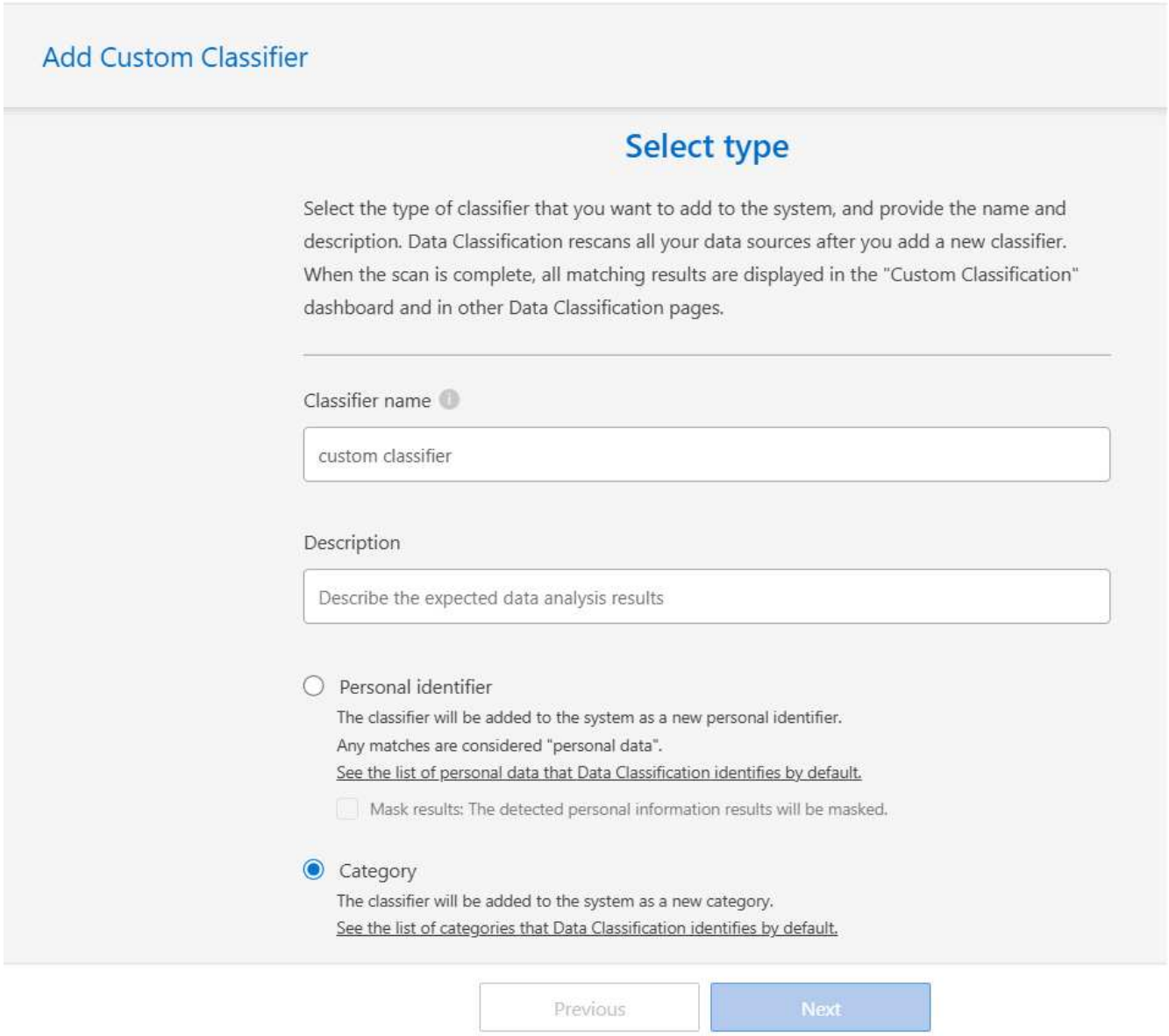
Avec NetApp Data Classification, vous pouvez créer une recherche personnalisée pour les informations sensibles. La recherche peut être limitée à une expression régulière (regex).

Créer une classification personnalisée

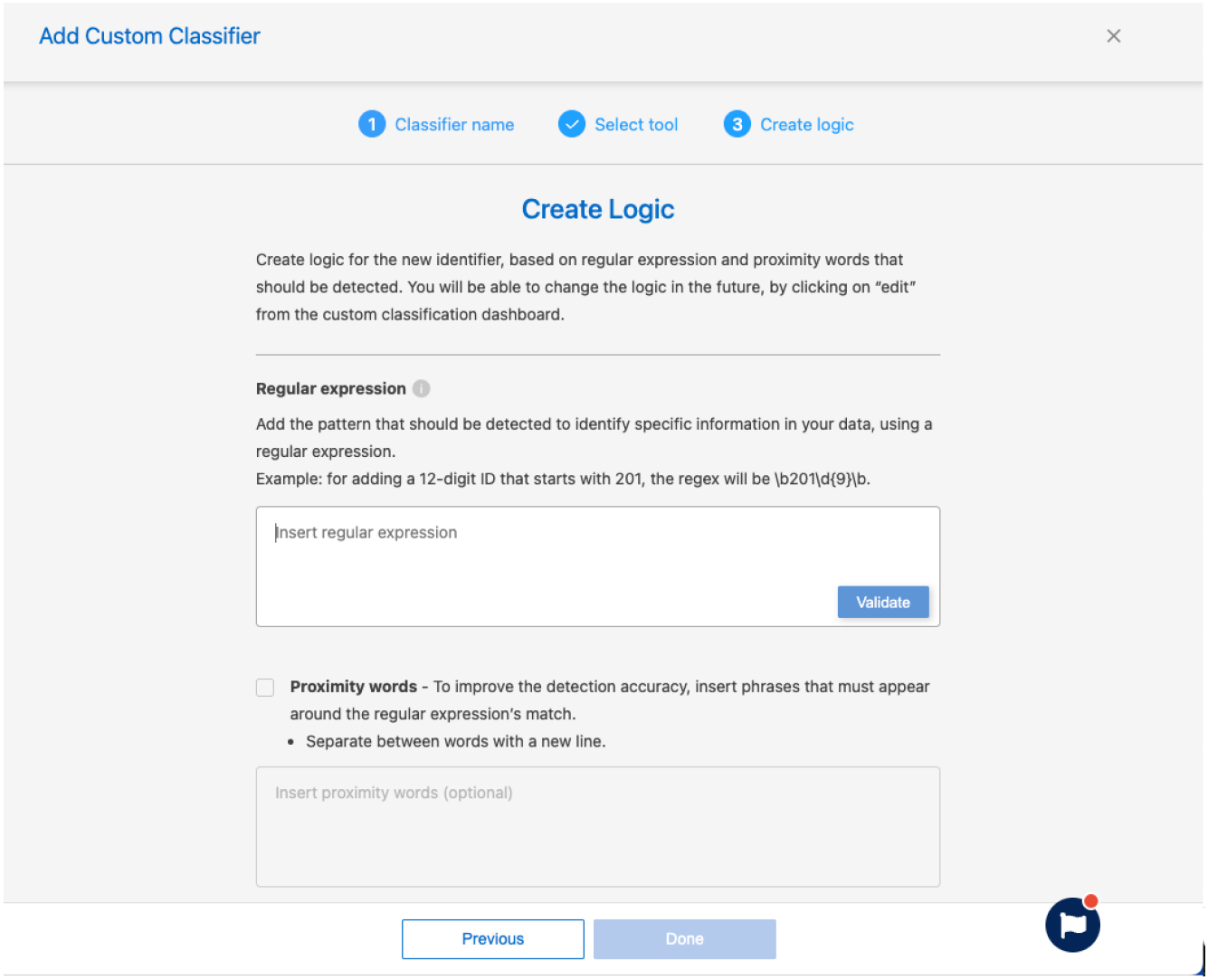
La classification personnalisée n'est disponible que pour les analyses Map & Classify, et non pour les analyses de cartographie uniquement. Cette fonctionnalité est actuellement en version préliminaire.

Étapes

1. Sélectionnez l'onglet **Classification personnalisée**.
2. Sélectionnez le bouton **Ajouter un nouveau classificateur**.
3. Ajoutez un nom de classificateur et une description pour le nouveau classificateur.
4. Choisissez d'ajouter le classificateur en tant qu'**identifiant personnel** ou **catégorie**.



5. Sélectionnez **Suivant**.
6. Pour ajouter la personnalisation sous forme d'expression régulière, sélectionnez **Expression régulière personnalisée** puis **Suivant**.
7. Ajoutez un modèle pour détecter les informations spécifiques de vos données. Sélectionnez **Valider** pour confirmer la syntaxe de votre
entrée.



8. Sélectionnez **Terminé** pour créer la classification personnalisée.

La nouvelle personnalisation est capturée lors de la prochaine analyse planifiée. Pour voir les résultats, voirGénérer des rapports de conformité .

Examinez les données stockées dans votre organisation avec la NetApp Data Classification

Le tableau de bord d'investigation des données affiche des informations au niveau des fichiers et des répertoires sur vos données, vous permettant de trier et de filtrer les résultats. La page Enquête sur les données présente des informations sur les métadonnées et les autorisations des fichiers et des répertoires, ainsi que sur l'identification des fichiers en double. Grâce aux informations au niveau des fichiers, des répertoires et des bases de données, vous pouvez prendre des mesures pour améliorer la conformité de votre organisation et économiser de l'espace de stockage. La page Enquête sur les données prend également en charge le déplacement, la copie et la suppression de fichiers.

Pour obtenir des informations à partir de la page Enquête, vous devez effectuer une analyse de classification complète sur vos sources de données. Les sources de données ayant fait l'objet d'une analyse de mappage uniquement n'affichent pas les détails au niveau du fichier.

Structure d'enquête sur les données

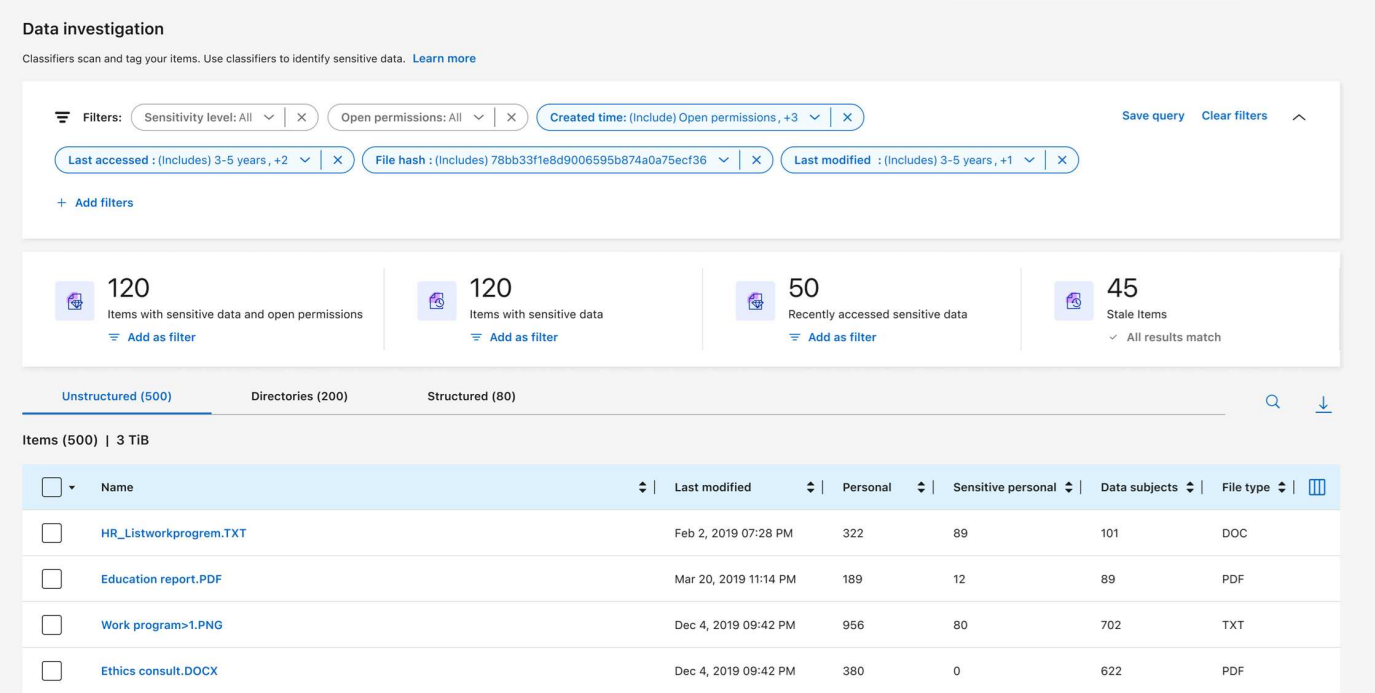
La page Enquête sur les données trie les données en trois onglets :

- **Données non structurées** : données de fichier
- **Répertoires** : dossiers et partages de fichiers
- **Structuré** : base de données

Filtres de données

La page Enquête sur les données fournit de nombreux filtres pour trier vos données afin que vous puissiez obtenir ce dont vous avez besoin. Vous pouvez utiliser plusieurs filtres de concert.

Pour ajouter un filtre, sélectionnez le bouton **Ajouter un filtre**.



Sensibilité et contenu du filtre

Utilisez les filtres suivants pour afficher la quantité d'informations sensibles contenues dans vos données.

Filtre	Détails
Catégorie	Sélectionnez le"types de catégories" .
Niveau de sensibilité	Sélectionnez le niveau de sensibilité : Personnel, Personnel sensible ou Non sensible.

Filtre	Détails
Nombre d'identifiants	Sélectionnez la plage d'identifiants sensibles détectés par fichier. Comprend les données personnelles et les données personnelles sensibles. Lors du filtrage dans les répertoires, la classification des données totalise les correspondances de tous les fichiers de chaque dossier (et sous-dossiers). REMARQUE : la version de décembre 2023 (version 1.26.6) a supprimé l'option permettant de calculer le nombre de données d'informations personnelles identifiables (PII) par répertoires.
Données personnelles	Sélectionnez le"types de données personnelles" .
Données personnelles sensibles	Sélectionnez le"types de données personnelles sensibles" .
Personne concernée	Saisissez le nom complet ou l'identifiant connu de la personne concernée. "En savoir plus sur les personnes concernées ici" .

Filtrer le propriétaire de l'utilisateur et les autorisations de l'utilisateur

Utilisez les filtres suivants pour afficher les propriétaires de fichiers et les autorisations d'accès à vos données.

Filtre	Détails
Autorisations d'ouverture	Sélectionnez le type d'autorisations dans les données et dans les dossiers/partages.
Autorisations utilisateur/groupe	Sélectionnez un ou plusieurs noms d'utilisateur et/ou noms de groupe, ou saisissez un nom partiel.
Propriétaire du fichier	Entrez le nom du propriétaire du fichier.
Nombre d'utilisateurs avec accès	Sélectionnez une ou plusieurs plages de catégories pour afficher les fichiers et dossiers ouverts à un certain nombre d'utilisateurs.

Filtrer chronologiquement

Utilisez les filtres suivants pour afficher les données en fonction de critères temporels.

Filtre	Détails
Temps créé	Sélectionnez une plage horaire pendant laquelle le fichier a été créé. Vous pouvez également spécifier une plage horaire personnalisée pour affiner davantage les résultats de la recherche.
Temps découvert	Sélectionnez une plage horaire pendant laquelle la classification des données a découvert le fichier. Vous pouvez également spécifier une plage horaire personnalisée pour affiner davantage les résultats de la recherche.
Dernière modification	Sélectionnez une plage horaire pendant laquelle le fichier a été modifié pour la dernière fois. Vous pouvez également spécifier une plage horaire personnalisée pour affiner davantage les résultats de la recherche.

Filtre	Détails
Dernier accès	Sélectionnez une plage horaire pendant laquelle le fichier ou le répertoire* a été consulté pour la dernière fois. Vous pouvez également spécifier une plage horaire personnalisée pour affiner davantage les résultats de la recherche. Pour les types de fichiers analysés par Data Classification, il s'agit de la dernière fois que Data Classification a analysé le fichier.

* L'heure du dernier accès à un répertoire n'est disponible que pour les partages NFS ou CIFS.

Filtrer les métadonnées

Utilisez les filtres suivants pour afficher les données en fonction de l'emplacement, de la taille et du type de répertoire ou de fichier.

Filtre	Détails
Chemin du fichier	Saisissez jusqu'à 20 chemins partiels ou complets que vous souhaitez inclure ou exclure de la requête. Si vous entrez à la fois des chemins d'inclusion et des chemins d'exclusion, Data Classification recherche d'abord tous les fichiers dans les chemins inclus, puis supprime les fichiers des chemins exclus, puis affiche les résultats. Notez que l'utilisation de « * » dans ce filtre n'a aucun effet et que vous ne pouvez pas exclure des dossiers spécifiques de l'analyse : tous les répertoires et fichiers sous un partage configuré seront analysés.
Type de répertoire	Sélectionnez le type de répertoire ; « Partager » ou « Dossier ».
Type de fichier	Sélectionnez le"types de fichiers" .
Taille du fichier	Sélectionnez la plage de taille du fichier.
Hachage de fichier	Saisissez le hachage du fichier pour rechercher un fichier spécifique, même si le nom est différent.

Type de stockage du filtre

Utilisez les filtres suivants pour afficher les données par type de stockage.

Filtre	Détails
Type de système	Sélectionnez le type de système.
Nom de l'environnement système	Sélectionnez des systèmes spécifiques.
Référentiel de stockage	Sélectionnez le référentiel de stockage, par exemple un volume ou un schéma.

Requête de filtrage

Utilisez le filtre suivant pour afficher les données par requêtes enregistrées.

Filtre	Détails
Requête enregistrée	Sélectionnez une requête enregistrée ou plusieurs. Aller à la"onglet requêtes enregistrées" pour afficher la liste des requêtes enregistrées existantes et en créer de nouvelles.
Mots-clés	Sélectionner"le tag ou les tags" qui sont attribués à vos fichiers.

Statut de l'analyse du filtre

Utilisez le filtre suivant pour afficher les données en fonction de l'état d'analyse de la classification des données.

Filtre	Détails
État de l'analyse	Sélectionnez une option pour afficher la liste des fichiers en attente de première analyse, en cours d'analyse, en attente de nouvelle analyse ou dont l'analyse a échoué.
Événement d'analyse d'analyse	Sélectionnez si vous souhaitez afficher les fichiers qui n'ont pas été classés parce que la classification des données n'a pas pu revenir à l'heure du dernier accès, ou les fichiers qui ont été classés même si la classification des données n'a pas pu revenir à l'heure du dernier accès.

"Voir les détails sur l'horodatage « dernier accès »"pour plus d'informations sur les éléments qui apparaissent dans la page Investigation lors du filtrage à l'aide de l'événement d'analyse d'analyse.

Filtrer les données par doublons

Utilisez le filtre suivant pour afficher les fichiers dupliqués dans votre stockage.

Filtre	Détails
Doublons	Sélectionnez si le fichier est dupliqué dans les référentiels.

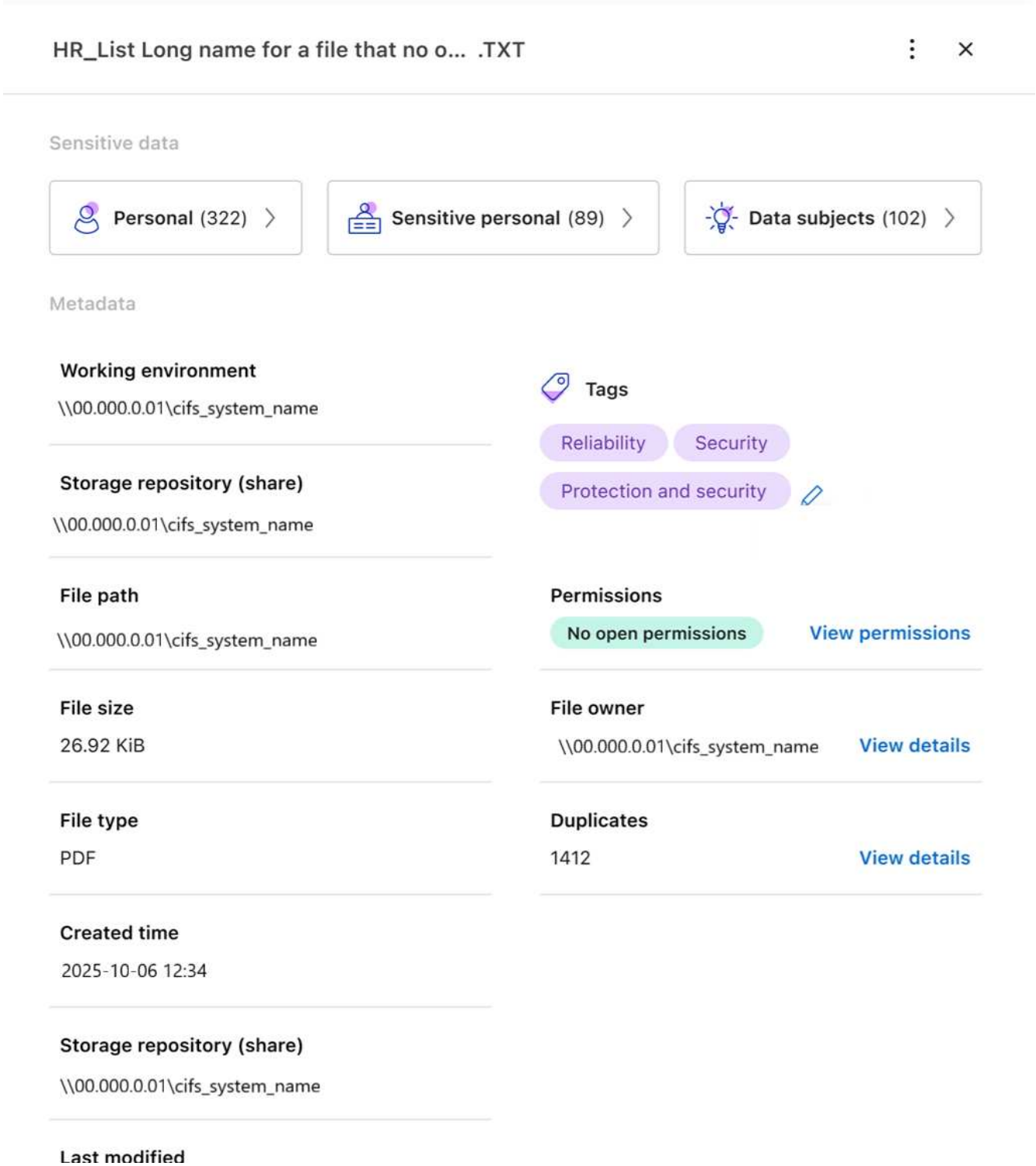
Afficher les métadonnées du fichier

En plus de vous montrer le système et le volume où réside le fichier, les métadonnées affichent beaucoup plus d'informations, notamment les autorisations du fichier, le propriétaire du fichier et s'il existe des doublons de ce fichier. Ces informations sont utiles si vous envisagez de"créer des requêtes enregistrées" car vous pouvez voir toutes les informations que vous pouvez utiliser pour filtrer vos données.

La disponibilité des informations dépend de la source des données. Par exemple, le nom du volume et les autorisations ne sont pas partagés pour les fichiers de base de données.

Étapes

1. Dans le menu Classification des données, sélectionnez **Enquête**.
2. Dans la liste Enquête sur les données à droite, sélectionnez le curseur vers le bas à droite pour n'importe quel fichier unique pour afficher les métadonnées du fichier.



3. En option, vous pouvez créer ou ajouter une balise au fichier avec le bouton **Créer une balise**. Sélectionnez une balise existante dans le menu déroulant ou ajoutez une nouvelle balise avec le bouton **+ Ajouter**. Les balises peuvent être utilisées pour filtrer les données.

Afficher les autorisations utilisateur pour les fichiers et les répertoires

Pour afficher une liste de tous les utilisateurs ou groupes ayant accès à un fichier ou à un répertoire et les types d'autorisations dont ils disposent, sélectionnez **Afficher toutes les autorisations**. Cette option est disponible uniquement pour les données dans les partages CIFS.

Si vous utilisez des identifiants de sécurité (SID) au lieu de noms d'utilisateur et de groupe, vous devez intégrer votre Active Directory dans la classification des données. Pour plus d'informations, consultez la section "ajouter Active Directory à la classification des données" .

Étapes

1. Dans le menu Classification des données, sélectionnez **Enquête**.
2. Dans la liste Enquête sur les données à droite, sélectionnez le curseur vers le bas à droite pour n'importe quel fichier unique pour afficher les métadonnées du fichier.
3. Pour afficher une liste de tous les utilisateurs ou groupes ayant accès à un fichier ou à un répertoire et les types d'autorisations dont ils disposent, dans le champ Autorisations d'ouverture, sélectionnez **Afficher toutes les autorisations**.

   La classification des données affiche jusqu'à 100 utilisateurs dans la liste.

4. Sélectionnez le curseur vers le bas bouton pour n'importe quel groupe pour voir la liste des utilisateurs qui font partie du groupe.

   Vous pouvez développer un niveau du groupe pour voir les utilisateurs qui font partie du groupe.

5. Sélectionnez le nom d'un utilisateur ou d'un groupe pour actualiser la page Enquête afin de voir tous les fichiers et répertoires auxquels l'utilisateur ou le groupe a accès.

Vérifiez les fichiers en double dans vos systèmes de stockage

Vous pouvez vérifier si des fichiers en double sont stockés dans vos systèmes de stockage. Ceci est utile si vous souhaitez identifier les zones dans lesquelles vous pouvez économiser de l'espace de stockage. Il est également bon de s'assurer que certains fichiers disposant d'autorisations spécifiques ou d'informations sensibles ne sont pas inutilement dupliqués dans vos systèmes de stockage.

Tous vos fichiers (à l'exception des bases de données) de 1 Mo ou plus, ou contenant des informations personnelles ou sensibles, sont comparés pour voir s'il existe des doublons.

La classification des données utilise la technologie de hachage pour déterminer les fichiers en double. Si un fichier possède le même code de hachage qu'un autre fichier, vous pouvez être sûr à 100 % que les fichiers sont des doublons exacts, même si les noms de fichiers sont différents.

Étapes

1. Dans le menu Classification des données, sélectionnez **Enquête**.
2. Dans le volet Filtre, sélectionnez « Taille du fichier » ainsi que « Doublons » (« Contient des doublons ») pour voir quels fichiers d'une certaine plage de taille sont dupliqués dans votre environnement.
3. Vous pouvez également télécharger la liste des fichiers en double et l'envoyer à votre administrateur de stockage afin qu'il puisse décider quels fichiers, le cas échéant, peuvent être supprimés.
4. En option, vous pouvez supprimer, étiqueter ou déplacer les fichiers en double. Sélectionnez les fichiers sur lesquels vous souhaitez effectuer une action, puis sélectionnez l'action appropriée.

Voir si un fichier spécifique est dupliqué

Vous pouvez voir si un seul fichier contient des doublons.

Étapes

1. Dans le menu Classification des données, sélectionnez **Enquête**.
2. Dans la liste Enquête sur les données, sélectionnez ⌄ à droite pour n'importe quel fichier unique pour afficher les métadonnées du fichier.

   Si des doublons existent pour un fichier, cette information apparaît à côté du champ *Doublons*.

3. Pour afficher la liste des fichiers en double et leur emplacement, sélectionnez **Afficher les détails**.
4. Sur la page suivante, sélectionnez **Afficher les doublons** pour afficher les fichiers dans la page Enquête.
5. En option, vous pouvez supprimer, étiqueter ou déplacer les fichiers en double. Sélectionnez les fichiers sur lesquels vous souhaitez effectuer une action, puis sélectionnez l'action appropriée.

Vous pouvez utiliser la valeur « hachage de fichier » fournie dans cette page et la saisir directement dans la page Enquête pour rechercher un fichier en double spécifique à tout moment - ou vous pouvez l'utiliser dans une requête enregistrée.

Téléchargez votre rapport

Vous pouvez télécharger vos résultats filtrés au format CSV ou JSON.

Il peut y avoir jusqu'à trois fichiers de rapport téléchargés si la classification des données analyse des fichiers (données non structurées), des répertoires (dossiers et partages de fichiers) et des bases de données (données structurées).

Les fichiers sont divisés en fichiers avec un nombre fixe de lignes ou d'enregistrements :

- JSON : 100 000 enregistrements par rapport dont la génération prend environ 5 minutes
- CSV : 200 000 enregistrements par rapport dont la génération prend environ 4 minutes

Vous pouvez télécharger une version du fichier CSV à visualiser dans ce navigateur. Cette version est limitée à 10 000 enregistrements.

Ce qui est inclus dans le rapport téléchargeable

Le **Rapport de données sur les fichiers non structurés** inclut les informations suivantes sur vos fichiers :

- Nom des fichiers
- Type d'emplacement
- Nom du système
- Référentiel de stockage (par exemple, un volume, un bucket, des partages)
- Type de référentiel
- Chemin du fichier
- Type de fichier
- Taille du fichier (en Mo)
- Temps de création
- Dernière modification
- Dernier accès

- Propriétaire du fichier
  - Les données du propriétaire du fichier englobent le nom du compte, le nom du compte SAM et l'adresse e-mail lorsque Active Directory est configuré.
- Catégorie
- Informations personnelles
- Informations personnelles sensibles
- Autorisations ouvertes
- Erreur d'analyse de numérisation
- Date de détection de suppression

  La date de détection de suppression identifie la date à laquelle le fichier a été supprimé ou déplacé. Cela vous permet d'identifier quand des fichiers sensibles ont été déplacés. Les fichiers supprimés ne contribuent pas au nombre de fichiers qui apparaît dans le tableau de bord ou sur la page Enquête. Les fichiers n'apparaissent que dans les rapports CSV.

Le **Rapport de données sur les répertoires non structurés** inclut les informations suivantes sur vos dossiers et partages de fichiers :

- Type de système
- Nom du système
- Nom du répertoire
- Référentiel de stockage (par exemple, un dossier ou des partages de fichiers)
- Propriétaire du répertoire
- Temps de création
- Temps découvert
- Dernière modification
- Dernier accès
- Autorisations ouvertes
- Type de répertoire

Le **rapport de données structurées** inclut les informations suivantes sur vos tables de base de données :

- Nom de la table de base de données
- Type d'emplacement
- Nom du système
- Référentiel de stockage (par exemple, un schéma)
- Nombre de colonnes
- Nombre de lignes
- Informations personnelles
- Informations personnelles sensibles

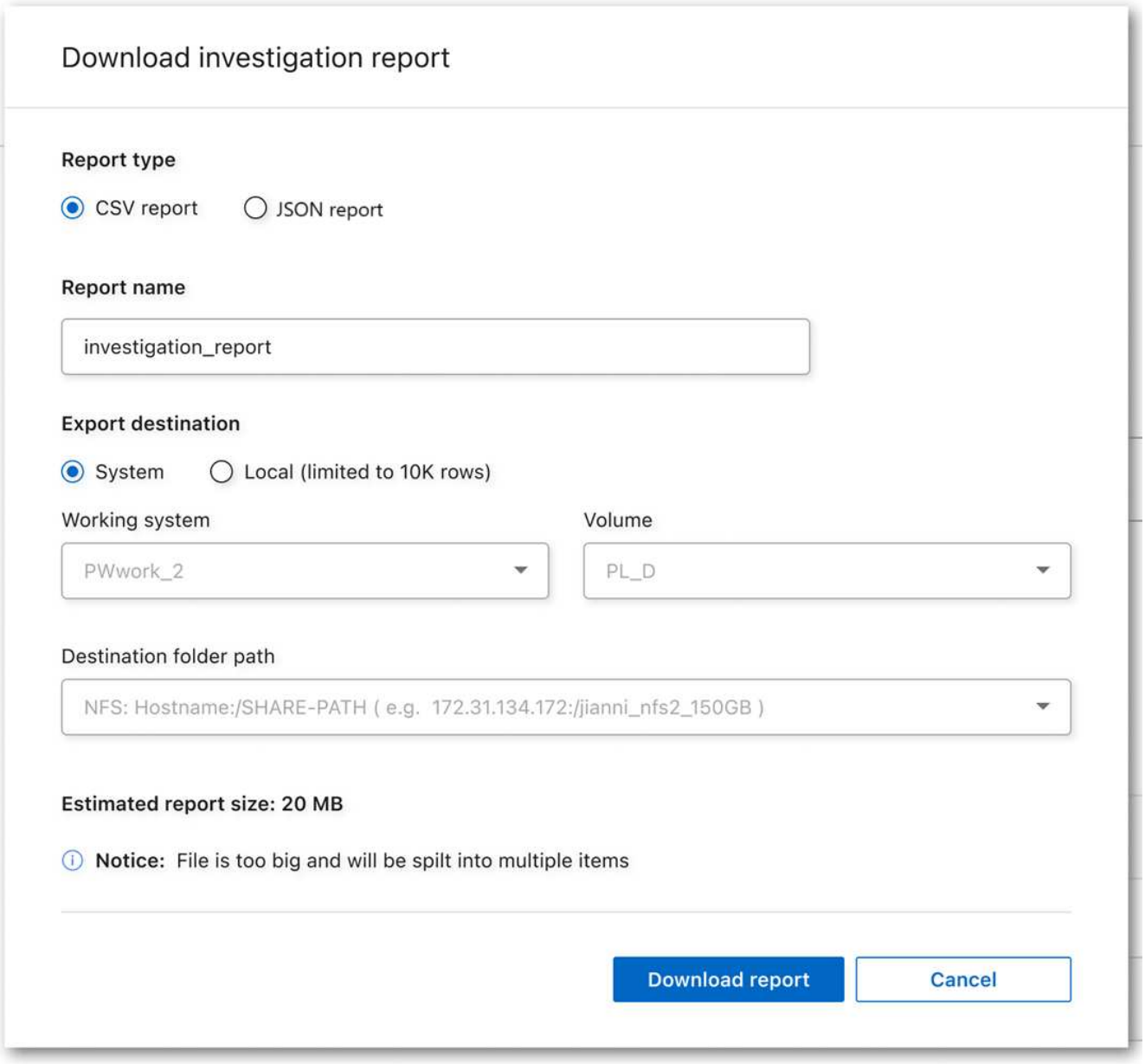
Étapes pour générer le rapport

1. À partir de la page Enquête sur les données, sélectionnez l'option bouton en haut à droite de la page.

2. Choisissez le type de rapport : CSV ou JSON.
3. Saisissez un **Nom de rapport**.
4. Pour télécharger le rapport complet, sélectionnez **Système** puis choisissez **Système** et **Volume** dans les menus déroulants respectifs. Fournissez un **chemin d'accès au dossier de destination**.

   Pour télécharger le rapport dans le navigateur, sélectionnez **Local** . Notez que cette option limite le rapport aux 10 000 premières lignes et est limitée au format **CSV**. Vous n'avez pas besoin de remplir d'autres champs si vous sélectionnez **Local**.

5. Sélectionnez **Télécharger le rapport**.



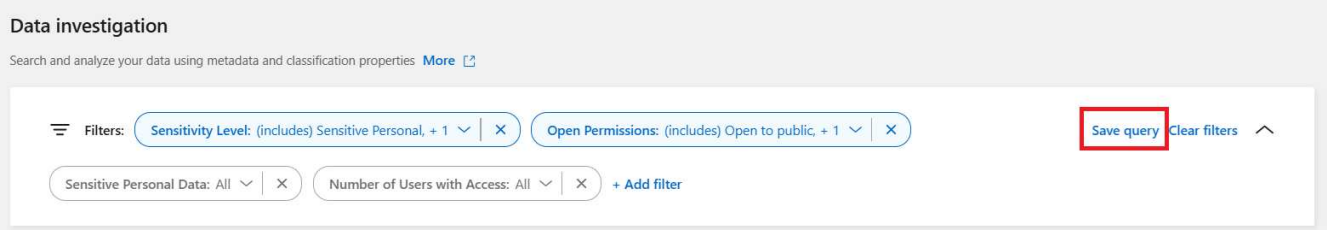
Résultat

Une boîte de dialogue affiche un message indiquant que les rapports sont en cours de téléchargement.

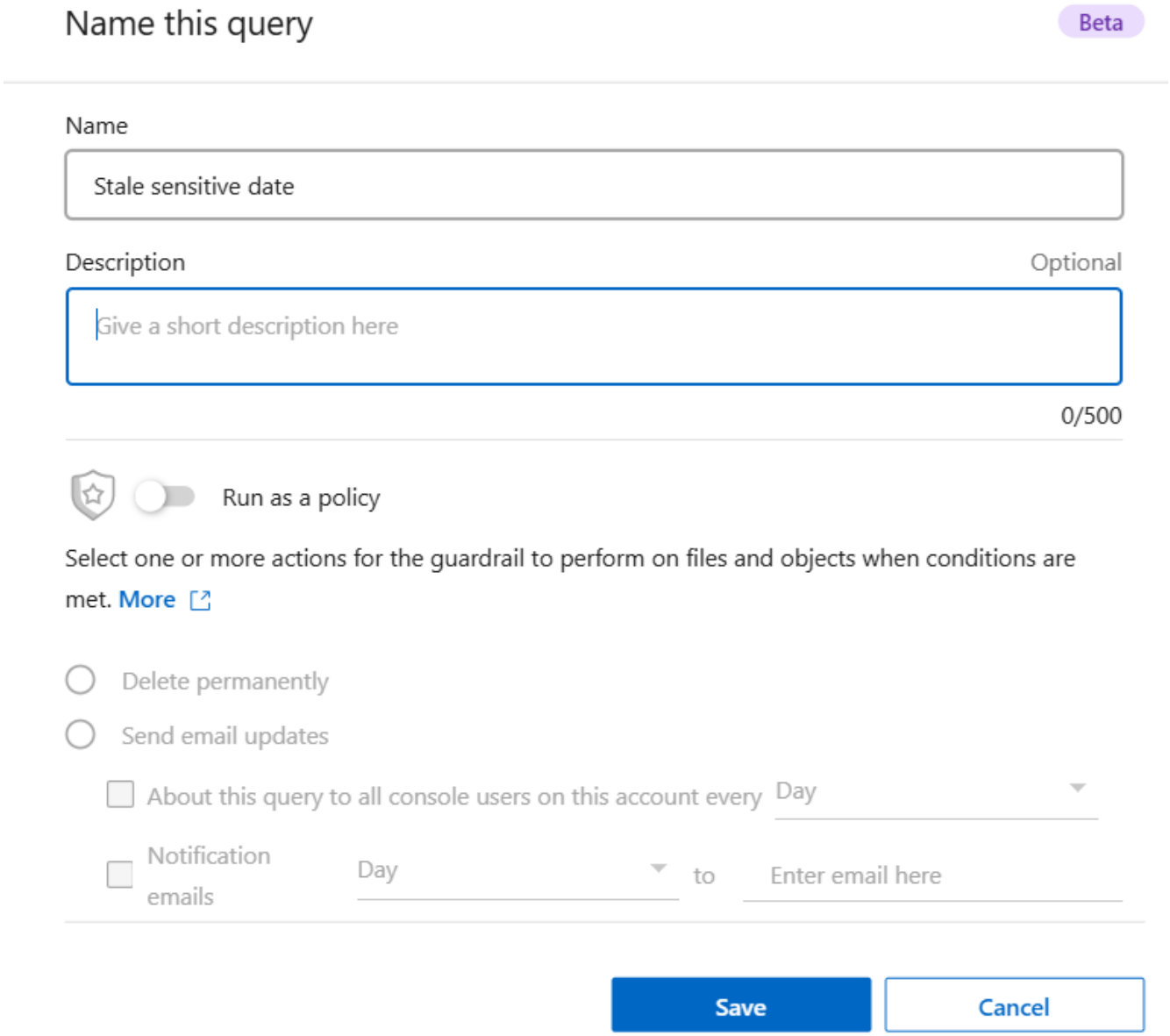
Créer une requête enregistrée en fonction des filtres sélectionnés

Étapes

1. Dans l'onglet Enquête, définissez une recherche en sélectionnant les filtres que vous souhaitez utiliser. Voir"Filtrage des données dans la page Investigation" pour plus de détails.
2. Une fois que vous avez défini toutes les caractéristiques du filtre à votre guise, sélectionnez **Enregistrer la requête**.



3. Nommez la requête enregistrée et ajoutez une description. Le nom doit être unique.
4. Vous pouvez éventuellement enregistrer la requête en tant que politique :
   a. Pour enregistrer la requête en tant que politique, activez le bouton **Exécuter en tant que politique**.
   b. Choisissez de **Supprimer définitivement** ou **Envoyer des mises à jour par e-mail**. Si vous choisissez les mises à jour par e-mail, vous pouvez envoyer les résultats de la requête par e-mail à *tous* les utilisateurs de la console quotidiennement, hebdomadairement ou mensuellement. Alternativement, vous pouvez envoyer la notification à une adresse e-mail spécifique aux mêmes fréquences.
5. Sélectionnez **Enregistrer**.



Une fois la recherche ou la politique créée, vous pouvez la visualiser dans l'onglet **Requêtes enregistrées**.

L'affichage des résultats sur la page Requêtes enregistrées peut prendre jusqu'à 15 minutes.

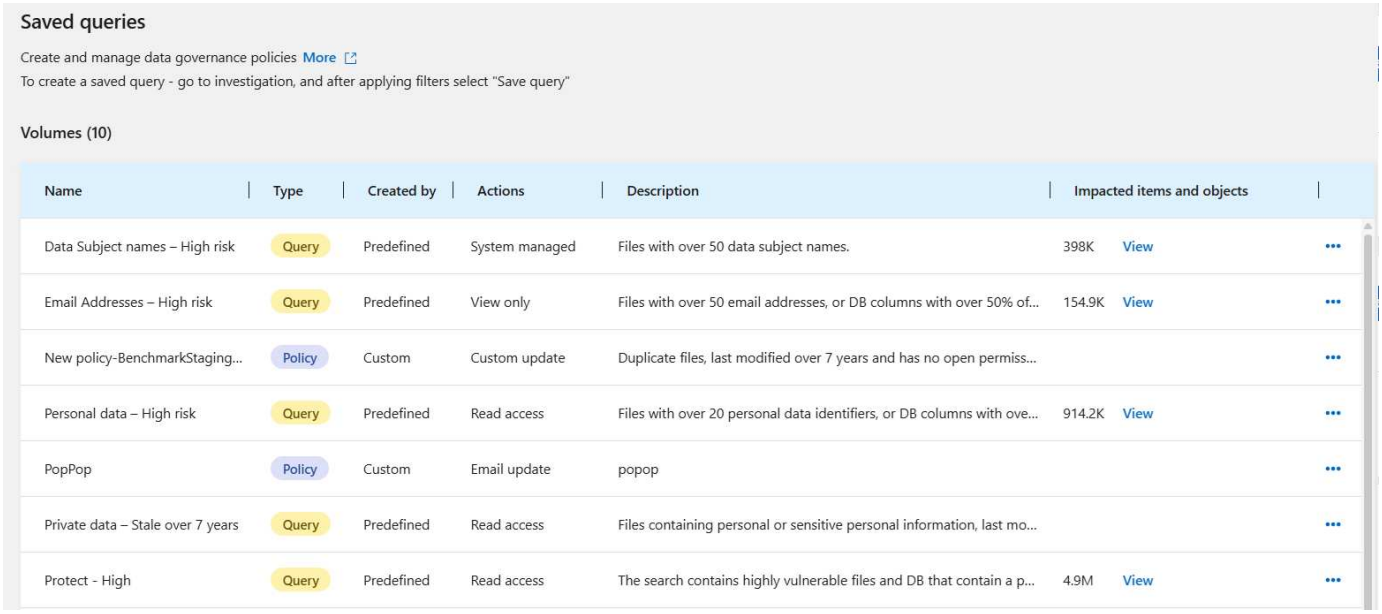
Gérer les requêtes enregistrées avec la NetApp Data Classification

NetaApp Data Classification prend en charge l'enregistrement de vos requêtes de recherche. Avec une requête enregistrée, vous pouvez créer des filtres personnalisés pour trier les requêtes fréquentes de votre page d'enquête sur les données. La classification des données comprend également des requêtes enregistrées prédéfinies basées sur des demandes courantes.

L'onglet **Requêtes enregistrées** du tableau de bord Conformité répertorie toutes les requêtes enregistrées prédéfinies et personnalisées disponibles sur cette instance de classification des données.

Les requêtes enregistrées peuvent également être enregistrées en tant que **politiques**. Alors que les requêtes filtrent les données, les politiques vous permettent d'agir sur les données. Avec une politique : vous pouvez supprimer les données découvertes ou envoyer des mises à jour par e-mail sur les données découvertes.

Les requêtes enregistrées apparaissent également dans la liste des filtres de la page Enquête.



Afficher les résultats des requêtes enregistrées dans la page Enquête

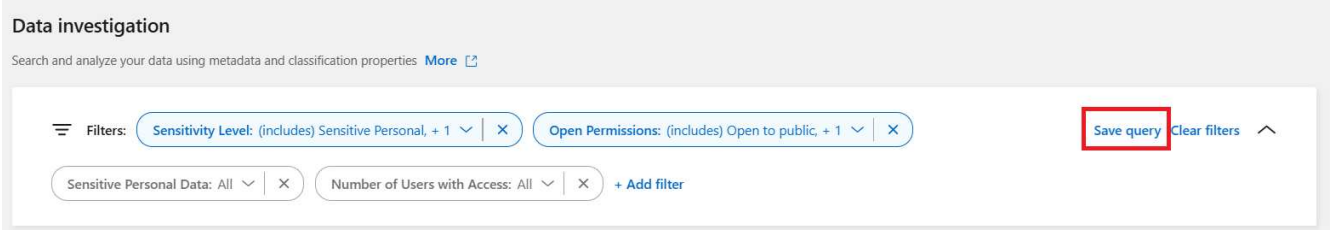
Pour afficher les résultats d'une requête enregistrée dans la page Investigation, sélectionnez l'icône bouton pour une recherche spécifique puis sélectionnez **Enquêter sur les résultats**.

Créer des requêtes et des politiques enregistrées

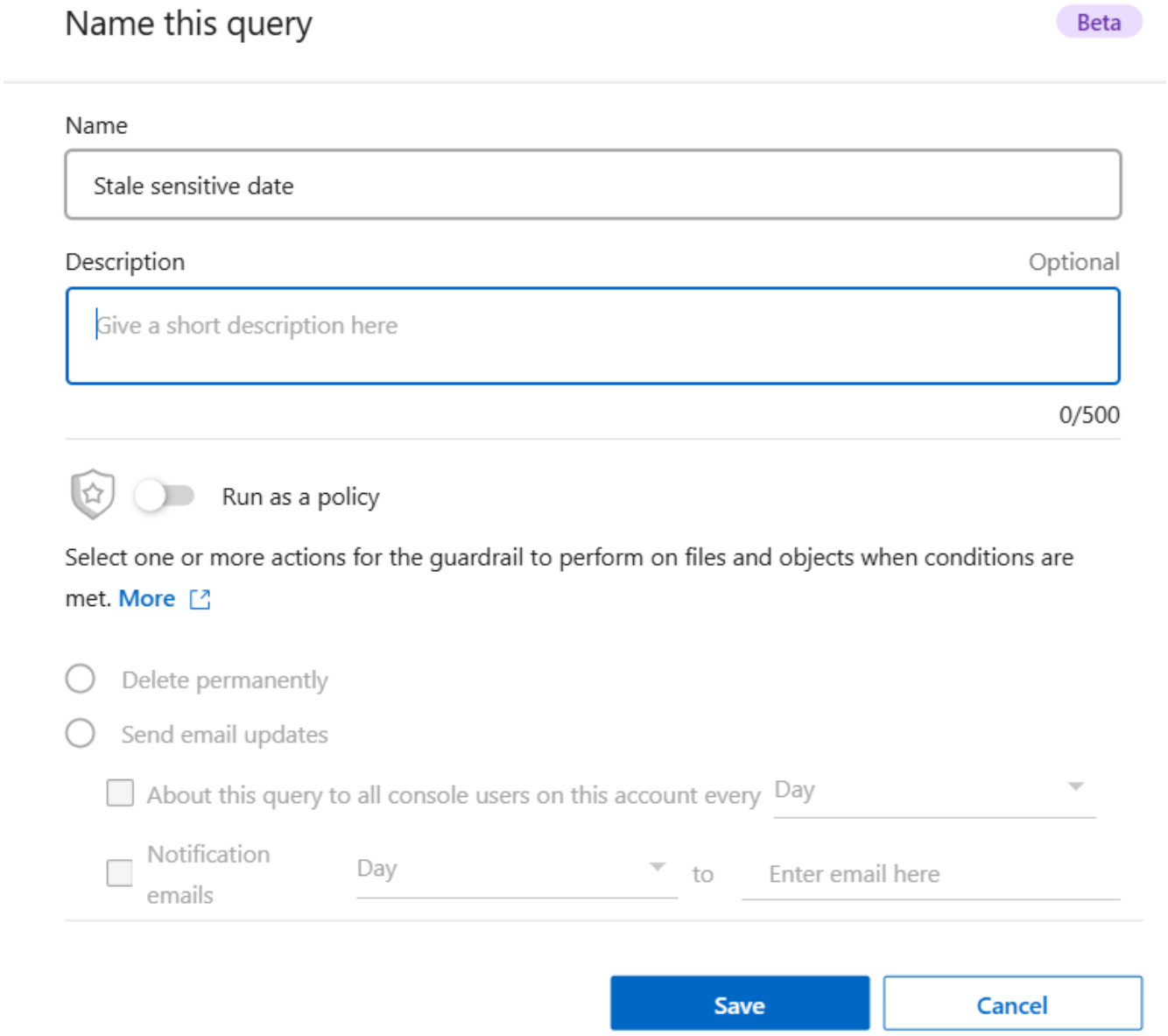
Vous pouvez créer vos propres requêtes enregistrées personnalisées qui fournissent des résultats pour des requêtes spécifiques à votre organisation. Les résultats sont renvoyés pour tous les fichiers et répertoires (partages et dossiers) qui correspondent aux critères de recherche.

Étapes

1. Dans l'onglet Enquête, définissez une recherche en sélectionnant les filtres que vous souhaitez utiliser. Voir"Filtrage des données dans la page Investigation" pour plus de détails.
2. Une fois que vous avez défini toutes les caractéristiques du filtre à votre guise, sélectionnez **Enregistrer la requête**.



3. Nommez la requête enregistrée et ajoutez une description. Le nom doit être unique.
4. Vous pouvez éventuellement enregistrer la requête en tant que politique :
   a. Pour enregistrer la requête en tant que politique, activez le bouton **Exécuter en tant que politique**.
   b. Choisissez de **Supprimer définitivement** ou **Envoyer des mises à jour par e-mail**. Si vous choisissez les mises à jour par e-mail, vous pouvez envoyer les résultats de la requête par e-mail à *tous* les utilisateurs de la console quotidiennement, hebdomadairement ou mensuellement. Alternativement, vous pouvez envoyer la notification à une adresse e-mail spécifique aux mêmes fréquences.
5. Sélectionnez **Enregistrer**.



Une fois la recherche ou la politique créée, vous pouvez la visualiser dans l'onglet **Requêtes enregistrées**.

Modifier les requêtes ou les politiques enregistrées

Vous pouvez modifier le nom et la description d'une requête enregistrée. Vous pouvez également convertir une requête en politique et vice versa.

Vous ne pouvez pas modifier les requêtes enregistrées par défaut. Vous ne pouvez pas modifier les filtres d'une requête enregistrée. Vous pouvez alternativement afficher les résultats de l'enquête d'une requête enregistrée, modifier les filtres, puis l'enregistrer en tant que nouvelle requête ou politique.

Étapes

1. Depuis la page Requêtes enregistrées, sélectionnez **Modifier la recherche** pour la recherche que vous souhaitez modifier.

2. Apportez les modifications aux champs nom et description. Pour modifier uniquement les champs nom et description.

   Vous pouvez éventuellement convertir la requête en politique ou convertir la politique en requête enregistrée. Activez le bouton **Exécuter en tant que politique** selon vos besoins. .. Si vous convertissez la requête en politique, choisissez **Supprimer définitivement** ou **Envoyer des mises à jour par e-mail**. Si vous choisissez les mises à jour par e-mail, vous pouvez envoyer les résultats de la requête par e-mail à *tous* les utilisateurs de la console quotidiennement, hebdomadairement ou mensuellement. Alternativement, vous pouvez envoyer la notification à une adresse e-mail spécifique aux mêmes fréquences.

3. Sélectionnez **Enregistrer** pour terminer les modifications.

Supprimer les requêtes enregistrées

Vous pouvez supprimer toute requête ou politique personnalisée enregistrée si vous n'en avez plus besoin. Vous ne pouvez pas supprimer les requêtes enregistrées par défaut.

Pour supprimer une requête enregistrée, sélectionnez l'icône bouton pour une recherche spécifique, sélectionnez **Supprimer la requête**, puis sélectionnez à nouveau **Supprimer la requête** dans la boîte de dialogue de confirmation.

Requêtes par défaut

La classification des données fournit les requêtes de recherche définies par le système suivantes :

- **Noms des personnes concernées - Risque élevé**

  Fichiers contenant plus de 50 noms de personnes concernées

- **Adresses e-mail - Risque élevé**

  Fichiers contenant plus de 50 adresses e-mail ou colonnes de base de données dont plus de 50 % de leurs lignes contiennent des adresses e-mail

- **Données personnelles - Risque élevé**

  Fichiers contenant plus de 20 identifiants de données personnelles ou colonnes de base de données dont plus de 50 % de leurs lignes contiennent des identifiants de données personnelles

- **Données privées - Obsolètes depuis plus de 7 ans**

  Fichiers contenant des informations personnelles ou sensibles, modifiés pour la dernière fois il y a plus de 7 ans

- **Protéger - Élevé**

  Fichiers ou colonnes de base de données contenant un mot de passe, des informations de carte de crédit,

un numéro IBAN ou un numéro de sécurité sociale

- **Protéger - Faible**

  Fichiers qui n'ont pas été consultés depuis plus de 3 ans

- **Protéger - Moyen**

  Fichiers contenant des fichiers ou des colonnes de base de données avec des identifiants de données personnelles, notamment des numéros d'identification, des numéros d'identification fiscale, des numéros de permis de conduire, des identifiants médicaux ou des numéros de passeport

- **Données personnelles sensibles - Risque élevé**

  Fichiers contenant plus de 20 identifiants de données personnelles sensibles ou colonnes de base de données dont plus de 50 % de leurs lignes contiennent des données personnelles sensibles

Modifier les paramètres d'analyse de NetApp Data Classification pour vos référentiels

Vous pouvez gérer la manière dont vos données sont analysées dans chacun de vos systèmes et sources de données. Vous pouvez effectuer les modifications sur une base de « référentiel » ; ce qui signifie que vous pouvez effectuer des modifications pour chaque volume, schéma, utilisateur, etc. en fonction du type de source de données que vous analysez.

Certaines des choses que vous pouvez modifier sont si un référentiel est analysé ou non et si NetApp Data Classification effectue une"scan de cartographie ou scan de cartographie et de classification" . Vous pouvez également suspendre et reprendre l'analyse, par exemple, si vous devez arrêter l'analyse d'un volume pendant un certain temps.

Afficher l'état de l'analyse de vos référentiels

Vous pouvez afficher les référentiels individuels que NetApp Data Classification analyse (volumes, buckets, etc.) pour chaque système et source de données. Vous pouvez également voir combien ont été « cartographiés » et combien ont été « classés ». La classification prend plus de temps car l'identification complète de l'IA est effectuée sur toutes les données.

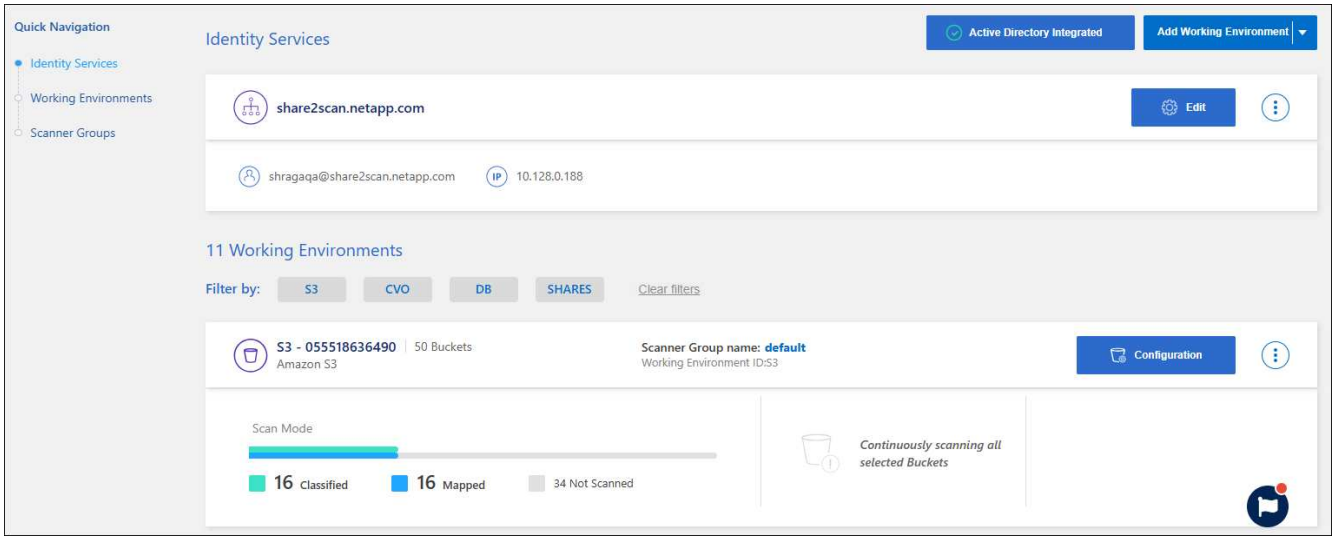
Vous pouvez afficher l'état de numérisation de chaque environnement de travail sur la page de configuration :

- **Initialisation** (point bleu clair) : La configuration de la carte ou de la classification est activée. Cela apparaît pendant quelques secondes avant de démarrer le statut « file d'attente en attente ».
- **File d'attente en attente** (point orange) : la tâche d'analyse attend d'être répertoriée dans la file d'attente d'analyse.
- **En file d'attente** (point orange) : la tâche a été ajoutée avec succès à la file d'attente d'analyse. Le système commencera à cartographier ou à classer le volume lorsque son tour dans la file d'attente arrivera.
- **En cours d'exécution** (point vert) : la tâche d'analyse, qui était dans la file d'attente, est en cours d'exécution sur le référentiel de stockage sélectionné.
- **Terminé** (point vert) : L'analyse du référentiel de stockage est terminée.

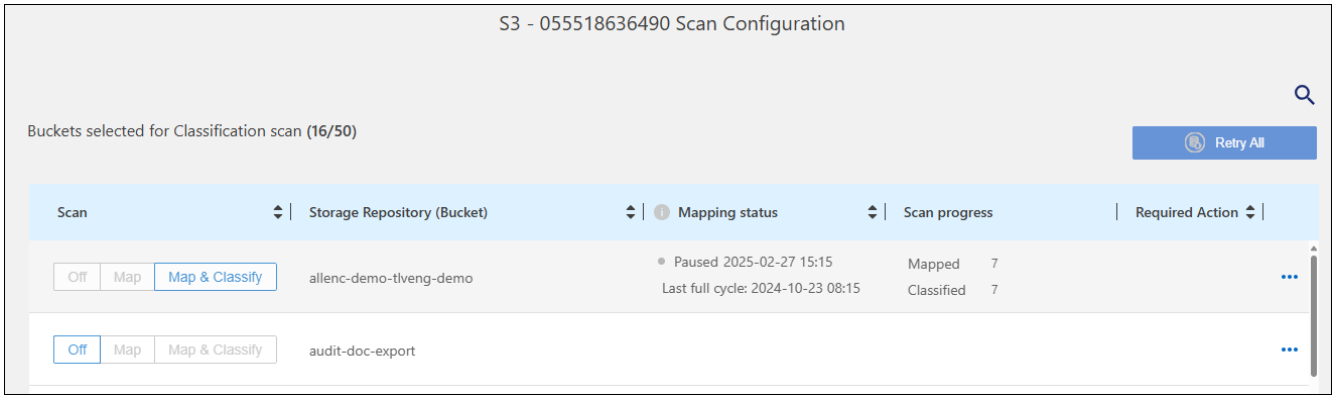
- **En pause** (point gris) : vous avez sélectionné l'option « Pause » pour interrompre la numérisation. Bien que les modifications de volume ne soient pas affichées dans le système, les informations analysées sont toujours affichées.
- **Erreur** (point rouge) : L'analyse ne peut pas se terminer car elle a rencontré des problèmes. Si vous devez effectuer une action, l'erreur apparaît dans l'info-bulle sous la colonne « Action requise ». Dans le cas contraire, le système affiche un statut « erreur » et tente de récupérer. Une fois terminé, le statut change.
- **Pas de numérisation** : la configuration du volume « Désactivé » a été sélectionnée et le système ne scanne pas le volume.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.



2. Dans l'onglet Configuration, sélectionnez le bouton **Configuration** pour le système.
3. Dans la page Configuration de l'analyse, affichez les paramètres d'analyse pour tous les référentiels.



4. Passez votre curseur sur le graphique dans la colonne *Statut de mappage* pour voir le nombre de fichiers qui restent à mapper ou à classer dans chaque référentiel (bucket dans cet exemple).

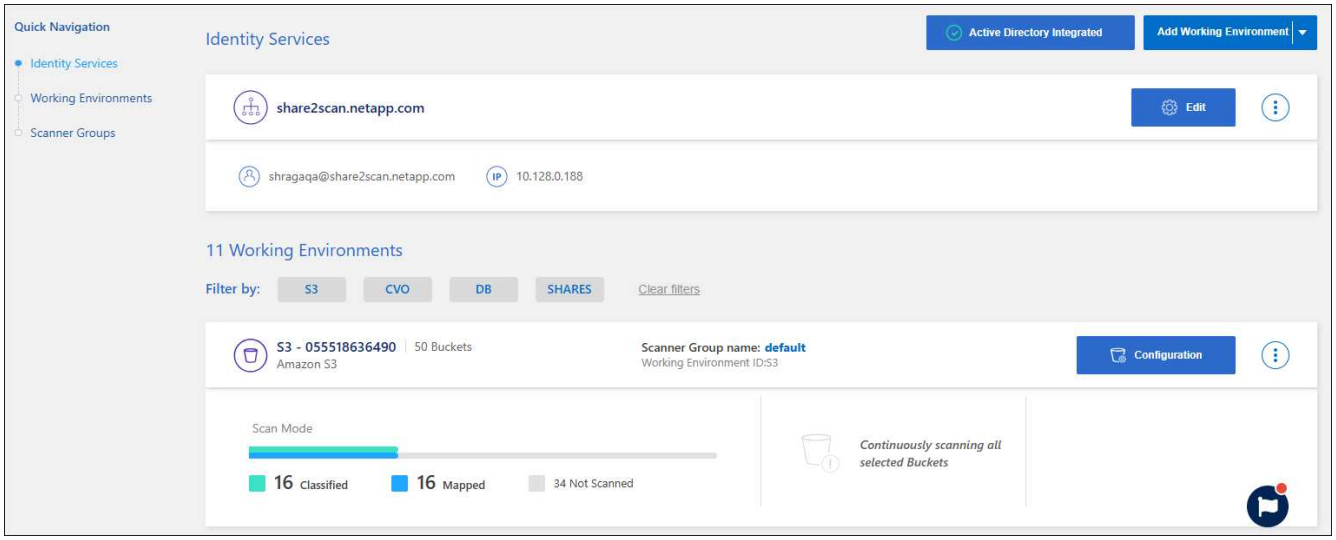
Modifier le type d'analyse d'un référentiel

Vous pouvez démarrer ou arrêter les analyses de mappage uniquement, ou les analyses de mappage et de classification, dans un système à tout moment à partir de la page Configuration. Vous pouvez également passer d'analyses de mappage uniquement à des analyses de mappage et de classification, et vice-versa.

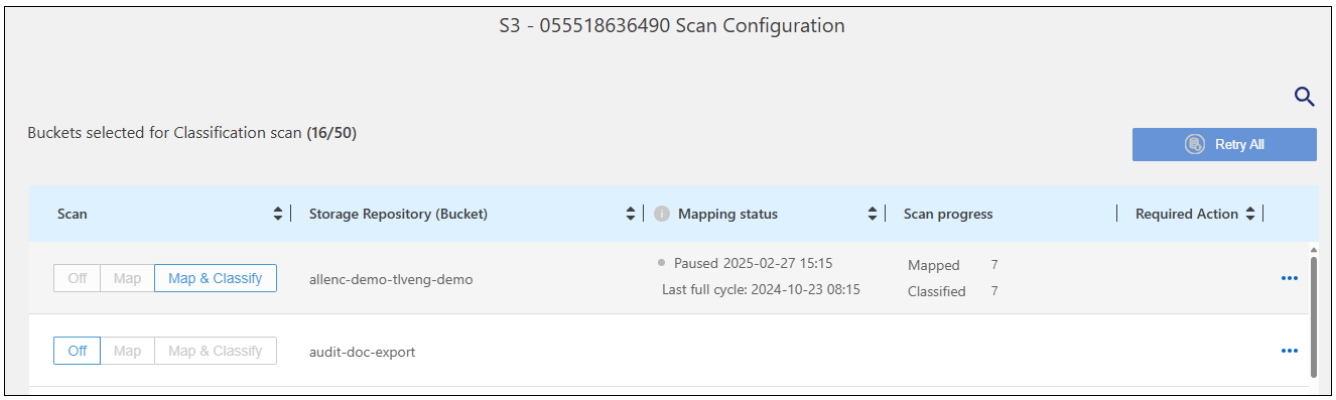
Les bases de données ne peuvent pas être configurées pour des analyses de mappage uniquement. L'analyse de la base de données peut être désactivée ou activée ; où activé est équivalent à Map & Classify.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Dans l'onglet Configuration, sélectionnez le bouton **Configuration** pour le système.

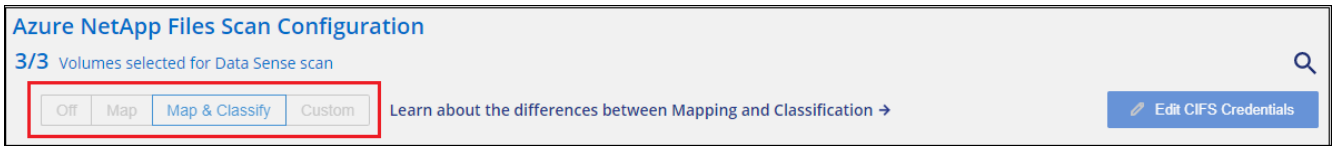


3. Dans la page Configuration de l'analyse, modifiez l'un des référentiels (compartiments dans cet exemple) pour effectuer des analyses **Map** ou **Map & Classify**.



Certains types de systèmes vous permettent de modifier le type d'analyse globalement pour tous les référentiels à l'aide d'une barre de boutons en haut de la page. Ceci est valable pour les systèmes Cloud Volumes ONTAP, ONTAP sur site, Azure NetApp Files et Amazon FSx pour ONTAP .

L'exemple ci-dessous montre cette barre de boutons pour un système Azure NetApp Files .



Prioriser les analyses

Vous pouvez prioriser les analyses de cartographie uniquement les plus importantes ou cartographier et classer les analyses pour garantir que les analyses hautement prioritaires sont effectuées en premier.

Par défaut, les analyses sont mises en file d'attente en fonction de l'ordre dans lequel elles sont lancées. Grâce à la possibilité de hiérarchiser les analyses, vous pouvez déplacer les analyses vers l'avant de la file d'attente. Plusieurs analyses peuvent être priorisées. La priorité est désignée selon un ordre premier entré, premier sorti, ce qui signifie que la première analyse que vous priorisez passe en tête de la file d'attente ; la deuxième analyse que vous priorisez devient la deuxième dans la file d'attente, et ainsi de suite.

La priorité est accordée une seule fois. Les réanalyses automatiques des données de cartographie se produisent dans l'ordre par défaut.

Étapes

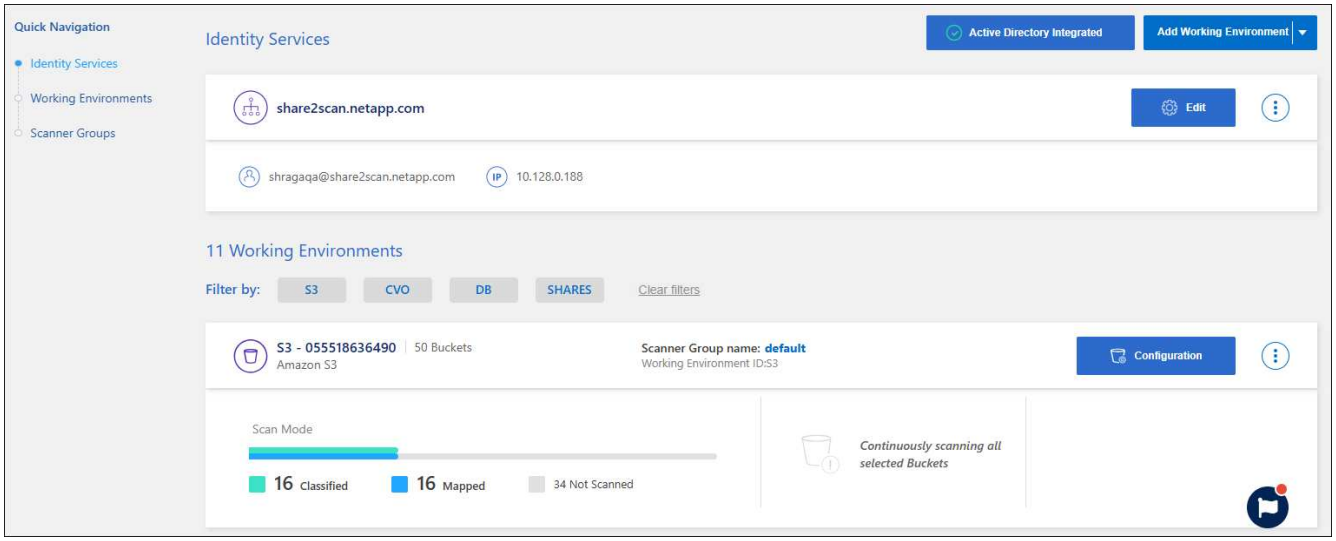
1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Sélectionnez les ressources que vous souhaitez prioriser.
3. Des actions … option, sélectionnez **Prioriser l'analyse**.

Arrêter la recherche d'un référentiel

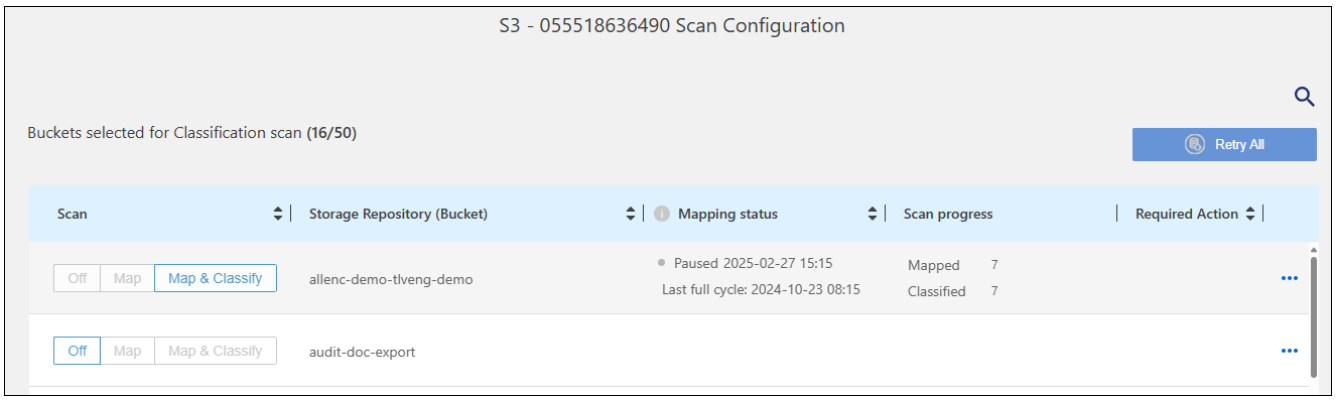
Vous pouvez arrêter l'analyse d'un référentiel (par exemple, un volume) si vous n'avez plus besoin de surveiller sa conformité. Vous pouvez le faire en désactivant la numérisation. Lorsque l'analyse est désactivée, toute l'indexation et les informations sur ce volume sont supprimées du système et la facturation de l'analyse des données est arrêtée.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Dans l'onglet Configuration, sélectionnez le bouton **Configuration** pour le système.



3. Dans la page Configuration de l'analyse, sélectionnez **Désactivé** pour arrêter l'analyse d'un compartiment particulier.



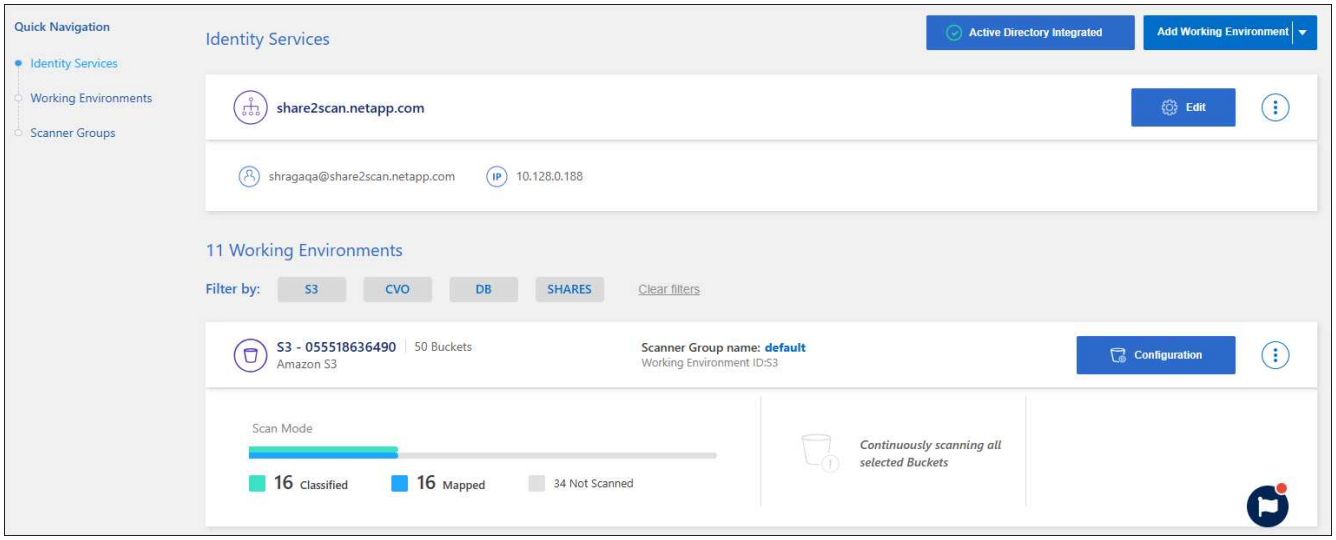
Mettre en pause et reprendre l'analyse d'un référentiel

Vous pouvez « suspendre » l'analyse d'un référentiel si vous souhaitez arrêter temporairement l'analyse de certains contenus. La suspension de l'analyse signifie que la classification des données n'effectuera aucune analyse future pour détecter les modifications ou les ajouts au référentiel, mais que tous les résultats actuels seront toujours affichés dans le système. La suspension de la numérisation n'arrête pas la facturation des données numérisées, car les données existent toujours.

Vous pouvez « reprendre » la numérisation à tout moment.

Étapes

1. Dans le menu Classification des données, sélectionnez **Configuration**.
2. Dans l'onglet Configuration, sélectionnez le bouton **Configuration** pour le système.



3. Dans la page Configuration de l'analyse, sélectionnez les actions ••• icône.
4. Sélectionnez **Pause** pour suspendre l'analyse d'un volume ou sélectionnez **Reprendre** pour reprendre l'analyse d'un volume qui avait été précédemment suspendu.

Afficher les rapports de conformité de la NetApp Data Classification

NetApp Data Classification fournit des rapports que vous pouvez utiliser pour mieux

comprendre l'état du programme de confidentialité des données de votre organisation.

Par défaut, les tableaux de bord de classification des données affichent les données de conformité et de gouvernance pour tous les systèmes, bases de données et sources de données. Si vous souhaitez afficher des rapports contenant des données pour certains systèmes uniquement, vous pouvez filtrer pour les voir uniquement.

- Les rapports de conformité ne sont disponibles que si vous effectuez une analyse de classification complète sur vos sources de données. Les sources de données ayant fait l'objet d'une analyse de mappage uniquement peuvent uniquement générer le rapport de mappage de données.
- NetApp ne peut pas garantir l'exactitude à 100 % des données personnelles et des données personnelles sensibles identifiées par Data Classification. Vous devez toujours valider les informations en examinant les données.

Les rapports suivants sont disponibles pour la classification des données :

- **Rapport d'évaluation de la découverte de données** : fournit une analyse de haut niveau de l'environnement analysé pour mettre en évidence les résultats du système et montrer les zones de préoccupation et les étapes de correction potentielles. Ce rapport est disponible dans le tableau de bord de gouvernance.
- **Rapport d'aperçu complet du mappage des données** : fournit des informations sur la taille et le nombre de fichiers dans vos systèmes. Cela inclut la capacité d'utilisation, l'âge des données, la taille des données et les types de fichiers. Ce rapport est disponible dans le tableau de bord de gouvernance.
- **Rapport de demande d'accès aux données personnelles** : vous permet d'extraire un rapport de tous les fichiers contenant des informations concernant le nom spécifique ou l'identifiant personnel d'une personne concernée. Ce rapport est disponible dans le tableau de bord Conformité.
- **Rapport HIPAA** : vous aide à identifier la répartition des informations de santé dans vos fichiers. Ce rapport est disponible dans le tableau de bord Conformité.
- **Rapport PCI DSS** : vous aide à identifier la répartition des informations de carte de crédit dans vos fichiers. Ce rapport est disponible dans le tableau de bord Conformité.
- **Rapport d'évaluation des risques liés à la confidentialité** : fournit des informations sur la confidentialité de vos données et un score de risque lié à la confidentialité. Ce rapport est disponible dans le tableau de bord Conformité.
- **Rapports sur un type d'informations spécifique** : Des rapports sont disponibles qui incluent des détails sur les fichiers identifiés contenant des données personnelles et des données personnelles sensibles. Vous pouvez également voir les fichiers classés par catégorie et par type de fichier.

Sélectionnez les systèmes pour les rapports

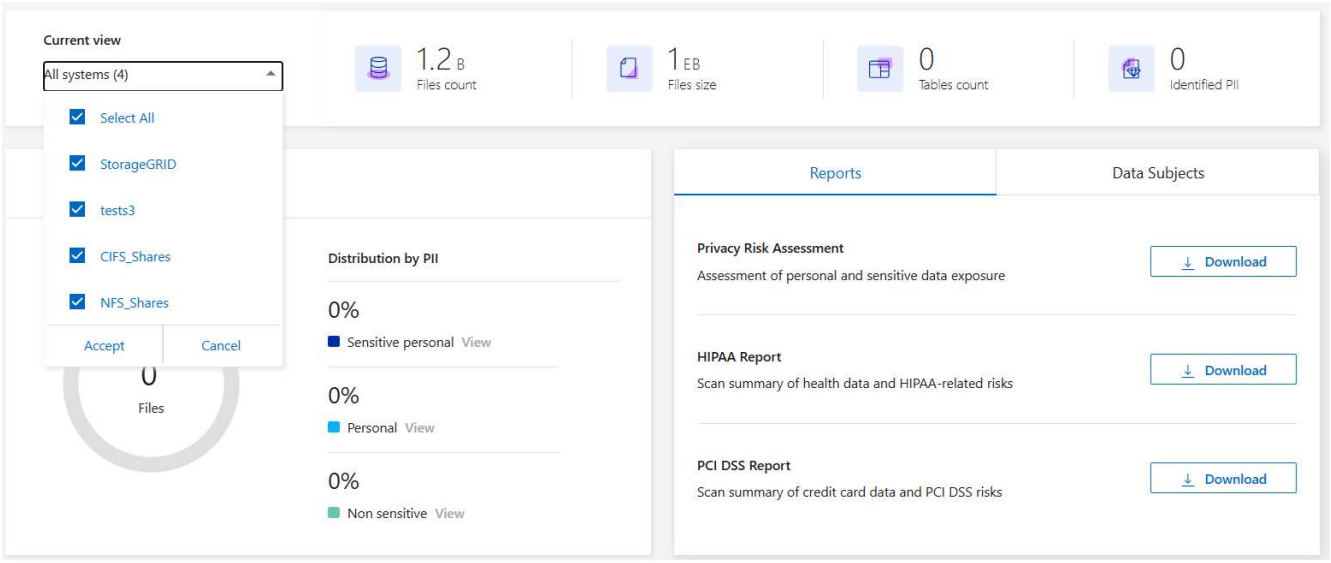
Vous pouvez filtrer le contenu du tableau de bord de conformité de la classification des données pour afficher les données de conformité pour tous les systèmes et bases de données, ou uniquement pour des systèmes spécifiques.

Lorsque vous filtrez le tableau de bord, la classification des données limite les données de conformité et les rapports uniquement aux systèmes que vous avez sélectionnés.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Sélectionnez la liste déroulante du filtre des systèmes, puis sélectionnez les systèmes.

3. Sélectionnez **Accepter** pour confirmer votre sélection.



Rapport de demande d'accès aux données personnelles

Les réglementations en matière de confidentialité telles que le RGPD européen accordent aux personnes concernées (telles que les clients ou les employés) le droit d'accéder à leurs données personnelles. Lorsqu'une personne concernée demande ces informations, on parle alors de demande d'accès aux données (DSAR). Les organisations sont tenues de répondre à ces demandes « sans retard injustifié », et au plus tard dans un délai d'un mois à compter de leur réception.

Vous pouvez répondre à un DSAR en recherchant le nom complet d'un sujet ou un identifiant connu (comme une adresse e-mail), puis en téléchargeant un rapport. Le rapport est conçu pour aider votre organisation à se conformer au RGPD ou à des lois similaires sur la confidentialité des données.

Comment la classification des données peut-elle vous aider à répondre à une DSAR ?

Lorsque vous effectuez une recherche sur une personne concernée, la classification des données recherche tous les fichiers contenant le nom ou l'identifiant de cette personne. La classification des données vérifie les dernières données pré-indexées pour le nom ou l'identifiant. Il ne lance pas de nouvelle analyse.

Une fois la recherche terminée, vous pouvez télécharger la liste des fichiers pour un rapport de demande d'accès aux données personnelles. Le rapport rassemble les informations issues des données et les traduit en termes juridiques que vous pouvez renvoyer à la personne.

La recherche de personnes concernées n'est actuellement pas prise en charge dans les bases de données.

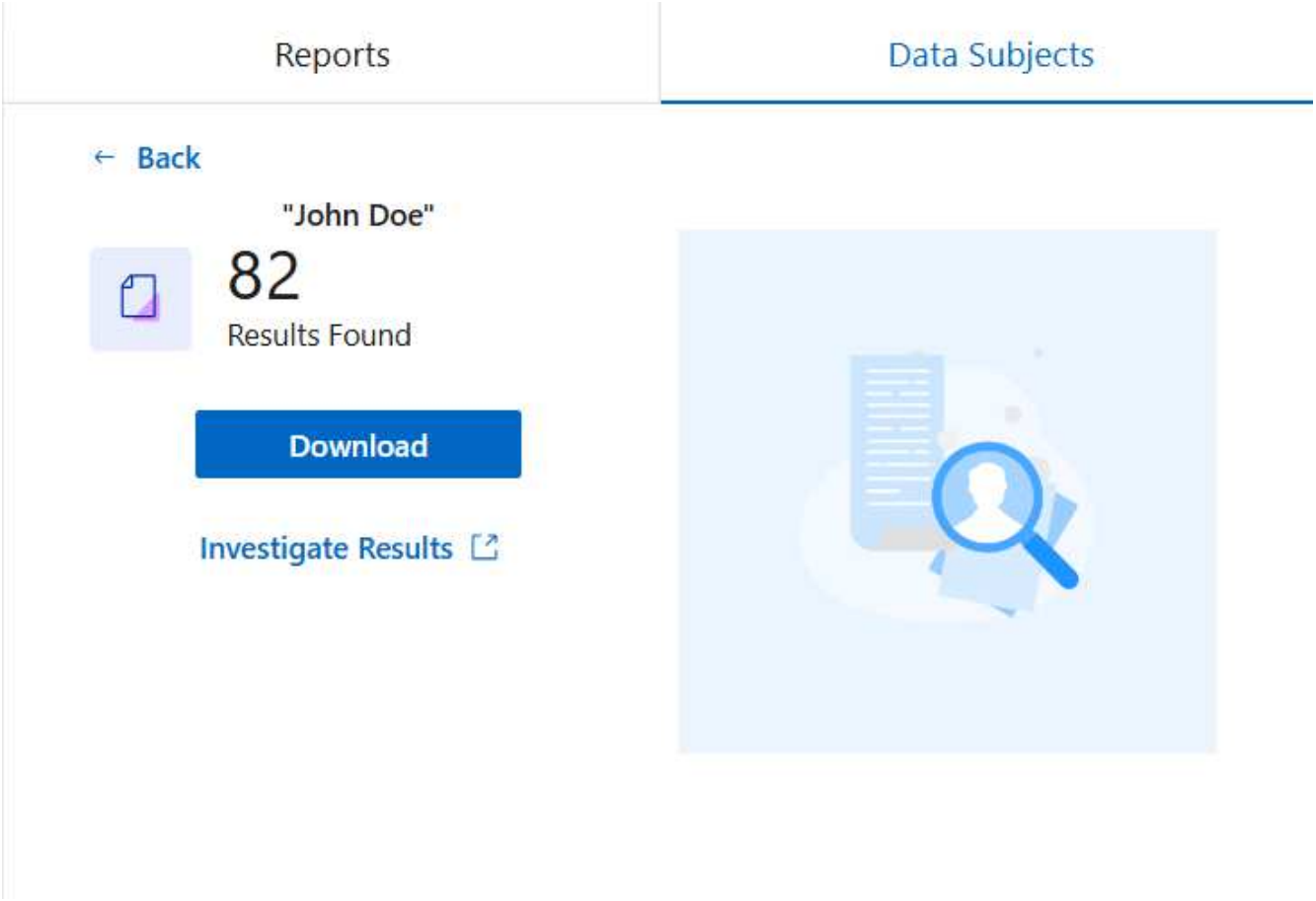
Rechercher des personnes concernées et télécharger des rapports

Recherchez le nom complet ou l'identifiant connu de la personne concernée, puis téléchargez un rapport de liste de fichiers ou un rapport DSAR. Vous pouvez rechercher par"tout type d'informations personnelles" .

L'anglais, l'allemand, le japonais et l'espagnol sont pris en charge lors de la recherche des noms des personnes concernées. La prise en charge de davantage de langues sera ajoutée ultérieurement.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Depuis la page Conformité, recherchez l'onglet **Personnes concernées**.
3. Dans la section **Personnes concernées**, saisissez un nom ou un identifiant connu, puis sélectionnez **Rechercher**.
4. Une fois la recherche terminée, sélectionnez **Télécharger** pour accéder à la réponse à la demande d'accès de la personne concernée. Sélectionnez **Enquêter sur les résultats** pour afficher plus d'informations sur la page Enquête sur les données.



5. Consultez les résultats dans la classification des données ou téléchargez-les sous forme de rapport en sélectionnant l'icône de téléchargement.

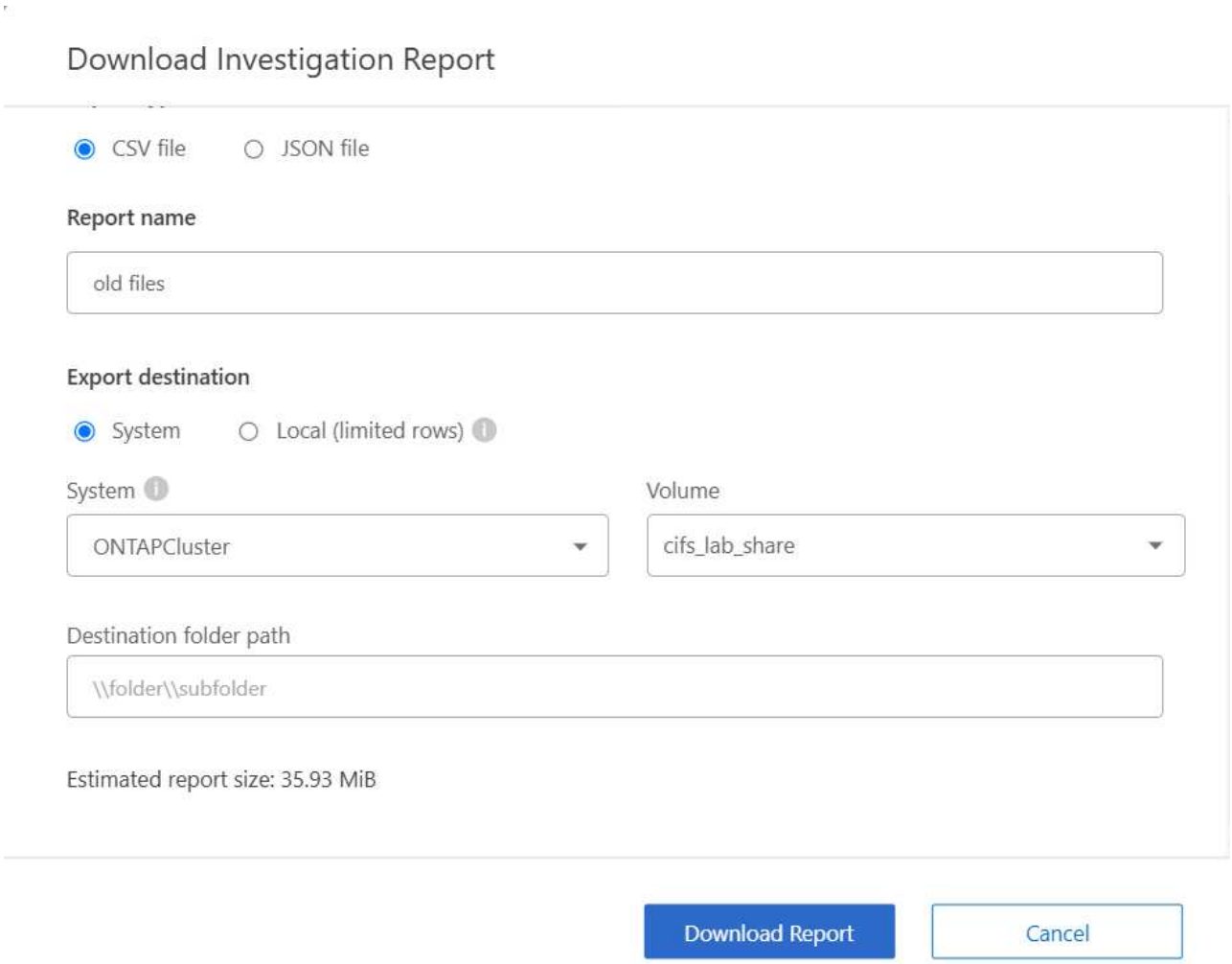
   a. Lorsque vous sélectionnez l'icône de téléchargement, configurez vos paramètres de téléchargement :

      - Choisissez le format du film : CSV ou JSON
      - Saisissez un **Nom du rapport**
      - Choisissez la destination d'exportation : **Système** ou votre machine **locale**.

        Si vous choisissez le système, toutes les données sont téléchargées. Vous devez également sélectionner le **Système**, le **Volume** et le **Chemin du dossier de destination**.

   Si vous choisissez **Local**, le rapport est limité aux 10 000 premières lignes de données non structurées, 5 000 lignes de données non structurées et 1 000 lignes de données structurées.

   a. Sélectionnez **Télécharger le rapport** pour lancer le téléchargement.



Rapport sur la loi sur la portabilité et la responsabilité en matière d'assurance maladie (HIPAA)

Le rapport sur la loi sur la portabilité et la responsabilité en matière d'assurance maladie (HIPAA) peut vous aider à identifier les fichiers contenant des informations sur la santé. Il est conçu pour aider votre organisation à se conformer aux lois sur la confidentialité des données HIPAA. Les informations recherchées par la classification des données comprennent :

- Modèle de référence de santé
- Code médical CIM-10-CM
- Code médical CIM-9-CM
- RH - Catégorie Santé
- Catégorie de données d'application de santé

Le rapport comprend les informations suivantes :

- Aperçu : Combien de fichiers contiennent des informations sur la santé et dans quels systèmes.
- Cryptage : pourcentage de fichiers contenant des informations sur la santé qui se trouvent sur des systèmes cryptés ou non cryptés. Ces informations sont spécifiques à Cloud Volumes ONTAP.
- Protection contre les ransomwares : pourcentage de fichiers contenant des informations sur la santé qui se

trouvent sur des systèmes sur lesquels la protection contre les ransomwares est activée ou non. Ces informations sont spécifiques à Cloud Volumes ONTAP.

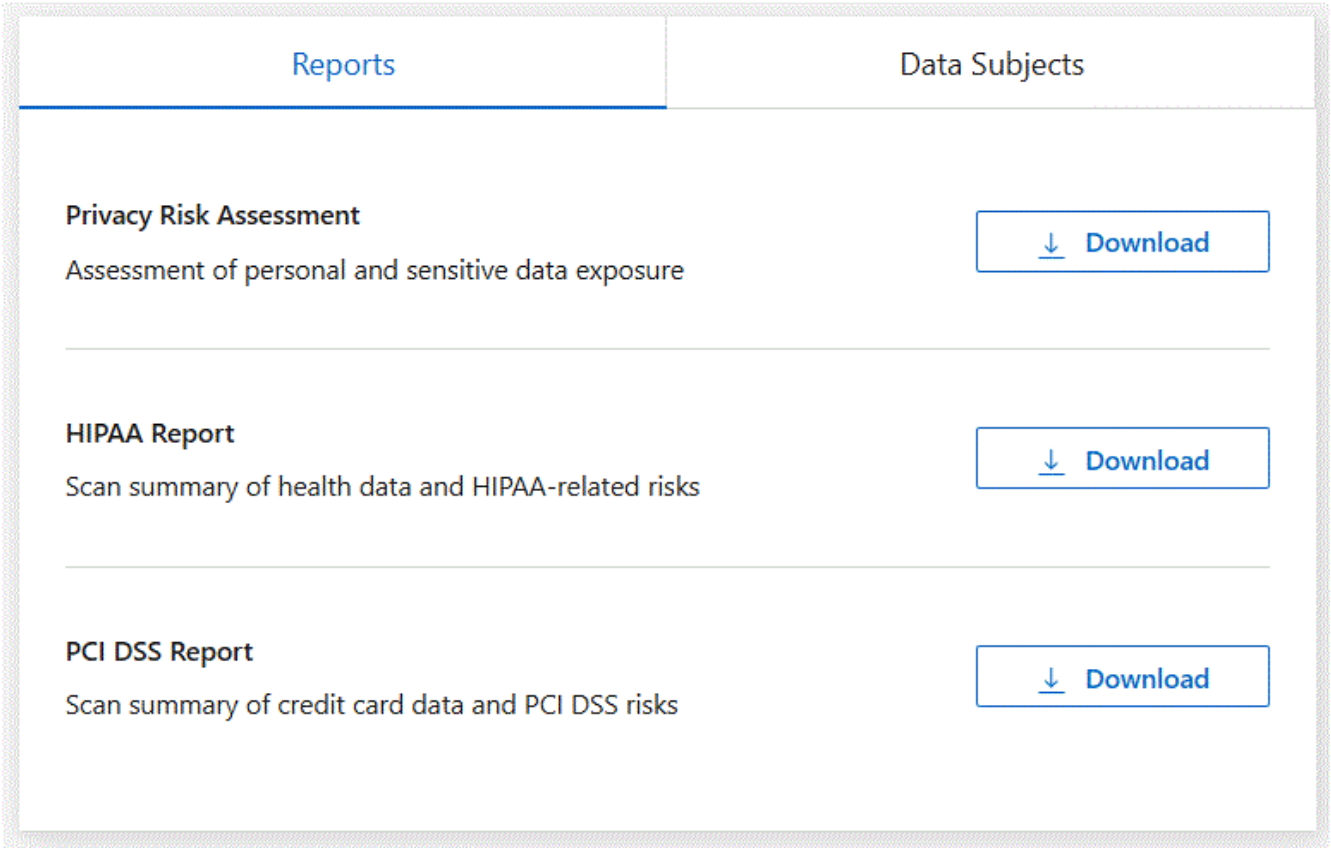
- Conservation : la période pendant laquelle les fichiers ont été modifiés pour la dernière fois. Cela est utile car vous ne devez pas conserver les informations de santé plus longtemps que nécessaire pour les traiter.
- Distribution des informations sur la santé : les systèmes dans lesquels les informations sur la santé ont été trouvées et si le cryptage et la protection contre les ransomwares sont activés.

Générer le rapport HIPAA

Accédez à l'onglet Conformité pour générer le rapport.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Localisez le **volet Rapports**. Sélectionnez l'icône de téléchargement à côté de **Rapport HIPAA**.



Résultat

La classification des données génère un rapport PDF.

Rapport sur la norme de sécurité des données de l'industrie des cartes de paiement (PCI DSS)

Le rapport sur la norme de sécurité des données de l'industrie des cartes de paiement (PCI DSS) peut vous aider à identifier la répartition des informations de carte de crédit dans vos fichiers.

Le rapport comprend les informations suivantes :

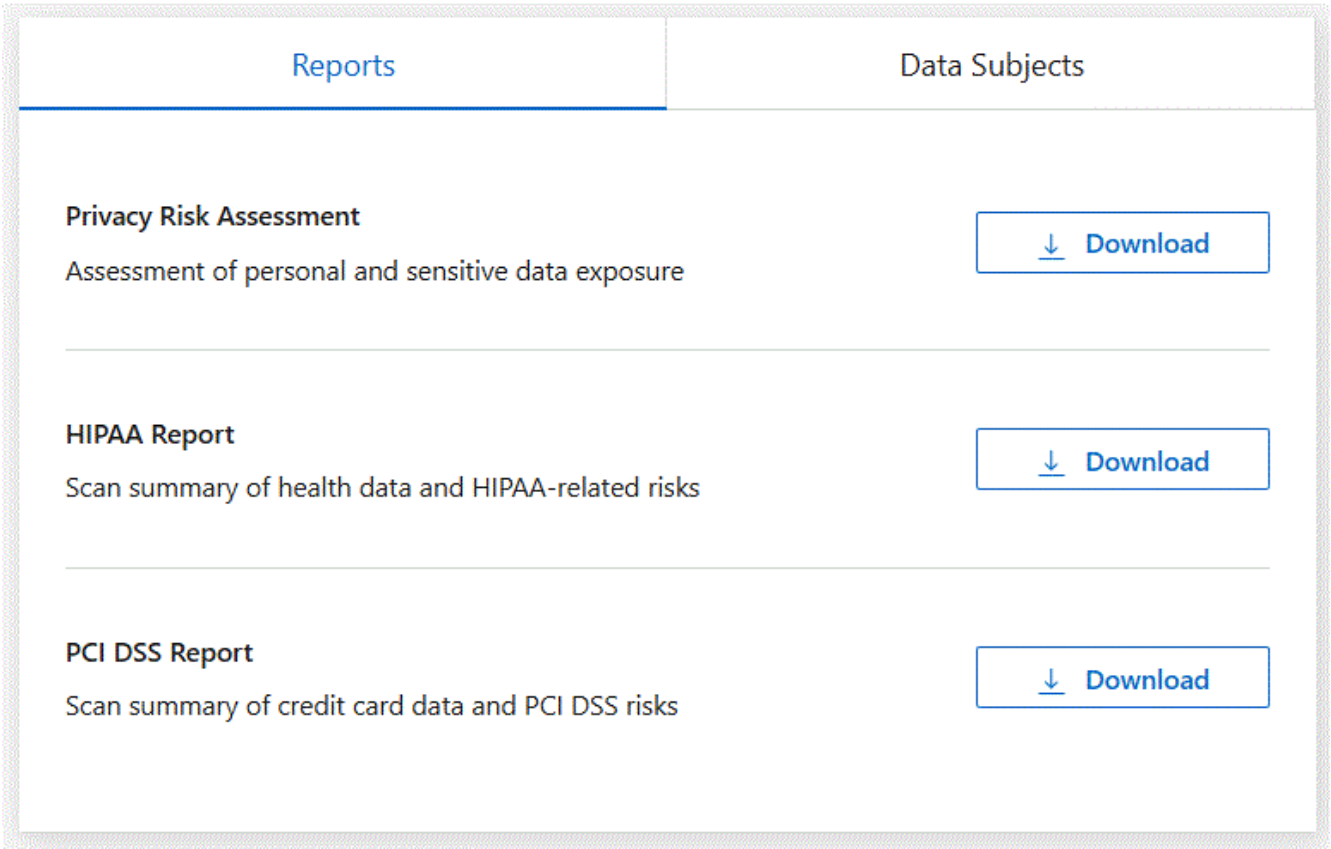
- Aperçu : Combien de fichiers contiennent des informations de carte de crédit et dans quels systèmes.
- Cryptage : pourcentage de fichiers contenant des informations de carte de crédit qui se trouvent sur des systèmes cryptés ou non cryptés. Ces informations sont spécifiques à Cloud Volumes ONTAP.
- Protection contre les ransomwares : pourcentage de fichiers contenant des informations de carte de crédit qui se trouvent sur des systèmes sur lesquels la protection contre les ransomwares est activée ou non. Ces informations sont spécifiques à Cloud Volumes ONTAP.
- Conservation : la période pendant laquelle les fichiers ont été modifiés pour la dernière fois. Cela est utile car vous ne devez pas conserver les informations de carte de crédit plus longtemps que nécessaire pour les traiter.
- Distribution des informations de carte de crédit : les systèmes sur lesquels les informations de carte de crédit ont été trouvées et si le cryptage et la protection contre les ransomwares sont activés.

Générer le rapport PCI DSS

Accédez à l'onglet Conformité pour générer le rapport.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Localisez le **volet Rapports**. Sélectionnez l'icône de téléchargement à côté de **Rapport PCI DSS**.



Résultat

La classification des données génère un rapport PDF que vous pouvez consulter et envoyer à d'autres groupes selon vos besoins.

Rapport d’évaluation des risques liés à la vie privée

Le rapport d’évaluation des risques liés à la confidentialité fournit un aperçu de l’état des risques liés à la confidentialité de votre organisation, comme l’exigent les réglementations sur la confidentialité telles que le RGPD et le CCPA.

Le rapport comprend les informations suivantes :

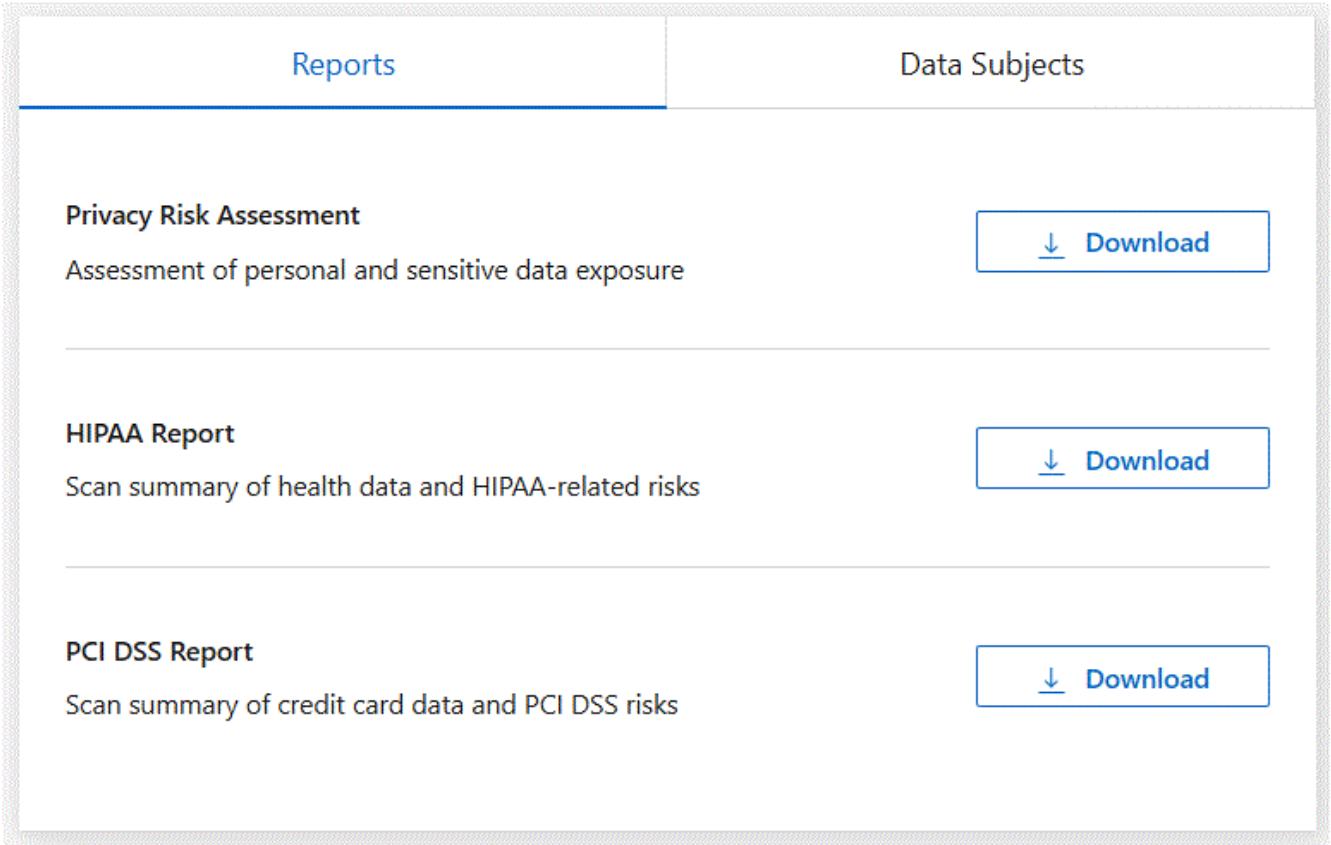
- État de conformité : un score de gravité et la distribution des données, qu’elles soient non sensibles, personnelles ou personnelles sensibles.
- Aperçu de l’évaluation : une répartition des types de données personnelles trouvées, ainsi que des catégories de données.
- Personnes concernées par cette évaluation : nombre de personnes, par lieu, pour lesquelles des identifiants nationaux ont été trouvés.

Générer le rapport d’évaluation des risques liés à la confidentialité

Accédez à l’onglet Conformité pour générer le rapport.

Étapes

1. Dans le menu Classification des données, sélectionnez **Conformité**.
2. Localisez le **volet Rapports**. Sélectionnez l’icône de téléchargement à côté de **Rapport d’évaluation des risques liés à la confidentialité**.



Résultat

La classification des données génère un rapport PDF que vous pouvez consulter et envoyer à d’autres groupes selon vos besoins.

Score de gravité

La classification des données calcule le score de gravité du rapport d'évaluation des risques liés à la confidentialité sur la base de trois variables :

- Le pourcentage de données personnelles sur l'ensemble des données.
- Le pourcentage de données personnelles sensibles sur l'ensemble des données.
- Le pourcentage de fichiers qui incluent des personnes concernées, déterminé par des identifiants nationaux tels que les cartes d'identité nationales, les numéros de sécurité sociale et les numéros d'identification fiscale.

La logique utilisée pour déterminer le score est la suivante :

Score de gravité	Logique
0	Les trois variables sont exactement 0 %
1	L'une des variables est supérieure à 0 %
2	L'une des variables est supérieure à 3 %
3	Deux des variables sont supérieures à 3 %
4	Trois des variables sont supérieures à 3 %
5	L'une des variables est supérieure à 6 %
6	Deux des variables sont supérieures à 6 %
7	Trois des variables sont supérieures à 6 %
8	L'une des variables est supérieure à 15 %
9	Deux des variables sont supérieures à 15 %
10	Trois des variables sont supérieures à 15 %

Gérer la classification des données

Exclure des répertoires spécifiques des analyses de NetApp Data Classification

Si vous souhaitez que NetApp Data Classification exclue des répertoires spécifiques des analyses, vous pouvez ajouter ces noms de répertoire à un fichier de configuration. Après avoir appliqué cette modification, le moteur de classification des données exclut ces répertoires des analyses.

Par défaut, les analyses de classification des données excluent les données d'instantané de volume, qui sont identiques à leur source dans le volume.

Sources de données prises en charge

L'exclusion de répertoires spécifiques des analyses de classification des données est prise en charge pour les partages NFS et CIFS dans les sources de données suivantes :

- ONTAP sur site
- Cloud Volumes ONTAP
- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- Partages de fichiers généraux

Définir les répertoires à exclure de l'analyse

Avant de pouvoir exclure des répertoires de l'analyse de classification, vous devez vous connecter au système de classification des données afin de pouvoir modifier un fichier de configuration et exécuter un script. Découvrez comment"connectez-vous au système de classification des données" selon que vous avez installé manuellement le logiciel sur une machine Linux ou si vous avez déployé l'instance dans le cloud.

Considérations

- Vous pouvez exclure un maximum de 50 chemins de répertoire par système de classification des données.
- L'exclusion des chemins de répertoire peut affecter les temps d'analyse.

Étapes

1. Sur le système de classification des données, accédez à « /opt/netapp/config/custom_configuration » puis ouvrez le fichier `data_provider.yaml` .
2. Dans la section « data_providers » sous la ligne « exclude : », entrez les chemins de répertoire à exclure. Par exemple:

```
exclude:
- "folder1"
- "folder2"
```

Ne modifiez rien d'autre dans ce fichier.

3. Enregistrez les modifications dans le fichier.
4. Accédez à « /opt/netapp/Datasense/tools/customer_configuration/data_providers » et exécutez le script suivant :

```
update_data_providers_from_config_file.sh
```

+ Cette commande valide les répertoires à exclure de l'analyse dans le moteur de classification.

Résultat

Toutes les analyses ultérieures de vos données excluront l'analyse des répertoires spécifiés.

Vous pouvez ajouter, modifier ou supprimer des éléments de la liste d'exclusion en suivant ces mêmes étapes. La liste d'exclusion révisée sera mise à jour après l'exécution du script pour valider vos modifications.

Exemples

Configuration 1 :

Chaque dossier contenant « folder1 » n'importe où dans le nom sera exclu de toutes les sources de données.

```
data_providers:
   exclude:
   - "folder1"
```

Résultats attendus pour les chemins qui seront exclus :

- /CVO1/dossier1
- /CVO1/nomdossier1
- /CVO1/dossier10
- /CVO1/*dossier1
- /CVO1/+nomdossier1
- /CVO1/notfolder10
- /CVO22/dossier1
- /CVO22/nomdossier1
- /CVO22/dossier10

Exemples de chemins qui ne seront pas exclus :

- /CVO1/*dossier
- /CVO1/nom du dossier
- /CVO22/*dossier20

Configuration 2 :

Tout dossier contenant uniquement « *folder1 » au début du nom sera exclu.

```
data_providers:
   exclude:
   - "\\*folder1"
```

Résultats attendus pour les chemins qui seront exclus :

- /CVO/*dossier1
- /CVO/*nomdossier1
- /CVO/*dossier10

Exemples de chemins qui ne seront pas exclus :

- /CVO/dossier1
- /CVO/nomdossier1
- /CVO/pas*dossier10

Configuration 3 :

Chaque dossier dans la source de données « CVO22 » qui contient « folder1 » n'importe où dans le nom sera exclu.

```
data_providers:
   exclude:
   - "CVO22/folder1"
```

Résultats attendus pour les chemins qui seront exclus :

- /CVO22/dossier1
- /CVO22/nomdossier1
- /CVO22/dossier10

Exemples de chemins qui ne seront pas exclus :

- /CVO1/dossier1
- /CVO1/nomdossier1
- /CVO1/dossier10

Échapper les caractères spéciaux dans les noms de dossiers

Si vous avez un nom de dossier qui contient l'un des caractères spéciaux suivants et que vous souhaitez exclure les données de ce dossier de l'analyse, vous devrez utiliser la séquence d'échappement \\ avant le nom du dossier.

```
 ., +, *, ?, ^, $, (, ), [, ], {, }, |
Par exemple:
```

Chemin dans la source : `/project/*not_to_scan`

Syntaxe dans le fichier d'exclusion : `"\\*not_to_scan"`

Afficher la liste d'exclusion actuelle

Il est possible que le contenu du `data_provider.yaml` fichier de configuration soit différent de ce qui a été réellement validé après l'exécution du `update_data_providers_from_config_file.sh` scénario. Pour afficher la liste actuelle des répertoires que vous avez exclus de l'analyse de classification des données, exécutez la commande suivante depuis « /opt/netapp/Datasense/tools/customer_configuration/data_providers » :

```
get_data_providers_configuration.sh
```

Définir des ID de groupe supplémentaires comme ouverts à l'organisation dans la NetApp Data Classification

Lorsque des ID de groupe (GID) sont attachés à des fichiers ou des dossiers dans des partages de fichiers NFS, ils définissent les autorisations pour le fichier ou le dossier ; par exemple s'ils sont « ouverts à l'organisation ». Si certains GID ne sont pas initialement configurés avec le niveau d'autorisation « Ouvrir à l'organisation », vous pouvez ajouter cette autorisation au GID afin que tous les fichiers et dossiers auxquels ce GID est attaché soient considérés comme « ouverts à l'organisation ».

Une fois cette modification effectuée et NetApp Data Classification réanalyse vos fichiers et dossiers, tous les fichiers et dossiers auxquels ces ID de groupe sont associés afficheront cette autorisation dans la page Détails de l'enquête et apparaîtront également dans les rapports où vous affichez les autorisations de fichiers.

Pour activer cette fonctionnalité, vous devez vous connecter au système de classification des données afin de pouvoir modifier un fichier de configuration et exécuter un script. Découvrez comment"connectez-vous au système de classification des données" selon que vous avez installé manuellement le logiciel sur une machine Linux ou si vous avez déployé l'instance dans le cloud.

Ajoutez l'autorisation « Ouvrir à l'organisation » aux identifiants de groupe

Vous devez disposer des numéros d'identification de groupe (GID) avant de commencer cette tâche.

Étapes

1. Sur le système de classification des données, accédez à « /opt/netapp/config/custom_configuration » et ouvrez le fichier `data_provider.yaml` .
2. Dans la ligne « organization_group_ids: [] », ajoutez les ID de groupe. Par exemple:

   ```
   organization_group_ids: [1014, 1015, 21, 2021, 1013, 2020, 1018, 1019]
   ```

   Ne changez rien d'autre dans ce fichier.
3. Enregistrez les modifications dans le fichier.
4. Accédez à « /opt/netapp/Datasense/tools/customer_configuration/data_providers » et exécutez le script suivant :

```
update_data_providers_from_config_file.sh
```

Cette commande valide les autorisations d'ID de groupe révisées dans le moteur de classification.

Résultat

Toutes les analyses ultérieures de vos données identifieront les fichiers ou dossiers auxquels ces identifiants de groupe sont associés comme étant « ouverts à l'organisation ».

Vous pouvez modifier la liste des identifiants de groupe et supprimer tous les identifiants de groupe que vous avez ajoutés dans le passé en suivant ces mêmes étapes. La liste révisée des ID de groupe sera mise à jour après avoir exécuté le script pour valider vos modifications.

Afficher la liste actuelle des identifiants de groupe

Il est possible que le contenu du `data_provider.yaml` fichier de configuration différent de ce qui a été réellement validé après l'exécution du `update_data_providers_from_config_file.sh` scénario. Pour afficher la liste actuelle des ID de groupe que vous avez ajoutés à la classification des données, exécutez la commande suivante depuis « /opt/netapp/Datasense/tools/customer_configuration/data_providers » :

```
get_data_providers_configuration.sh
```

Supprimer les sources de données de la NetApp Data Classification

Si nécessaire, vous pouvez empêcher NetApp Data Classification d'analyser un ou plusieurs systèmes, bases de données ou groupes de partage de fichiers.

Désactiver les analyses de conformité pour un système

Lorsque vous désactivez les analyses, Data Classification n'analyse plus les données du système et supprime les informations de conformité indexées de l'instance Data Classification (les données du système lui-même ne sont pas supprimées).

1. Depuis la page *Configuration*, sélectionnez l'option bouton dans la ligne du système puis **Désactiver la classification des données**.

Vous pouvez également désactiver les analyses de conformité pour un système à partir du panneau Services lorsque vous sélectionnez le système.

Supprimer une base de données de la classification des données

Si vous ne souhaitez plus analyser une certaine base de données, vous pouvez la supprimer de l'interface de classification des données et arrêter toutes les analyses.

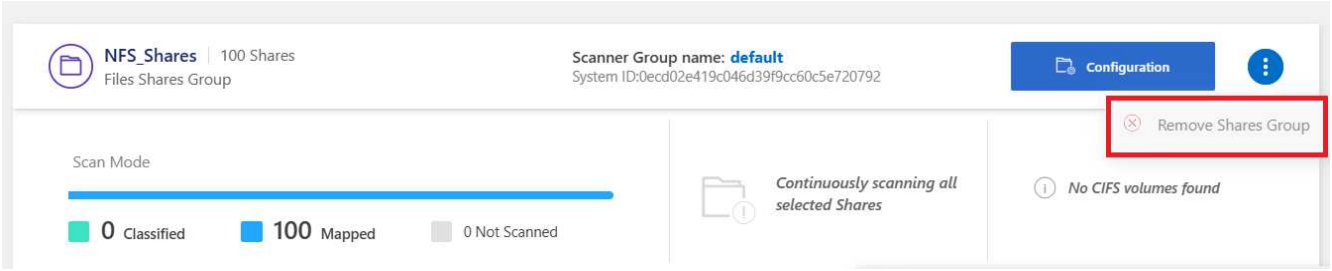
1. Depuis la page *Configuration*, sélectionnez l'option bouton dans la ligne de la base de données puis **Supprimer le serveur de base de données**.

Supprimer un groupe de partages de fichiers de la classification des données

Si vous ne souhaitez plus analyser les fichiers utilisateur d'un groupe de partages de fichiers, vous pouvez supprimer le groupe de partages de fichiers de l'interface de classification des données et arrêter toutes les analyses.

Étapes

1. Depuis la page *Configuration*, sélectionnez l'option bouton dans la ligne du groupe de partages de fichiers puis **Supprimer le groupe de partages de fichiers**.



2. Sélectionnez **Supprimer le groupe de partages** dans la boîte de dialogue de confirmation.

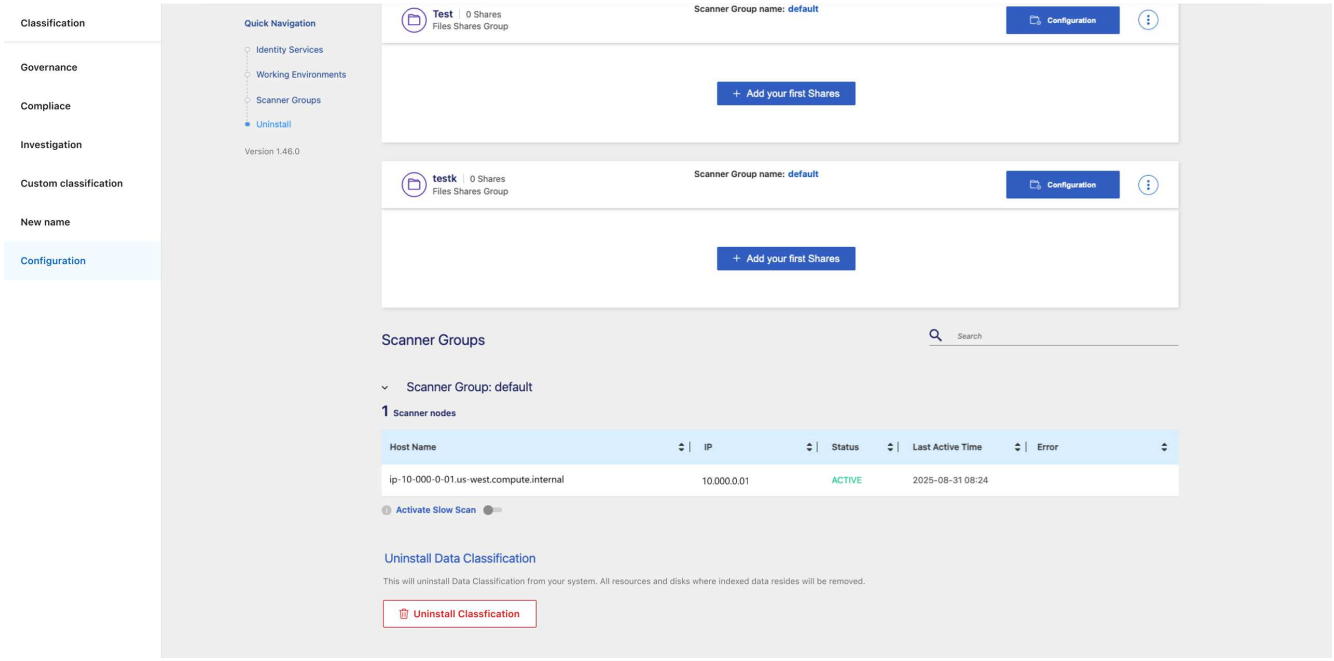
Désinstaller NetApp Data Classification

Vous pouvez désinstaller NetApp Data Classification pour résoudre les problèmes ou pour supprimer définitivement le logiciel de l'hôte. La suppression de l'instance supprime également les disques associés sur lesquels résident les données indexées, ce qui signifie que toutes les informations analysées par Data Classification seront définitivement supprimées.

Les étapes à suivre dépendent du fait que vous avez déployé la classification des données dans le cloud ou sur un hôte local.

Désinstaller Data Classification d'un fournisseur cloud

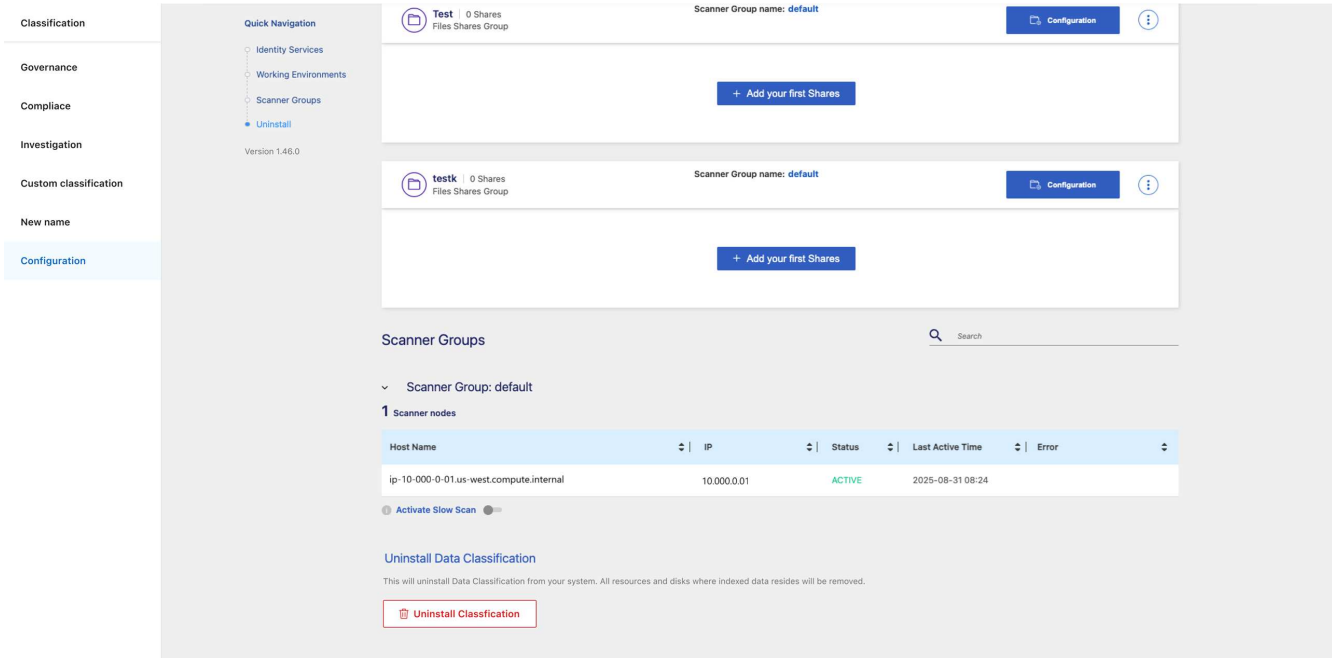
1. Dans Classification des données, sélectionnez **Configuration**.
2. Au bas de la page de configuration, sélectionnez **Désinstaller la classification**.



3. Dans la boîte de dialogue, saisissez « désinstaller » pour procéder à la déconnexion de l'instance de classification des données de l'agent de la console. Sélectionnez **Désinstaller** pour confirmer.
4. Dans la boîte de dialogue *Désinstaller la classification*, saisissez **uninstall** pour confirmer que vous souhaitez déconnecter l'instance de classification des données de l'agent de la console, puis sélectionnez **Désinstaller**.
5. Pour finaliser le processus de désinstallation, accédez à la console de votre fournisseur de cloud et supprimez l'instance de classification des données. L'instance est nommée *CloudCompliance* avec un hachage généré (UUID) concaténé. Par exemple : *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

Désinstaller la classification des données d'un déploiement sur site

1. Dans Classification des données, sélectionnez **Configuration**.
2. Au bas de la page de configuration, sélectionnez **Désinstaller la classification**.



3. Dans la boîte de dialogue, saisissez « désinstaller » pour procéder à la déconnexion de l'instance de classification des données de l'agent de la console. Sélectionnez **Désinstaller** pour confirmer.
4. Pour désinstaller le logiciel de l'hôte, exécutez le `cleanup.sh` script sur la machine hôte de classification des données, par exemple :

```
cleanup.sh
```

Le script est situé dans le `/install/light_probe/onprem_installer/cleanup.sh` annuaire. Découvrez comment"connectez-vous à la machine hôte de classification des données" .

Référence

Types d'instances de NetApp Data Classification pris en charge

Le logiciel de NetApp Data Classification doit s'exécuter sur un hôte qui répond aux exigences spécifiques du système d'exploitation, de la RAM, des logiciels, etc. Lors du déploiement de la classification des données dans le cloud, nous vous recommandons d'utiliser un système doté de caractéristiques « larges » pour bénéficier de toutes les fonctionnalités.

Vous pouvez déployer la classification des données sur un système avec moins de processeurs et moins de RAM, mais il existe certaines limitations lors de l'utilisation de ces systèmes moins puissants. "En savoir plus sur ces limitations" .

Dans les tableaux suivants, si le système marqué comme « par défaut » n'est pas disponible dans la région où vous installez Data Classification, le système suivant dans le tableau sera déployé.

Types d'instances AWS

Taille du système	Spécifications	Type d'instance
Très grand	32 processeurs, 128 Go de RAM, 1 To de SSD gp3	"m6i.8xlarge"(défaut)
Grand	16 processeurs, 64 Go de RAM, 500 Go de SSD	"m6i.4xlarge"(par défaut) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
Moyen	8 processeurs, 32 Go de RAM, 200 Go de SSD	"m6i.2xlarge"(par défaut) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
Petit	8 processeurs, 16 Go de RAM, 100 Go de SSD	"c6a.2xlarge"(par défaut) c5a.2xlarge c5.2xlarge c4.2xlarge

Types d'instances Azure

Taille du système	Spécifications	Type d'instance
Très grand	32 processeurs, 128 Go de RAM, disque système (2 048 Gio, débit minimal de 250 Mo/s) et disque de données (SSD 1 Tio, débit minimal de 750 Mo/s)	"Standard_D32_v3"(défaut)
Grand	16 processeurs, 64 Go de RAM, 500 Go de SSD	"Standard_D16s_v3"(défaut)

Types d'instances GCP

Taille du système	Spécifications	Type d'instance
Grand	16 processeurs, 64 Go de RAM, 500 Go de SSD	"n2-standard-16"(par défaut) n2d-standard-16 n1-standard-16

Métadonnées collectées à partir de sources de données dans la NetApp Data Classification

NetApp Data Classification collecte certaines métadonnées lors de l'exécution d'analyses de classification sur les données de vos sources de données et systèmes. La classification des données peut accéder à la plupart des métadonnées dont nous avons besoin pour classer vos données, mais il existe certaines sources pour lesquelles nous ne pouvons pas accéder aux données dont nous avons besoin.

	Métadonnées	CIFS	NFS
Horodatages	*Heure de création*	disponible	Non disponible (non pris en charge sous Linux)
	Heure du dernier accès	disponible	disponible
	Heure de la dernière modification	disponible	disponible
Autorisations	*Ouvrir les autorisations*	Si le groupe « TOUT LE MONDE » a accès au fichier, il est considéré comme « Ouvert à l'organisation »	Si « Autres » a accès au fichier, il est considéré comme « Ouvert à l'organisation »
	Accès utilisateurs/groupes	Les informations sur les utilisateurs et les groupes sont extraites de LDAP	Non disponible (les utilisateurs NFS sont généralement gérés localement sur le serveur, par conséquent, le même individu peut avoir un UID différent sur chaque serveur)

- La classification des données n'extrait pas la « dernière heure d'accès » des sources de données de la base de données.
- Les anciennes versions du système d'exploitation Windows (par exemple, Windows 7 et Windows 8) désactivent par défaut la collecte de l'attribut « heure du dernier accès » car cela peut avoir un impact sur les performances du système. Lorsque cet attribut n'est pas collecté, les analyses de classification des données basées sur « l'heure du dernier accès » seront affectées. Vous pouvez activer la collecte de l'heure du dernier accès sur ces anciens systèmes Windows si nécessaire.

Horodatage du dernier accès

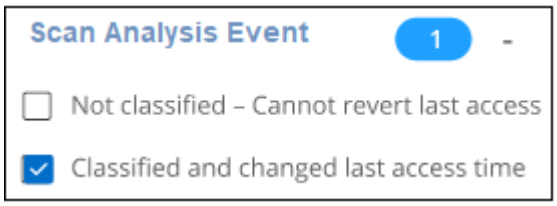
Lorsque Data Classification extrait des données à partir de partages de fichiers, le système d'exploitation considère qu'il accède aux données et modifie l'« heure du dernier accès » en conséquence. Après l'analyse, la classification des données tente de rétablir l'heure du dernier accès à l'horodatage d'origine. Si la classification des données ne dispose pas d'autorisations d'attributs d'écriture dans CIFS ou d'autorisations

d'écriture dans NFS, le système ne peut pas rétablir l'heure du dernier accès à l'horodatage d'origine. Les volumes ONTAP configurés avec SnapLock disposent d'autorisations en lecture seule et ne peuvent pas non plus rétablir l'heure du dernier accès à l'horodatage d'origine.

Par défaut, si Data Classification ne dispose pas de ces autorisations, le système n'analysera pas ces fichiers dans vos volumes car Data Classification ne peut pas rétablir l'« heure du dernier accès » à l'horodatage d'origine. Cependant, si vous ne vous souciez pas de savoir si l'heure du dernier accès est réinitialisée à l'heure d'origine dans vos fichiers, vous pouvez sélectionner le commutateur **Analyser en cas d'absence d'autorisations « attributs d'écriture »** en bas de la page de configuration afin que la classification des données analyse les volumes quelles que soient les autorisations.

Cette fonctionnalité s'applique aux systèmes ONTAP sur site, à Cloud Volumes ONTAP, à Azure NetApp Files, à Amazon FSx for NetApp ONTAP et aux partages de fichiers tiers.

Il existe un filtre dans la page Investigation appelé *Événement d'analyse d'analyse* qui vous permet d'afficher soit les fichiers qui n'ont pas été classés parce que la classification des données n'a pas pu revenir à l'heure du dernier accès, soit les fichiers qui ont été classés même si la classification des données n'a pas pu revenir à l'heure du dernier accès.



Les sélections de filtres sont :

- « Non classé – Impossible de revenir à l'heure du dernier accès » – Ceci affiche les fichiers qui n'ont pas été classés en raison d'autorisations d'écriture manquantes.
- « Heure du dernier accès classé et mis à jour » : cela affiche les fichiers qui ont été classés et la classification des données n'a pas pu réinitialiser l'heure du dernier accès à la date d'origine. Ce filtre n'est pertinent que pour les environnements dans lesquels vous avez activé **Analyser en cas d'absence d'autorisations « attributs d'écriture »**.

Si nécessaire, vous pouvez exporter ces résultats vers un rapport afin de voir quels fichiers sont ou ne sont pas analysés en raison des autorisations. "En savoir plus sur les rapports d'enquête sur les données".

Connectez-vous au système de NetApp Data Classification

Vous devez vous connecter au système de NetApp Data Classification pour pouvoir accéder aux fichiers journaux ou modifier les fichiers de configuration.

Lorsque Data Classification est installé sur une machine Linux dans vos locaux ou sur une machine Linux que vous avez déployée dans le cloud, vous pouvez accéder directement au fichier de configuration et au script.

Lorsque la classification des données est déployée dans le cloud, vous devez vous connecter en SSH à l'instance de classification des données. Vous vous connectez au système via SSH en saisissant le nom d'utilisateur et le mot de passe, ou en utilisant la clé SSH que vous avez fournie lors de l'installation de l'agent de console. La commande SSH est :

```
ssh -i <path_to_the_ssh_key> <machine_user>@<datasense_ip>
```

- `<path_to_the_ssh_key>`= emplacement des clés d'authentification ssh
- `<machine_user>`:
  - Pour AWS : utilisez <ec2-user>
  - Pour Azure : utilisez l'utilisateur créé pour l'instance de la console
  - Pour GCP : utilisez l'utilisateur créé pour l'instance de la console
- `<datasense_ip>`= Adresse IP de l'instance de la machine virtuelle

Vous devez modifier les règles entrantes du groupe de sécurité pour accéder au système dans le cloud. Pour plus de détails, voir :

- "Règles de groupe de sécurité dans AWS"
- "Règles de groupe de sécurité dans Azure"
- "Règles de pare-feu dans Google Cloud"

API de NetApp Data Classification

Les fonctionnalités de NetApp Data Classification disponibles via l'interface utilisateur Web sont également disponibles via l'API REST.

Il existe quatre catégories définies dans la classification des données qui correspondent aux onglets de l'interface utilisateur :

- Enquête
- Conformité
- Gouvernance
- Configuration

Les API de la documentation Swagger vous permettent de rechercher, d'agréger des données, de suivre vos analyses et d'effectuer des actions telles que copier, déplacer et supprimer.

Aperçu

L'API vous permet d'exécuter les fonctions suivantes :

- Informations sur l'exportation
  - Tout ce qui est disponible dans l'interface utilisateur peut être exporté via l'API (à l'exception des rapports)

    - Les données sont exportées au format JSON (faciles à analyser et à transmettre à des applications tierces, comme Splunk)
- Créez des requêtes à l'aide d'instructions « AND » et « OR », incluez et excluez des informations, et bien plus encore.

  Par exemple, vous pouvez localiser des fichiers *sans* informations personnelles identifiables (PII) spécifiques (fonctionnalité non disponible dans l'interface utilisateur). Vous pouvez également exclure des champs spécifiques de l'opération d'exportation.

- Effectuer des actions
    - Mettre à jour les informations d'identification CIFS
    - Afficher et annuler les actions
    - Réanalyser les répertoires
    - Exporter des données

L'API est sécurisée et utilise la même méthode d'authentification que l'interface utilisateur. Vous trouverez des informations sur l'authentification dans le"Documentation REST API" .

Accéder à la référence de l'API Swagger

Pour accéder à Swagger, vous aurez besoin de l'adresse IP de votre instance de classification des données. Dans le cas d'un déploiement cloud, vous utiliserez l'adresse IP publique. Ensuite, vous devrez accéder à ce point de terminaison :

https://<classification_ip>/documentation

Exemple utilisant les API

L'exemple suivant montre un appel d'API pour copier des fichiers.

Demande d'API

Vous devrez d'abord obtenir tous les champs et options pertinents pour qu'un système puisse afficher tous les filtres dans l'onglet d'enquête.

```
curl -X GET "http://{classification_ip}/api/{classification_version}
/search/options?data_mode=ALL_EXTRACTABLE" -H "accept: application/json"
-H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… " -H "x-agent-id:
hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients"
```

Réponse

```
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
      "field": "string",
      "is_rulable": true,
```

```
      "name": "string",
      "operators": [
        "EQUALS"
      ],
      "optional_values": [
        {}
      ],
      "secondary": {},
      "server_data": false,
      "type": "TEXT"
    }
  ]
}
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "POLICIES",
      "name": "Policies",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "EXTRACTION_STATUS_RANGE",
      "name": "Scan Analysis Status",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "SCAN_ANALYSIS_ERROR",
      "name": "Scan Analysis Event",
      "operators": [
        "IN"
      ],
```

```
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "PUBLIC_ACCESS",
    "name": "Open Permissions",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": true,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "USERS_PERMISSIONS_COUNT_RANGE",
    "name": "Number of Users with Access",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": true,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "USER_GROUP_PERMISSIONS",
    "name": "User / Group Permissions",
    "operators": [
      "IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "FILE_OWNER",
    "name": "File Owner",
    "operators": [
      "EQUALS",
      "CONTAINS"
```

```
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT_TYPE",
      "name": "system-type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT",
      "name": "system",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
      "field": "SCAN_TASK",
      "name": "Storage Repository",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_PATH",
      "name": "File / Directory Path",
      "operators": [
```

```
        "MULTI_CONTAINS",
        "MULTI_EXCLUDE"
      ],
      "server_data": true,
      "type": "MULTI_TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_DASHBOARD_EXTRACTABLE",
      "field": "CATEGORY",
      "name": "Category",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_SENSITIVITY_LEVEL",
      "name": "Sensitivity Level",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "NUMBER_OF_IDENTIFIERS",
      "name": "Number of identifiers",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_PERSONAL",
      "name": "Personal Data",
```

```
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_SENSITIVE",
      "name": "Sensitive Personal Data",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DATA_SUBJECT",
      "name": "Data Subject",
      "operators": [
        "EQUALS",
        "CONTAINS"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "DIRECTORIES",
      "field": "DIRECTORY_TYPE",
      "name": "Directory Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
```

```
      "field": "FILE_TYPE",
      "name": "File Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "FILE_SIZE_RANGE",
      "name": "File Size",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_CREATION_RANGE_RETENTION",
      "name": "Created Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DISCOVERED_TIME_RANGE",
      "name": "Discovered Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
```

```
      "field": "FILE_LAST_MODIFICATION_RETENTION",
      "name": "Last Modified",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_LAST_ACCESS_RANGE_RETENTION",
      "name": "Last Accessed",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "IS_DUPLICATE",
      "name": "Duplicates",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "FILE_HASH",
      "name": "File Hash",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
```

```
      "field": "USER_DEFINED_STATUS",
      "name": "Tags",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ASSIGNED_TO",
      "name": "Assigned to",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    }
  ]
}
```

Nous utiliserons cette réponse dans nos paramètres de requête pour filtrer les fichiers souhaités que nous souhaitons copier.

Vous pouvez appliquer une action sur plusieurs éléments. Les types d'actions pris en charge incluent : déplacer, supprimer et copier.

Nous allons créer l'action de copie :

Demande d'API

Cette API suivante est cette API d'action et elle vous permet de créer plusieurs actions.

```
curl -X POST "http://
{classification_ip}/api//{classification_version}/actions" -H "accept:
application/json" -H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… "
-H "x-agent-id: hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients " -H "Content-
Type: application/json" -d "{ \"action_type\": \"COPY\", \"data_mode\":
\"FILES\", \"policy_id\": 0, \"request_params\": { destination_nfs_path:
"{ontap_ip}:/{share_name} " },
\"requested_query\":{"condition":"AND","rules":[{"field":"ENVIRONMENT_TYPE
","operator":"IN","value":["ONPREM"]},{"field":"CATEGORY","operator":"IN",
"value":["21"]}]}}"
```

Réponse

La réponse renverra l'objet d'action, vous pouvez donc utiliser les API get et delete pour obtenir l'état de l'action ou pour l'annuler.


```
{
  "action_type": "COPY",
  "creation_time": "2023-08-08T12:37:21.705Z",
  "data_mode": "FILES",
  "end_time": "2023-08-08T12:37:21.705Z",
  "estimated_time_to_complete": 0,
  "id": 0,
  "policy_id": 0,
  "policy_name": "string",
  "priority": 0,
  "request_params": {},
  "requested_query": {},
  "result": {
    "error_message": "string",
    "failed": 0,
    "in_progress": 0,
    "succeeded": 0,
    "total": 0
  },
  "start_time": "2023-08-08T12:37:21.705Z",
  "status": "QUEUED",
  "title": "string",
  "user_id": "string"
}
```

Connaissances et soutien

Inscrivez-vous au support de la NetApp Console

L'enregistrement du support est requis pour bénéficier d'un support technique spécifique à la NetApp Console et à ses solutions de stockage et services de données. L'enregistrement du support est également requis pour activer les flux de travail clés pour les systèmes Cloud Volumes ONTAP .

L'inscription au support n'active pas la prise en charge NetApp pour un service de fichiers de fournisseur cloud. Pour obtenir une assistance technique relative à un service de fichiers de fournisseur cloud, à son infrastructure ou à toute solution utilisant le service, reportez-vous à « Obtenir de l'aide » dans la documentation de ce produit.

- "Amazon FSx pour ONTAP"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

Présentation de l'enregistrement de l'assistance

Il existe deux formes d'inscription pour activer le droit au support :

- Enregistrement du numéro de série de votre compte NetApp Console (votre numéro de série 960xxxxxxxxx à 20 chiffres situé sur la page Ressources de support de la console).

  Il s'agit de votre identifiant d'abonnement d'assistance unique pour tout service au sein de la console. Chaque compte de console doit être enregistré.

- Enregistrement des numéros de série Cloud Volumes ONTAP associés à un abonnement sur la place de marché de votre fournisseur de cloud (il s'agit de numéros de série 909201xxxxxxxx à 20 chiffres).

  Ces numéros de série sont communément appelés *numéros de série PAYGO* et sont générés par la NetApp Console au moment du déploiement de Cloud Volumes ONTAP .

L'enregistrement des deux types de numéros de série permet des fonctionnalités telles que l'ouverture de tickets d'assistance et la génération automatique de dossiers. L'enregistrement est terminé en ajoutant des comptes NetApp Support Site (NSS) à la console comme décrit ci-dessous.

Enregistrez la NetApp Console pour le support NetApp

Pour vous inscrire au support et activer le droit de support, un utilisateur de votre compte NetApp Console doit associer un compte de site de support NetApp à sa connexion à la console. La manière dont vous vous inscrivez au support NetApp dépend du fait que vous possédez déjà ou non un compte NetApp Support Site (NSS).

Client existant avec un compte NSS

Si vous êtes un client NetApp avec un compte NSS, il vous suffit de vous inscrire pour bénéficier de l'assistance via la console.

Étapes

1. Sélectionnez **Administration** > **Informations d'identification**.
2. Sélectionnez **Informations d'identification de l'utilisateur**.
3. Sélectionnez **Ajouter des informations d'identification NSS** et suivez l'invite d'authentification du site de support NetApp (NSS).
4. Pour confirmer que le processus d'inscription a réussi, sélectionnez l'icône Aide, puis sélectionnez **Assistance**.

   La page **Ressources** devrait indiquer que votre compte Console est enregistré pour l'assistance.

   Notez que les autres utilisateurs de la console ne verront pas ce même statut d'enregistrement de support s'ils n'ont pas associé un compte de site de support NetApp à leur connexion. Cependant, cela ne signifie pas que votre compte n'est pas enregistré pour bénéficier de l'assistance. Tant qu'un utilisateur de l'organisation a suivi ces étapes, votre compte a été enregistré.

Client existant mais pas de compte NSS

Si vous êtes un client NetApp existant avec des licences et des numéros de série existants mais *pas* de compte NSS, vous devez créer un compte NSS et l'associer à votre connexion à la console.

Étapes

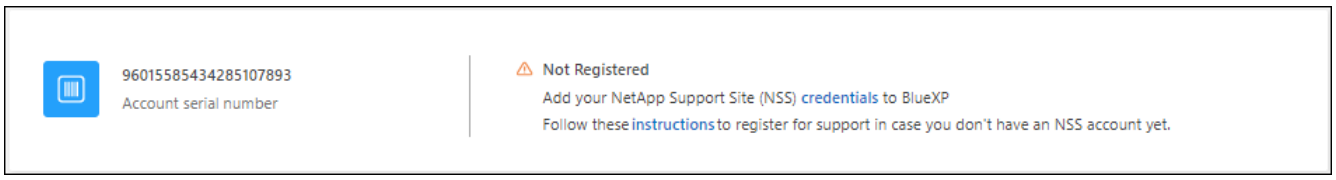
1. Créez un compte sur le site de support NetApp en remplissant le "Formulaire d'inscription des utilisateurs du site de support NetApp"
   a. Assurez-vous de sélectionner le niveau d'utilisateur approprié, qui est généralement * Client/Utilisateur final NetApp *.
   b. Assurez-vous de copier le numéro de série du compte de console (960xxxx) utilisé ci-dessus pour le champ du numéro de série. Cela accélérera le traitement du compte.
2. Associez votre nouveau compte NSS à votre connexion à la console en suivant les étapes ci-dessousClient existant avec un compte NSS .

Tout nouveau chez NetApp

Si vous êtes nouveau sur NetApp et que vous n'avez pas de compte NSS, suivez chaque étape ci-dessous.

Étapes

1. Dans le coin supérieur droit de la console, sélectionnez l'icône Aide, puis sélectionnez **Support**.
2. Recherchez le numéro de série de votre identifiant de compte sur la page d'inscription au support.



3. Accéder à "Site d'inscription au support de NetApp" et sélectionnez **Je ne suis pas un client NetApp enregistré**.
4. Remplissez les champs obligatoires (ceux avec des astérisques rouges).
5. Dans le champ **Gamme de produits**, sélectionnez **Cloud Manager**, puis sélectionnez votre fournisseur de facturation applicable.
6. Copiez le numéro de série de votre compte à l'étape 2 ci-dessus, effectuez la vérification de sécurité, puis

confirmez que vous avez lu la politique de confidentialité des données mondiales de NetApp.

Un email est immédiatement envoyé à la boîte mail prévue à cet effet pour finaliser cette transaction sécurisée. Assurez-vous de vérifier vos dossiers spam si l'e-mail de validation n'arrive pas dans quelques minutes.

7. Confirmez l'action depuis l'e-mail.

   La confirmation soumet votre demande à NetApp et vous recommande de créer un compte sur le site de support NetApp .

8. Créez un compte sur le site de support NetApp en remplissant le "Formulaire d'inscription des utilisateurs du site de support NetApp"
   a. Assurez-vous de sélectionner le niveau d'utilisateur approprié, qui est généralement * Client/Utilisateur final NetApp *.
   b. Assurez-vous de copier le numéro de série du compte (960xxxx) utilisé ci-dessus pour le champ du numéro de série. Cela accélérera le traitement.

Après avoir terminé

NetApp devrait vous contacter au cours de ce processus. Il s'agit d'un exercice d'intégration unique pour les nouveaux utilisateurs.

Une fois que vous avez votre compte de site de support NetApp , associez le compte à votre connexion à la console en suivant les étapes ci-dessousClient existant avec un compte NSS .

Associer les informations d'identification NSS pour la prise en charge de Cloud Volumes ONTAP

L'association des informations d'identification du site de support NetApp à votre compte de console est requise pour activer les workflows clés suivants pour Cloud Volumes ONTAP:

- Enregistrement des systèmes Cloud Volumes ONTAP prépayés pour le support

  Fournir votre compte NSS est nécessaire pour activer le support de votre système et pour accéder aux ressources de support technique NetApp .

- Déploiement de Cloud Volumes ONTAP lorsque vous apportez votre propre licence (BYOL)

  Il est nécessaire de fournir votre compte NSS pour que la console puisse télécharger votre clé de licence et activer l'abonnement pour la durée que vous avez achetée. Cela inclut les mises à jour automatiques pour les renouvellements de mandat.

- Mise à niveau du logiciel Cloud Volumes ONTAP vers la dernière version

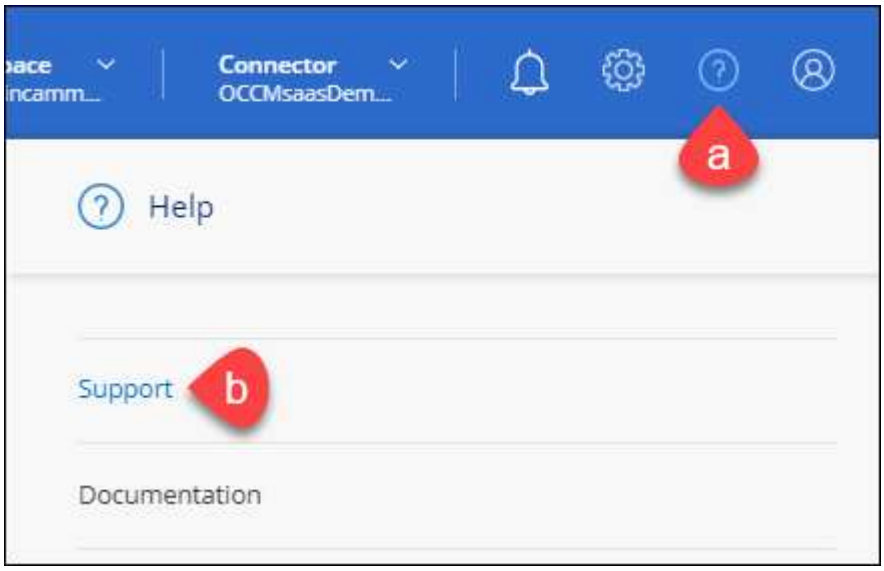
L'association des informations d'identification NSS à votre compte de NetApp Console est différente du compte NSS associé à une connexion utilisateur de console.

Ces informations d'identification NSS sont associées à votre ID de compte de console spécifique. Les utilisateurs appartenant à l'organisation Console peuvent accéder à ces informations d'identification depuis **Support > Gestion NSS**.

- Si vous disposez d'un compte client, vous pouvez ajouter un ou plusieurs comptes NSS.
- Si vous disposez d'un compte partenaire ou revendeur, vous pouvez ajouter un ou plusieurs comptes NSS, mais ils ne peuvent pas être ajoutés aux côtés des comptes de niveau client.

Étapes

1. Dans le coin supérieur droit de la console, sélectionnez l'icône Aide, puis sélectionnez **Support**.



2. Sélectionnez **Gestion NSS > Ajouter un compte NSS**.
3. Lorsque vous y êtes invité, sélectionnez **Continuer** pour être redirigé vers une page de connexion Microsoft.

   NetApp utilise Microsoft Entra ID comme fournisseur d'identité pour les services d'authentification spécifiques au support et aux licences.

4. Sur la page de connexion, indiquez votre adresse e-mail et votre mot de passe enregistrés sur le site de support NetApp pour effectuer le processus d'authentification.

   Ces actions permettent à la console d'utiliser votre compte NSS pour des tâches telles que les téléchargements de licences, la vérification des mises à niveau de logiciels et les futures inscriptions au support.

   Notez ce qui suit :

   - Le compte NSS doit être un compte de niveau client (pas un compte invité ou temporaire). Vous pouvez avoir plusieurs comptes NSS au niveau client.
   - Il ne peut y avoir qu'un seul compte NSS si ce compte est un compte de niveau partenaire. Si vous essayez d'ajouter des comptes NSS au niveau client et qu'un compte au niveau partenaire existe, vous obtiendrez le message d'erreur suivant :

     « Le type de client NSS n'est pas autorisé pour ce compte car il existe déjà des utilisateurs NSS de types différents. »

   Il en va de même si vous disposez de comptes NSS préexistants au niveau client et que vous essayez d'ajouter un compte au niveau partenaire.

   - Une fois la connexion réussie, NetApp stockera le nom d'utilisateur NSS.

     Il s'agit d'un identifiant généré par le système qui correspond à votre e-mail. Sur la page **Gestion NSS**, vous pouvez afficher votre e-mail à partir du ••• menu.

- Si vous avez besoin d'actualiser vos jetons d'identification de connexion, il existe également une option **Mettre à jour les informations d'identification** dans le ••• menu.

  L'utilisation de cette option vous invite à vous reconnecter. Notez que le jeton de ces comptes expire après 90 jours. Une notification sera publiée pour vous en informer.

Obtenez de l'aide pour la NetApp Data Classification

NetApp fournit un support pour NetApp Console et ses services cloud de diverses manières. De nombreuses options d'auto-assistance gratuites sont disponibles 24h/24 et 7j/7, telles que des articles de la base de connaissances (KB) et un forum communautaire. Votre inscription au support inclut une assistance technique à distance via un ticket web.

Obtenir de l'aide pour un service de fichiers d'un fournisseur cloud

Pour obtenir une assistance technique relative à un service de fichiers de fournisseur cloud, à son infrastructure ou à toute solution utilisant le service, reportez-vous à la documentation de ce produit.

- "Amazon FSx pour ONTAP"
- "Azure NetApp Files"
- "Google Cloud NetApp Volumes"

Pour bénéficier d'un support technique spécifique à NetApp et à ses solutions de stockage et services de données, utilisez les options de support décrites ci-dessous.

Utiliser les options d'auto-assistance

Ces options sont disponibles gratuitement, 24 heures sur 24, 7 jours sur 7 :

- Documentation

  La documentation de la NetApp Console que vous consultez actuellement.

- "Base de connaissances"

  Recherchez dans la base de connaissances NetApp pour trouver des articles utiles pour résoudre les problèmes.

- "Communautés"

  Rejoignez la communauté NetApp Console pour suivre les discussions en cours ou en créer de nouvelles.

Créer un dossier auprès du support NetApp

En plus des options d'auto-assistance ci-dessus, vous pouvez travailler avec un spécialiste du support NetApp pour résoudre tout problème après avoir activé le support.

Avant de commencer

- Pour utiliser la fonctionnalité **Créer un dossier**, vous devez d'abord associer vos informations

d’identification du site de support NetApp à votre connexion à la console. "Découvrez comment gérer les informations d’identification associées à votre connexion à la console" .

- Si vous ouvrez un dossier pour un système ONTAP doté d’un numéro de série, votre compte NSS doit être associé au numéro de série de ce système.

Étapes

1. Dans la NetApp Console, sélectionnez **Aide > Support**.
2. Sur la page **Ressources**, choisissez l’une des options disponibles sous Support technique :

   a. Sélectionnez **Appelez-nous** si vous souhaitez parler à quelqu’un au téléphone. Vous serez redirigé vers une page sur netapp.com qui répertorie les numéros de téléphone que vous pouvez appeler.

   b. Sélectionnez **Créer un dossier** pour ouvrir un ticket avec un spécialiste du support NetApp :

      - **Service** : sélectionnez le service auquel le problème est associé. Par exemple, * NetApp Console* lorsqu’il s’agit d’un problème de support technique lié aux flux de travail ou aux fonctionnalités de la console.
      - **Système** : Si applicable au stockage, sélectionnez * Cloud Volumes ONTAP* ou **On-Prem**, puis l’environnement de travail associé.

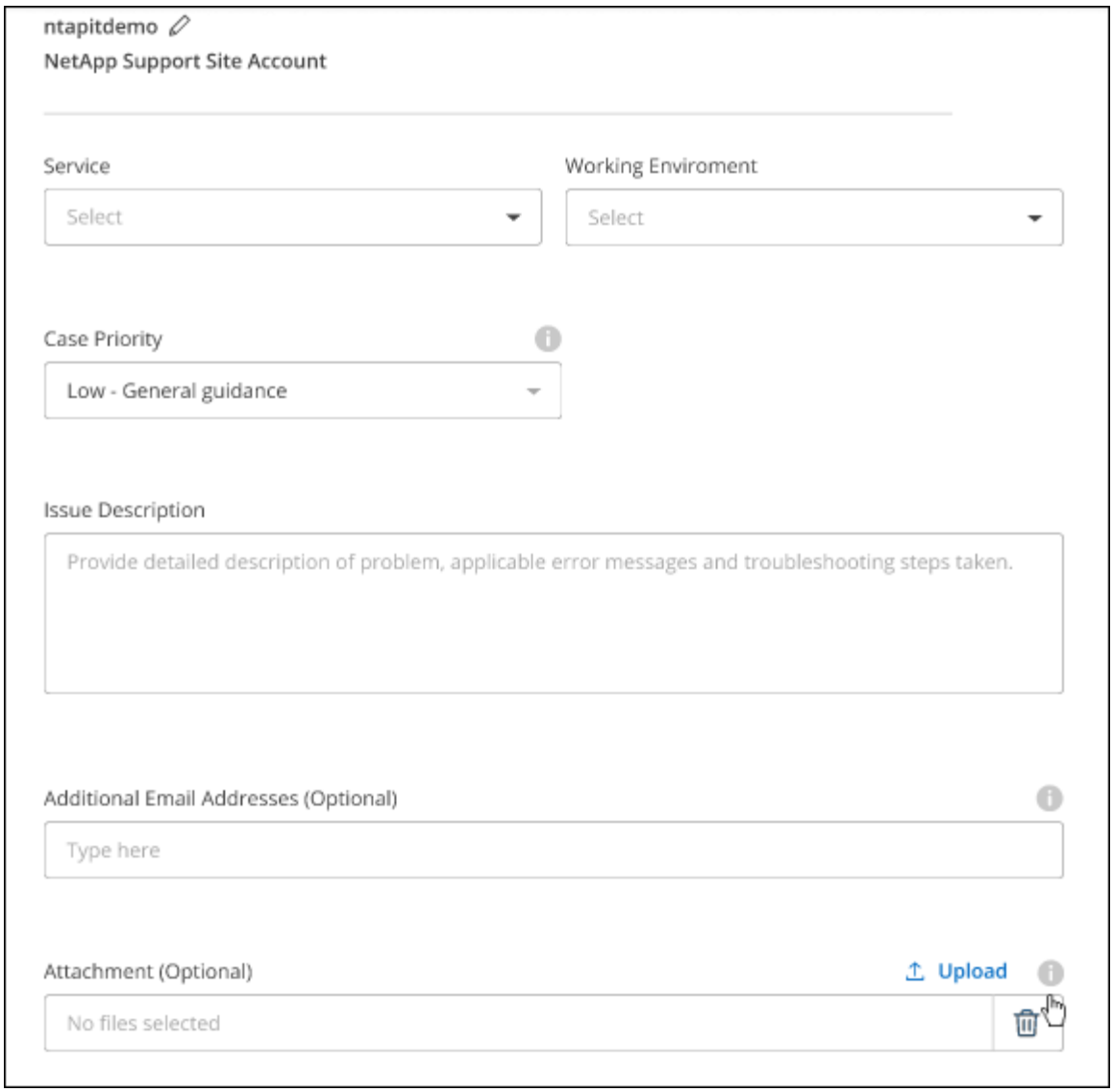
        La liste des systèmes est dans le périmètre de l’organisation de la console et de l’agent de console que vous avez sélectionné dans la bannière supérieure.

      - **Priorité du cas** : Choisissez la priorité du cas, qui peut être Faible, Moyenne, Élevée ou Critique.

        Pour en savoir plus sur ces priorités, passez votre souris sur l’icône d’information à côté du nom du champ.

      - **Description du problème** : Fournissez une description détaillée de votre problème, y compris tous les messages d’erreur applicables ou les étapes de dépannage que vous avez effectuées.
      - **Adresses e-mail supplémentaires** : saisissez des adresses e-mail supplémentaires si vous souhaitez informer quelqu’un d’autre de ce problème.
      - **Pièce jointe (facultatif)** : Téléchargez jusqu’à cinq pièces jointes, une à la fois.

        Les pièces jointes sont limitées à 25 Mo par fichier. Les extensions de fichiers suivantes sont prises en charge : txt, log, pdf, jpg/jpeg, rtf, doc/docx, xls/xlsx et csv.



Après avoir terminé

Une fenêtre contextuelle apparaîtra avec votre numéro de dossier d'assistance. Un spécialiste du support NetApp examinera votre cas et vous répondra dans les plus brefs délais.

Pour un historique de vos demandes d'assistance, vous pouvez sélectionner **Paramètres > Chronologie** et rechercher les actions nommées « créer une demande d'assistance ». Un bouton à l'extrême droite vous permet de développer l'action pour voir les détails.

Il est possible que vous rencontriez le message d'erreur suivant lorsque vous essayez de créer un dossier :

« Vous n'êtes pas autorisé à créer un dossier contre le service sélectionné »

Cette erreur peut signifier que le compte NSS et la société d'enregistrement à laquelle il est associé ne sont pas la même société d'enregistrement pour le numéro de série du compte NetApp Console (c'est-à-dire. 960xxxx) ou le numéro de série de l'environnement de travail. Vous pouvez demander de l'aide en utilisant l'une des options suivantes :

- Soumettez un cas non technique à https://mysupport.netapp.com/site/help

Gérez vos cas d’assistance

Vous pouvez afficher et gérer les cas d’assistance actifs et résolus directement depuis la console. Vous pouvez gérer les cas associés à votre compte NSS et à votre entreprise.

Notez ce qui suit :

- Le tableau de bord de gestion des cas en haut de la page offre deux vues :
  - La vue de gauche montre le nombre total de dossiers ouverts au cours des 3 derniers mois par le compte utilisateur NSS que vous avez fourni.
  - La vue de droite affiche le nombre total de dossiers ouverts au cours des 3 derniers mois au niveau de votre entreprise en fonction de votre compte utilisateur NSS.

  Les résultats du tableau reflètent les cas liés à la vue que vous avez sélectionnée.
- Vous pouvez ajouter ou supprimer des colonnes d’intérêt et filtrer le contenu des colonnes telles que Priorité et Statut. D’autres colonnes fournissent simplement des capacités de tri.

  Consultez les étapes ci-dessous pour plus de détails.
- Au niveau de chaque cas, nous offrons la possibilité de mettre à jour les notes du cas ou de fermer un cas qui n’est pas déjà au statut Fermé ou En attente de fermeture.

Étapes

1. Dans la NetApp Console, sélectionnez **Aide > Support**.
2. Sélectionnez **Gestion des cas** et si vous y êtes invité, ajoutez votre compte NSS à la console.

   La page **Gestion des cas** affiche les cas ouverts liés au compte NSS associé à votre compte utilisateur de la console. Il s’agit du même compte NSS qui apparaît en haut de la page **Gestion NSS**.
3. Modifiez éventuellement les informations qui s’affichent dans le tableau :
   - Sous **Cas de l’organisation**, sélectionnez **Afficher** pour afficher tous les cas associés à votre entreprise.
   - Modifiez la plage de dates en choisissant une plage de dates exacte ou en choisissant une période différente.
   - Filtrer le contenu des colonnes.
   - Modifiez les colonnes qui apparaissent dans le tableau en sélectionnant et ensuite choisir les colonnes que vous souhaitez afficher.
4. Gérer un dossier existant en sélectionnant et en sélectionnant l’une des options disponibles :
   - **Voir le cas** : Afficher tous les détails sur un cas spécifique.
   - **Mettre à jour les notes du cas** : fournissez des détails supplémentaires sur votre problème ou sélectionnez **Télécharger des fichiers** pour joindre jusqu’à un maximum de cinq fichiers.

     Les pièces jointes sont limitées à 25 Mo par fichier. Les extensions de fichiers suivantes sont prises en charge : txt, log, pdf, jpg/jpeg, rtf, doc/docx, xls/xlsx et csv.
   - **Fermer le dossier** : Fournissez des détails sur les raisons pour lesquelles vous fermez le dossier et sélectionnez **Fermer le dossier**.

Questions fréquemment posées sur la NetApp Data Classification

Cette FAQ peut vous aider si vous cherchez simplement une réponse rapide à une question.

NetApp Data Classification

Les questions suivantes fournissent une compréhension générale de la classification des données.

Comment fonctionne la classification des données ?

La classification des données déploie une autre couche d'IA aux côtés de votre système de NetApp Console et de vos systèmes de stockage. Il analyse ensuite les données sur les volumes, les buckets, les bases de données et d'autres comptes de stockage et indexe les informations de données trouvées. La classification des données exploite à la fois l'intelligence artificielle et le traitement du langage naturel, contrairement aux solutions alternatives généralement construites autour d'expressions régulières et de correspondances de modèles.

La classification des données utilise l'IA pour fournir une compréhension contextuelle des données pour une détection et une classification précises. Il est piloté par l'IA car il est conçu pour les types de données et l'échelle modernes. Il comprend également le contexte des données afin de fournir une découverte et une classification solides et précises.

"En savoir plus sur le fonctionnement de la classification des données" .

Data Classification dispose-t-il d'une API REST et fonctionne-t-il avec des outils tiers ?

Oui, Data Classification dispose d'une API REST pour les fonctionnalités prises en charge dans la version Data Classification qui fait partie de la plate-forme principale de la console. Voir "Documentation API" .

La classification des données est-elle disponible via les places de marché cloud ?

La classification des données fait partie des fonctionnalités principales de la NetApp Console . Vous n'avez donc pas besoin d'utiliser les places de marché pour ce service.

Classification des données, numérisation et analyse

Les questions suivantes concernent les performances d'analyse de la classification des données et les analyses.

À quelle fréquence Data Classification analyse-t-il mes données ?

Bien que l'analyse initiale de vos données puisse prendre un peu de temps, les analyses suivantes inspectent uniquement les modifications incrémentielles, ce qui réduit les temps d'analyse du système. La classification des données analyse vos données en continu de manière circulaire, six référentiels à la fois, de sorte que toutes les données modifiées sont classées très rapidement.

"Découvrez comment fonctionnent les analyses" .

La classification des données analyse les bases de données une seule fois par jour ; les bases de données ne sont pas analysées en continu comme les autres sources de données.

Les analyses de données ont un impact négligeable sur vos systèmes de stockage et sur vos données.

Les performances de numérisation varient-elles ?

Les performances d'analyse peuvent varier en fonction de la bande passante du réseau et de la taille moyenne des fichiers dans votre environnement. Cela peut également dépendre des caractéristiques de taille du système hôte (dans le cloud ou sur site). Voir "L'instance de classification des données" et "Déploiement de la classification des données" pour plus d'informations.

Lors de l'ajout initial de nouvelles sources de données, vous pouvez également choisir d'effectuer uniquement une analyse de « mappage » (Mapping uniquement) au lieu d'une analyse de « classification » complète (Map & Classify). La cartographie peut être effectuée sur vos sources de données très rapidement car elle n'accède pas aux fichiers pour voir les données à l'intérieur. "Découvrez la différence entre une analyse de cartographie et une analyse de classification" .

Puis-je rechercher mes données à l'aide de la classification des données ?

La classification des données offre des capacités de recherche étendues qui facilitent la recherche d'un fichier ou d'un élément de données spécifique dans toutes les sources connectées. La classification des données permet aux utilisateurs de rechercher plus en profondeur que ce que reflètent les métadonnées. Il s'agit d'un service indépendant de la langue qui peut également lire les fichiers et analyser une multitude de types de données sensibles, tels que les noms et les identifiants. Par exemple, les utilisateurs peuvent effectuer des recherches dans les magasins de données structurés et non structurés pour trouver des données qui peuvent avoir fui des bases de données vers les fichiers utilisateur, en violation de la politique de l'entreprise. Les recherches peuvent être enregistrées pour plus tard et des politiques peuvent être créées pour rechercher et agir sur les résultats à une fréquence définie.

Une fois les fichiers d'intérêt trouvés, les caractéristiques peuvent être répertoriées, notamment les balises, le compte système, le bucket, le chemin du fichier, la catégorie (à partir de la classification), la taille du fichier, la dernière modification, l'état de l'autorisation, les doublons, le niveau de sensibilité, les données personnelles, les types de données sensibles dans le fichier, le propriétaire, le type de fichier, la taille du fichier, l'heure de création, le hachage du fichier, si les données ont été attribuées à une personne recherchant son attention, et plus encore. Des filtres peuvent être appliqués pour éliminer les caractéristiques qui ne sont pas pertinentes.

Data Classification dispose également d'un contrôle d'accès basé sur les rôles (RBAC) pour permettre le déplacement ou la suppression des fichiers, si les autorisations appropriées sont présentes. Si les autorisations appropriées ne sont pas présentes, les tâches peuvent être attribuées à une personne de l'organisation qui dispose des autorisations appropriées.

Gestion de la classification des données et confidentialité

Les questions suivantes fournissent des informations sur la gestion des paramètres de classification des données et de confidentialité.

Comment activer ou désactiver la classification des données ?

Vous devez d'abord déployer une instance de classification des données dans la console ou sur un système local. Une fois l'instance en cours d'exécution, vous pouvez activer le service sur les systèmes, bases de données et autres sources de données existants à partir de l'onglet **Configuration** ou en sélectionnant un système spécifique. "Apprenez comment démarrer" .

L’activation de la classification des données sur une source de données entraîne une analyse initiale immédiate. Les résultats de l’analyse s’affichent peu de temps après.

Vous pouvez désactiver la classification des données pour qu’elle analyse un système individuel, une base de données ou un groupe de partage de fichiers à partir de la page Configuration de la classification des données. Voir "Supprimer les sources de données de la classification des données" .

Pour supprimer complètement l’instance de classification des données, supprimez manuellement l’instance de classification des données du portail de votre fournisseur de cloud ou de l’emplacement sur site.

Le service peut-il exclure l’analyse des données dans certains répertoires ?

Oui. Si vous souhaitez que la classification des données exclue les données d’analyse qui résident dans certains répertoires de sources de données, vous pouvez fournir cette liste au moteur de classification. Une fois cette modification appliquée, la classification des données exclura les données d’analyse dans les répertoires spécifiés. "Apprendre encore plus" .

Les instantanés résidant sur les volumes ONTAP sont-ils analysés ?

Non. La classification des données n’analyse pas les instantanés car le contenu est identique au contenu du volume.

Que se passe-t-il si la hiérarchisation des données est activée sur vos volumes ONTAP ?

Lorsque la classification des données analyse les volumes contenant des données froides hiérarchisées vers le stockage d’objets à l’aide des analyses de mappage uniquement, elle analyse toutes les données : les données qui se trouvent sur les disques locaux et les données froides hiérarchisées vers le stockage d’objets. Ceci est également vrai pour les produits non NetApp qui implémentent la hiérarchisation.

L’analyse de cartographie uniquement ne réchauffe pas les données froides : elles restent froides et restent dans le stockage d’objets. En revanche, si vous effectuez l’analyse Map & Classify, certaines configurations risquent de réchauffer les données froides.

Types de systèmes sources et types de données

Les questions suivantes concernent les types de stockage qui peuvent être analysés et les types de données analysées.

Existe-t-il des restrictions lors d’un déploiement dans une région gouvernementale ?

La classification des données est prise en charge lorsque l’agent de console est déployé dans une région gouvernementale (AWS GovCloud, Azure Gov ou Azure DoD) - également appelée « mode restreint ».

Quelles sources de données puis-je analyser si j’installe Data Classification sur un site sans accès Internet ?

Le mode privé BlueXP (interface BlueXP héritée) est généralement utilisé avec des environnements locaux qui n'ont pas de connexion Internet et avec des régions cloud sécurisées, notamment AWS Secret Cloud, AWS Top Secret Cloud et Azure IL6. NetApp continue de prendre en charge ces environnements avec l'interface BlueXP héritée. Pour la documentation du mode privé dans l'ancienne interface BlueXP , voir"Documentation PDF pour le mode privé BlueXP" .

La classification des données ne peut analyser que les données provenant de sources de données locales sur le site local. À l'heure actuelle, la classification des données peut analyser les sources de données locales suivantes en « mode privé » – également appelé site « dark » :

- Systèmes ONTAP sur site
- Schémas de bases de données
- Stockage d'objets utilisant le protocole Simple Storage Service (S3)

Quels types de fichiers sont pris en charge ?

La classification des données analyse tous les fichiers pour obtenir des informations sur les catégories et les métadonnées et affiche tous les types de fichiers dans la section Types de fichiers du tableau de bord.

Lorsque la classification des données détecte des informations personnelles identifiables (PII) ou lorsqu'elle effectue une recherche DSAR, seuls les formats de fichiers suivants sont pris en charge :

`.CSV, .DCM, .DOC, .DOCX, .JSON, .PDF, .PPTX, .RTF, .TXT, .XLS, .XLSX, Docs, Sheets, and Slides`

Quels types de données et de métadonnées la classification des données capture-t-elle ?

La classification des données vous permet d'exécuter une analyse de « mappage » générale ou une analyse de « classification » complète sur vos sources de données. La cartographie fournit uniquement un aperçu de haut niveau de vos données, tandis que la classification fournit une analyse approfondie de vos données. La cartographie peut être effectuée sur vos sources de données très rapidement car elle n'accède pas aux fichiers pour voir les données à l'intérieur.

- **Analyse de mappage de données (Analyse de mappage uniquement)** : la classification des données analyse uniquement les métadonnées. Ceci est utile pour la gestion et la gouvernance globales des données, la définition rapide de la portée des projets, les très grands domaines et la priorisation. Le mappage des données est basé sur les métadonnées et est considéré comme une analyse **rapide**.

  Après une analyse rapide, vous pouvez générer un rapport de mappage de données. Ce rapport est un aperçu des données stockées dans vos sources de données d'entreprise pour vous aider à prendre des décisions concernant l'utilisation des ressources, la migration, la sauvegarde, la sécurité et les processus de conformité.

- **Analyse approfondie de la classification des données (analyse Map & Classify)** : la classification des données analyse les données à l'aide de protocoles standard et d'une autorisation en lecture seule dans tous vos environnements. Certains fichiers sont ouverts et analysés à la recherche de données sensibles liées à l'entreprise, d'informations privées et de problèmes liés aux ransomwares.

  Après une analyse complète, vous pouvez appliquer de nombreuses fonctionnalités supplémentaires de classification des données à vos données, telles que l'affichage et l'affinage des données dans la page Enquête sur les données, la recherche de noms dans les fichiers, la copie, le déplacement et la

suppression des fichiers sources, et bien plus encore.

La classification des données capture des métadonnées telles que : le nom du fichier, les autorisations, l'heure de création, le dernier accès et la dernière modification. Cela inclut toutes les métadonnées qui apparaissent dans la page Détails de l'enquête sur les données et dans les rapports d'enquête sur les données.

La classification des données peut identifier de nombreux types de données privées telles que les informations personnelles (PII) et les informations personnelles sensibles (SPII). Pour plus de détails sur les données privées, reportez-vous àCatégories de données privées analysées par la classification des données .

Puis-je limiter les informations de classification des données à des utilisateurs spécifiques ?

Oui, la classification des données est entièrement intégrée à la NetApp Console. Les utilisateurs de la NetApp Console ne peuvent voir que les informations des systèmes qu'ils sont autorisés à consulter en fonction de leurs autorisations.

De plus, si vous souhaitez autoriser certains utilisateurs à afficher uniquement les résultats de l'analyse de classification des données sans avoir la possibilité de gérer les paramètres de classification des données, vous pouvez attribuer à ces utilisateurs le rôle **Visionneuse de classification** (lors de l'utilisation de la NetApp Console en mode standard) ou le rôle **Visionneuse de conformité** (lors de l'utilisation de la NetApp Console en mode restreint). "Apprendre encore plus" .

Quelqu'un peut-il accéder aux données privées envoyées entre mon navigateur et Data Classification ?

Non. Les données privées envoyées entre votre navigateur et l'instance de classification des données sont sécurisées par un cryptage de bout en bout à l'aide de TLS 1.2, ce qui signifie que NetApp et les parties non NetApp ne peuvent pas les lire. La classification des données ne partagera aucune donnée ni aucun résultat avec NetApp, sauf si vous demandez et approuvez l'accès.

Les données analysées restent dans votre environnement.

Comment les données sensibles sont-elles traitées ?

NetApp n'a pas accès aux données sensibles et ne les affiche pas dans l'interface utilisateur. Les données sensibles sont masquées, par exemple, les quatre derniers chiffres sont affichés pour les informations de carte de crédit.

Où sont stockées les données ?

Les résultats de l'analyse sont stockés dans Elasticsearch au sein de votre instance de classification des données.

Comment accède-t-on aux données ?

La classification des données accède aux données stockées dans Elasticsearch via des appels API, qui nécessitent une authentification et sont cryptés à l'aide d'AES-128. L'accès direct à Elasticsearch nécessite un accès root.

Licences et coûts

La question suivante concerne les licences et les coûts d'utilisation de la classification des données.

Combien coûte la classification des données ?

La classification des données est une fonctionnalité principale de la NetApp Console . Ce n'est pas facturé.

Déploiement de l'agent de console

Les questions suivantes concernent l'agent de console.

Qu'est-ce que l'agent Console ?

L'agent de console est un logiciel exécuté sur une instance de calcul au sein de votre compte cloud ou sur site, qui permet à la NetApp Console de gérer en toute sécurité les ressources cloud. Vous devez déployer un agent de console pour utiliser la classification des données.

Où l'agent de console doit-il être installé ?

Lors de l'analyse des données, l'agent NetApp Console Console doit être installé aux emplacements suivants :

- Pour Cloud Volumes ONTAP dans AWS ou Amazon FSx pour ONTAP: l'agent de console se trouve dans AWS.
- Pour Cloud Volumes ONTAP dans Azure ou dans Azure NetApp Files: l'agent de console est dans Azure.
- Pour Cloud Volumes ONTAP dans GCP : l'agent de console est dans GCP.
- Pour les systèmes ONTAP sur site : l'agent de console est sur site.

Si vous avez des données à ces emplacements, vous devrez peut-être utiliser "plusieurs agents de console" .

La classification des données nécessite-t-elle l'accès à des informations d'identification ?

La classification des données elle-même ne récupère pas les informations d'identification de stockage. Au lieu de cela, ils sont stockés dans l'agent de la console.

La classification des données utilise les informations d'identification du plan de données, par exemple les informations d'identification CIFS pour monter les partages avant l'analyse.

La communication entre le service et l'agent de la console utilise-t-elle HTTP ?

Oui, Data Classification communique avec l'agent de la console via HTTP.

Déploiement de la classification des données

Les questions suivantes concernent l'instance distincte de classification des données.

Quels modèles de déploiement la classification des données prend-elle en charge ?

La NetApp Console permet à l'utilisateur d'analyser et de générer des rapports sur les systèmes pratiquement n'importe où, y compris sur site, dans le cloud et dans les environnements hybrides. La classification des données est normalement déployée à l'aide d'un modèle SaaS, dans lequel le service est activé via l'interface de la console et ne nécessite aucune installation matérielle ou logicielle. Même dans ce mode de déploiement « click-and-run », la gestion des données peut être effectuée indépendamment du fait que les magasins de données se trouvent sur site ou dans le cloud public.

Quel type d'instance ou de machine virtuelle est requis pour la classification des données ?

Quand"déployé dans le cloud" :

- Dans AWS, la classification des données s'exécute sur une instance m6i.4xlarge avec un disque GP2 de 500 Gio. Vous pouvez sélectionner un type d'instance plus petit lors du déploiement.
- Dans Azure, la classification des données s'exécute sur une machine virtuelle Standard_D16s_v3 avec un disque de 500 Gio.
- Dans GCP, la classification des données s'exécute sur une machine virtuelle n2-standard-16 avec un disque persistant standard de 500 Gio.

"En savoir plus sur le fonctionnement de la classification des données" .

Puis-je déployer la classification des données sur mon propre hôte ?

Oui. Vous pouvez installer le logiciel de classification des données sur un hôte Linux disposant d'un accès Internet sur votre réseau ou dans le cloud. Tout fonctionne de la même manière et vous continuez à gérer votre configuration d'analyse et vos résultats via la console. Voir"Déploiement de la classification des données sur site" pour la configuration système requise et les détails d'installation.

Qu'en est-il des sites sécurisés sans accès Internet ?

Oui, c'est également pris en charge. Tu peux"déployer la classification des données sur un site local qui n'a pas d'accès Internet" pour des sites entièrement sécurisés.

Mentions légales

Les mentions légales donnent accès aux déclarations de droits d'auteur, aux marques déposées, aux brevets et bien plus encore.

Copyright

"https://www.netapp.com/company/legal/copyright/"

Marques de commerce

NETAPP, le logo NETAPP et les marques répertoriées sur la page Marques NetApp sont des marques commerciales de NetApp, Inc. Les autres noms de sociétés et de produits peuvent être des marques commerciales de leurs propriétaires respectifs.

"https://www.netapp.com/company/legal/trademarks/"

Brevets

Une liste actuelle des brevets détenus par NetApp est disponible à l'adresse suivante :

https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf

Politique de confidentialité

"https://www.netapp.com/company/legal/privacy-policy/"

Open source

Les fichiers d'avis fournissent des informations sur les droits d'auteur et les licences tiers utilisés dans les logiciels NetApp .

- "Avis concernant la NetApp Console"
- "Avis relatif à la NetApp Data Classification"

Informations sur le copyright

Informations sur les marques commerciales