Empezar

NetApp Data Classification

NetApp
February 06, 2026

Tabla de contenidos

Empezar

Obtenga más información sobre la NetApp Data Classification

NetApp Data Classification es un servicio de gobernanza de datos para la NetApp Console que escanea sus fuentes de datos locales y en la nube corporativas para mapear y clasificar datos e identificar información privada. Esto puede ayudar a reducir el riesgo de seguridad y cumplimiento, disminuir los costos de almacenamiento y ayudarlo con sus proyectos de migración de datos.

A partir de la versión 1.31, la clasificación de datos está disponible como una capacidad principal dentro de la NetApp Console. No hay ningún cargo adicional No se requiere licencia de clasificación ni suscripción. + Si ha estado utilizando la versión heredada 1.30 o anterior, esa versión estará disponible hasta que expire su suscripción.

NetApp Console

Se puede acceder a la clasificación de datos a través de la NetApp Console.

La NetApp Console proporciona una gestión centralizada de los servicios de datos y almacenamiento de NetApp en entornos locales y en la nube a nivel empresarial. La consola es necesaria para acceder y utilizar los servicios de datos de NetApp . Como interfaz de administración, le permite administrar muchos recursos de almacenamiento desde una sola interfaz. Los administradores de la consola pueden controlar el acceso al almacenamiento y los servicios para todos los sistemas dentro de la empresa.

No necesita una licencia o suscripción para comenzar a usar NetApp Console y solo incurre en cargos cuando necesita implementar agentes de Console en su nube para garantizar la conectividad con sus sistemas de almacenamiento o servicios de datos de NetApp . Sin embargo, algunos servicios de datos de NetApp accesibles desde la consola requieren licencia o suscripción.

Obtenga más información sobre el"NetApp Console" .

Funciones

La clasificación de datos utiliza inteligencia artificial (IA), procesamiento del lenguaje natural (PLN) y aprendizaje automático (ML) para comprender el contenido que escanea con el fin de extraer entidades y categorizar el contenido en consecuencia. Esto permite que la clasificación de datos proporcione las siguientes áreas de funcionalidad.

"Conozca los casos de uso para la clasificación de datos" .

Mantener el cumplimiento

La clasificación de datos proporciona varias herramientas que pueden ayudarle con sus esfuerzos de cumplimiento. Puede utilizar la clasificación de datos para:

- Identificar información de identificación personal (PII).
- Identifique un amplio alcance de información personal confidencial según lo exigen las regulaciones de privacidad GDPR, CCPA, PCI y HIPAA.

- Responder a las solicitudes de acceso del titular de los datos (DSAR) basadas en el nombre o la dirección de correo electrónico.

Fortalecer la seguridad

La clasificación de datos puede identificar datos que potencialmente corren el riesgo de ser accedidos con fines delictivos. Puede utilizar la clasificación de datos para:

- Identifique todos los archivos y directorios (recursos compartidos y carpetas) con permisos abiertos que estén expuestos a toda su organización o al público.
- Identifique datos confidenciales que residen fuera de la ubicación inicial dedicada.
- Cumplir con las políticas de retención de datos.
- Utilice *Políticas* para detectar automáticamente nuevos problemas de seguridad para que el personal de seguridad pueda tomar medidas de inmediato.

Optimizar el uso del almacenamiento

La clasificación de datos proporciona herramientas que pueden ayudarle con el costo total de propiedad (TCO) de su almacenamiento. Puede utilizar la clasificación de datos para:

- Aumente la eficiencia del almacenamiento identificando datos duplicados o no relacionados con el negocio.
- Ahorre costos de almacenamiento identificando datos inactivos que puede clasificar en un almacenamiento de objetos menos costoso. "Obtenga más información sobre la organización en niveles de los sistemas Cloud Volumes ONTAP" . "Obtenga más información sobre la organización en niveles de los sistemas ONTAP locales" .

Sistemas y fuentes de datos compatibles

La clasificación de datos puede escanear y analizar datos estructurados y no estructurados de los siguientes tipos de sistemas y fuentes de datos:

Sistemas

- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- Cloud Volumes ONTAP (implementado en AWS, Azure o GCP)
- Clústeres ONTAP locales
- StorageGRID
- Google Cloud NetApp Volumes

Fuentes de datos

- Recursos compartidos de archivos de NetApp
- Bases de datos:
  - Servicio de base de datos relacional de Amazon (Amazon RDS)
  - MongoDB
  - MySQL
  - Oráculo

- PostgreSQL
- SAP HANA
- Servidor SQL (MSSQL)

La clasificación de datos admite las versiones NFS 3.x, 4.0 y 4.1, y las versiones CIFS 1.x, 2.0, 2.1 y 3.0.

Costo

La clasificación de datos es de uso gratuito. No se requiere licencia de clasificación ni suscripción paga.

Costos de infraestructura

- La instalación de Data Classification en la nube requiere implementar una instancia en la nube, lo que genera cargos por parte del proveedor de la nube donde se implementa. Ver el tipo de instancia que se implementa para cada proveedor de nube . No hay ningún costo si instala Data Classification en un sistema local.
- Para la clasificación de datos es necesario que haya implementado un agente de consola. En muchos casos, ya tienes un agente de consola debido a otro almacenamiento y servicios que estás usando en la consola. La instancia del agente de consola genera cargos del proveedor de la nube donde se implementa. Ver el "tipo de instancia que se implementa para cada proveedor de nube" . No hay ningún costo si instala el agente de consola en un sistema local.

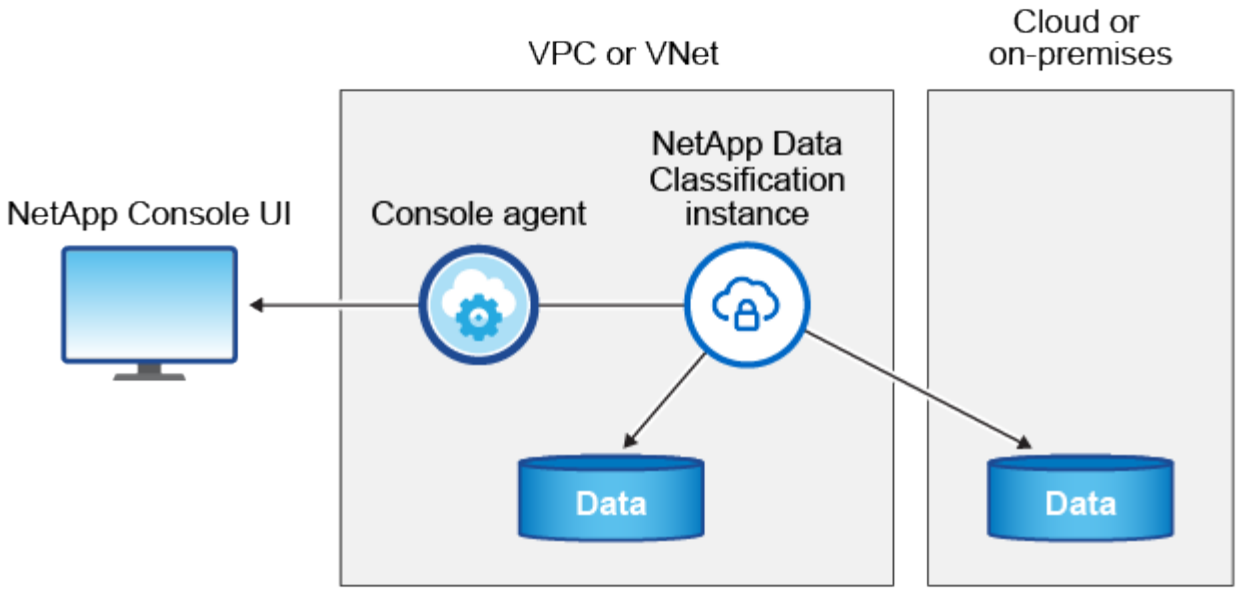
Costos de transferencia de datos

Los costos de transferencia de datos dependen de su configuración. Si la instancia de clasificación de datos y la fuente de datos están en la misma zona de disponibilidad y región, no hay costos de transferencia de datos. Pero si la fuente de datos, como un sistema Cloud Volumes ONTAP , está en una zona de disponibilidad o región *diferente*, su proveedor de nube le cobrará los costos de transferencia de datos. Consulte estos enlaces para obtener más detalles:

- "AWS: Precios de Amazon Elastic Compute Cloud (Amazon EC2)"
- "Microsoft Azure: Detalles de precios del ancho de banda"
- "Google Cloud: Precios del servicio de transferencia de almacenamiento"

La instancia de Clasificación de Datos

Cuando implementa la clasificación de datos en la nube, la consola implementa la instancia en la misma subred que el agente de la consola. "Obtenga más información sobre el agente de consola."



Tenga en cuenta lo siguiente sobre la instancia predeterminada:

- En AWS, la clasificación de datos se ejecuta en un "instancia m6i.4xlarge" con un disco GP2 de 500 GiB. La imagen del sistema operativo es Amazon Linux 2. Al implementar en AWS, puede elegir un tamaño de instancia más pequeño si está escaneando una pequeña cantidad de datos.
- En Azure, la clasificación de datos se ejecuta en un"Standard_D16s_v3 VM" con un disco de 500 GiB. La imagen del sistema operativo es Ubuntu 22.04.
- En GCP, la clasificación de datos se ejecuta en un"Máquina virtual n2-standard-16" con un disco persistente estándar de 500 GiB. La imagen del sistema operativo es Ubuntu 22.04.
- En las regiones donde la instancia predeterminada no está disponible, la clasificación de datos se ejecuta en una instancia alternativa. "Ver los tipos de instancias alternativos" .
- La instancia se llama *CloudCompliance* con un hash generado (UUID) concatenado a ella. Por ejemplo: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*
- Solo se implementa una instancia de clasificación de datos por agente de consola.

También puede implementar la clasificación de datos en un host Linux en sus instalaciones o en un host en su proveedor de nube preferido. El software funciona exactamente de la misma manera independientemente del método de instalación que elija. Las actualizaciones del software de clasificación de datos se automatizan siempre que la instancia tenga acceso a Internet.

La instancia debe permanecer en ejecución en todo momento porque la clasificación de datos escanea continuamente los datos.

Implementar en diferentes tipos de instancias

Revise las siguientes especificaciones para los tipos de instancias:

Tamaño del sistema	Especificaciones	Limitaciones
Extra grande	32 CPU, 128 GB de RAM, 1 TiB SSD	Puede escanear hasta 500 millones de archivos.

Tamaño del sistema	Especificaciones	Limitaciones
Grande (predeterminado)	16 CPU, 64 GB de RAM, SSD de 500 GiB	Puede escanear hasta 250 millones de archivos.

Al implementar la clasificación de datos en Azure o GCP, envíe un correo electrónico a ng-contact-data-sense@netapp.com para obtener ayuda si desea utilizar un tipo de instancia más pequeño.

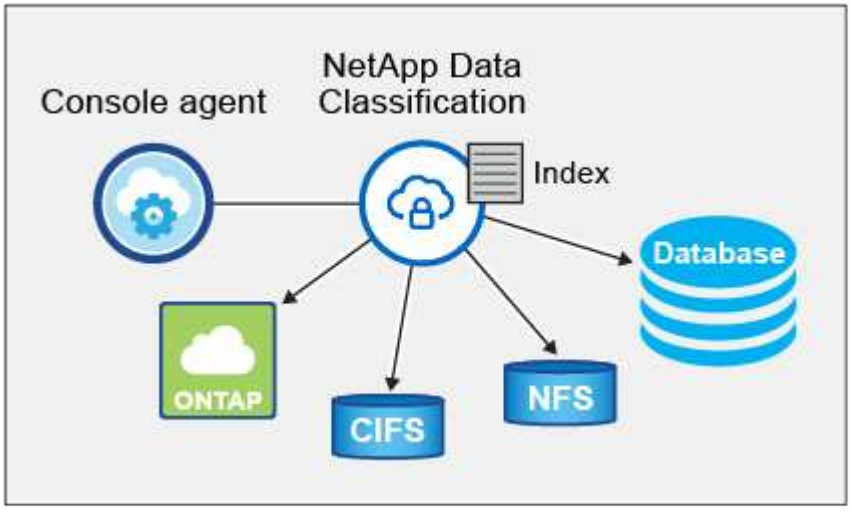
Cómo funciona el escaneo de clasificación de datos

A un alto nivel, el escaneo de clasificación de datos funciona así:

1. Implementa una instancia de Clasificación de datos en la consola.
2. Habilita el mapeo de alto nivel (llamados escaneos *Solo mapeo*) o el escaneo de nivel profundo (llamados escaneos *Mapeo y clasificación*) en una o más fuentes de datos.
3. La clasificación de datos escanea datos utilizando un proceso de aprendizaje de IA.
4. Puede utilizar los paneles y las herramientas de generación de informes proporcionados para ayudarle en sus esfuerzos de cumplimiento y gobernanza.

Después de habilitar la Clasificación de datos y seleccionar los repositorios que desea escanear (estos son los volúmenes, esquemas de bases de datos u otros datos de usuario), inmediatamente comienza a escanear los datos para identificar datos personales y confidenciales. En la mayoría de los casos, debe centrarse en escanear datos de producción en vivo en lugar de copias de seguridad, espejos o sitios de recuperación ante desastres. Luego, la clasificación de datos mapea los datos de su organización, categoriza cada archivo e identifica y extrae entidades y patrones predefinidos en los datos. El resultado del escaneo es un índice de información personal, información personal confidencial, categorías de datos y tipos de archivos.

La clasificación de datos se conecta a los datos como cualquier otro cliente montando volúmenes NFS y CIFS. A los volúmenes NFS se accede automáticamente como de solo lectura, mientras que es necesario proporcionar credenciales de Active Directory para escanear volúmenes CIFS.



Después del escaneo inicial, la clasificación de datos escanea continuamente sus datos de manera rotatoria para detectar cambios incrementales. Por eso es importante mantener la instancia en ejecución.

Puede habilitar y deshabilitar escaneos a nivel de volumen o a nivel de esquema de base de datos.

La clasificación de datos no impone un límite en la cantidad de datos que puede escanear. Cada agente de consola admite el escaneo y la visualización de 500 TiB de datos. Para escanear más de 500 TiB de datos,"instalar otro agente de consola" entonces"Implementar otra instancia de clasificación de datos" . + La interfaz de usuario de la consola muestra datos de un solo conector. Para obtener sugerencias sobre cómo ver datos de varios agentes de la consola, consulte"Trabajar con múltiples agentes de consola" .

¿Cuál es la diferencia entre los escaneos de mapeo y clasificación?

Puede realizar dos tipos de escaneos en Clasificación de datos:

- Los escaneos de solo mapeo brindan únicamente una descripción general de alto nivel de sus datos y se realizan en fuentes de datos seleccionadas. Los escaneos de solo mapeo toman menos tiempo que los escaneos de mapas y clasificación porque no acceden a los archivos para ver los datos dentro de ellos. Es posible que desees hacer esto inicialmente para identificar áreas de investigación y luego realizar un escaneo de Mapa y Clasificación en esas áreas.
- **Los escaneos de mapas y clasificación** proporcionan un escaneo de nivel profundo de sus datos.

Para obtener detalles sobre las diferencias entre los escaneos de mapeo y clasificación, consulte"¿Cuál es la diferencia entre los escaneos de mapeo y clasificación?" .

Información que la clasificación de datos categoriza

La clasificación de datos recopila, indexa y asigna categorías a los siguientes datos:

- **Metadatos estándar** sobre los archivos: el tipo de archivo, su tamaño, fechas de creación y modificación, etc.
- **Datos personales**: Información de identificación personal (PII), como direcciones de correo electrónico, números de identificación o números de tarjetas de crédito, que la clasificación de datos identifica mediante palabras, cadenas y patrones específicos en los archivos. "Obtenga más información sobre los datos personales" .
- **Datos personales sensibles**: Tipos especiales de información personal sensible (IPS), como datos de salud, origen étnico u opiniones políticas, según lo define el Reglamento General de Protección de Datos (RGPD) y otras regulaciones de privacidad. "Obtenga más información sobre datos personales sensibles" .
- **Categorías**: La clasificación de datos toma los datos que escanea y los divide en diferentes tipos de categorías. Las categorías son temas basados en el análisis de IA del contenido y los metadatos de cada archivo. "Obtenga más información sobre las categorías".
- **Reconocimiento de entidades de nombre**: la clasificación de datos utiliza IA para extraer los nombres naturales de las personas de los documentos. "Obtenga información sobre cómo responder a las solicitudes de acceso de los interesados" .

Descripción general de la red

Data Classification implementa un solo servidor o clúster donde usted elija: en la nube o en las instalaciones. Los servidores se conectan a través de protocolos estándar a las fuentes de datos e indexan los resultados en un clúster Elasticsearch, que también está implementado en los mismos servidores. Esto permite compatibilidad con entornos multicloud, cross-cloud, cloud privado y locales.

La consola implementa la instancia de clasificación de datos con un grupo de seguridad que habilita conexiones HTTP entrantes desde el agente de la consola.

Cuando usa la consola en modo SaaS, la conexión a la consola se proporciona a través de HTTPS y los datos privados enviados entre su navegador y la instancia de clasificación de datos están protegidos con cifrado de extremo a extremo mediante TLS 1.2, lo que significa que NetApp y terceros no pueden leerlos.

Las reglas de salida están completamente abiertas. Se necesita acceso a Internet para instalar y actualizar el software de clasificación de datos y para enviar métricas de uso.

Si tiene requisitos de red estrictos,"Obtenga información sobre los puntos finales con los que se comunica la clasificación de datos" .

Acceso a la NetApp Data Classification

Puede acceder a la NetApp Data Classification a través de la NetApp Console.

Para iniciar sesión en la consola, puede usar sus credenciales del sitio de soporte de NetApp o puede registrarse para iniciar sesión en la NetApp Console usando su correo electrónico y una contraseña. "Obtenga más información sobre cómo iniciar sesión en la consola" .

Tareas específicas requieren roles de usuario de consola específicos. "Obtenga información sobre los roles de acceso a la consola para todos los servicios" .

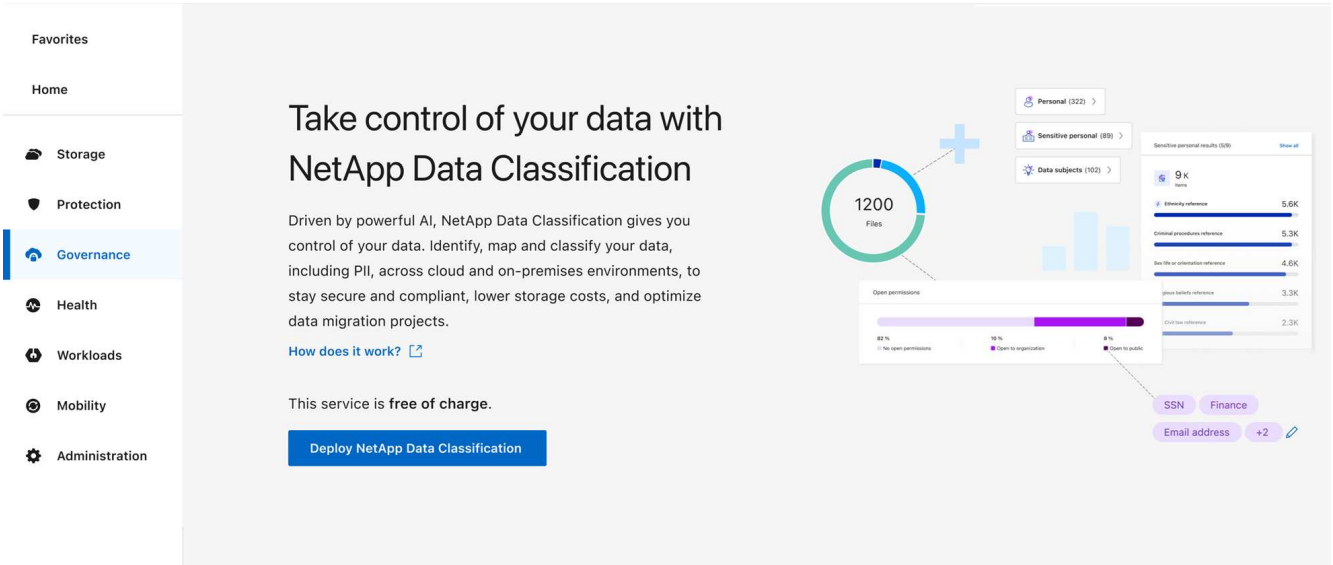
Antes de empezar

- "Debes agregar un agente de consola."
- "Comprenda qué estilo de implementación de clasificación de datos se adapta a su carga de trabajo."

Pasos

1. En un navegador web, navegue hasta el"Consola" .
2. Inicie sesión en la consola.
3. Desde la página principal de la NetApp Console, seleccione **Gobernanza** > **Clasificación de datos**.
4. Si es la primera vez que accede a Clasificación de datos, aparecerá la página de destino.

   Seleccione **Implementar clasificación local o en la nube** para comenzar a implementar su instancia de clasificación. Para obtener más información, consulte"¿Qué implementación de clasificación de datos debería utilizar?"



De lo contrario, aparecerá el Panel de clasificación de datos.

Implementar la clasificación de datos

¿Qué implementación de NetApp Data Classification debería utilizar?

Puede implementar NetApp Data Classification de diferentes maneras. Conozca qué método se adapta a sus necesidades.

La clasificación de datos se puede implementar de las siguientes maneras:

- "Implementar en la nube usando la consola" . La consola implementa la instancia de clasificación de datos en la misma red del proveedor de nube que el agente de la consola.
- "Instalar en un host Linux con acceso a Internet" . Instale Data Classification en un host Linux en su red, o en un host Linux en la nube, que tenga acceso a Internet. Este tipo de instalación puede ser una buena opción si prefiere escanear sistemas ONTAP locales utilizando una instancia de clasificación de datos que también se encuentre en las instalaciones, aunque esto no es un requisito.
- "Instalar en un host Linux en un sitio local sin acceso a Internet", también conocido como *modo privado*. Este tipo de instalación, que utiliza un script de instalación, no tiene conectividad con la capa SaaS de la consola.

> El modo privado de BlueXP (interfaz BlueXP heredada) generalmente se usa con entornos locales que no tienen conexión a Internet y con regiones de nube seguras, que incluyen AWS Secret Cloud, AWS Top Secret Cloud y Azure IL6. NetApp continúa brindando soporte a estos entornos con la interfaz BlueXP heredada. Para obtener documentación del modo privado en la interfaz heredada de BlueXP , consulte"Documentación en PDF para el modo privado de BlueXP" .

Tanto la instalación en un host Linux con acceso a Internet como la instalación local en un host Linux sin acceso a Internet utilizan un script de instalación. El script comienza verificando si el sistema y el entorno cumplen los requisitos previos. Si se cumplen los requisitos previos, se inicia la instalación. Si desea verificar los requisitos previos independientemente de ejecutar la instalación de Clasificación de datos, hay un paquete de software separado que puede descargar que solo prueba los requisitos previos.

Consulte "Compruebe que su host Linux esté listo para instalar Data Classification" .

Implemente la NetApp Data Classification en la nube mediante la NetApp Console

Puede implementar NetApp Data Classification en la nube con la NetApp Console. La consola implementa la instancia de clasificación de datos en la misma red del proveedor de nube que el agente de la consola.

Tenga en cuenta que también puede"Instalar Data Classification en un host Linux que tenga acceso a Internet" . Este tipo de instalación puede ser una buena opción si prefiere escanear sistemas ONTAP locales utilizando una instancia de clasificación de datos que también se encuentra en las instalaciones, pero esto no es un requisito. El software funciona exactamente de la misma manera independientemente del método de instalación que elija.

Inicio rápido

Comience rápidamente siguiendo estos pasos o desplácese hacia abajo hasta las secciones restantes para obtener detalles completos.

Crear un agente de consola

Si aún no tiene un agente de consola, cree uno. Ver"Creación de un agente de consola en AWS" ,"Creación de un agente de consola en Azure" , o"Creación de un agente de consola en GCP" .

También puedes"Instalar el agente de consola local" en un host Linux de su red o en un host Linux en la nube.

Prerrequisitos

Asegúrate de que tu entorno puede cumplir los requisitos previos. Esto incluye acceso saliente a internet para la instancia, conectividad entre el agente de la Console y Data Classification por el puerto 443 y más. Ver la lista completa.

Implementar la clasificación de datos

Inicie el asistente de instalación para implementar la instancia de Clasificación de datos en la nube.

Crear un agente de consola

Si aún no tiene un agente de consola, cree uno en su proveedor de nube. Ver "Creación de un agente de consola en AWS" o "Creación de un agente de consola en Azure" , o "Creación de un agente de consola en GCP" . En la mayoría de los casos, probablemente ya tendrá configurado un agente de consola antes de intentar activar la clasificación de datos, ya que la mayoría "Las funciones de la consola requieren un agente de consola" , pero hay casos en los que necesitarás configurarlo ahora.

Hay algunos escenarios en los que es necesario utilizar un agente de consola implementado en un proveedor de nube específico:

- Al escanear datos en Cloud Volumes ONTAP en AWS o Amazon FSx para buckets de ONTAP , se utiliza un agente de consola en AWS.
- Al escanear datos en Cloud Volumes ONTAP en Azure o en Azure NetApp Files, se utiliza un agente de consola en Azure.
  - Para Azure NetApp Files, debe implementarse en la misma región que los volúmenes que desea escanear.
- Al escanear datos en Cloud Volumes ONTAP en GCP, se utiliza un agente de consola en GCP.

Los sistemas ONTAP locales, los recursos compartidos de archivos de NetApp y las bases de datos se pueden escanear al usar cualquiera de estos agentes de consola en la nube.

Ten en cuenta que también puedes "Instalar el agente de consola local" en un host Linux de su red o en la nube. Algunos usuarios que planean instalar Data Classification localmente también pueden optar por instalar el agente de consola localmente.

Puede haber situaciones en las que necesites usar "varios agentes de consola" .

La clasificación de datos no impone un límite en la cantidad de datos que puede escanear. Cada agente de consola admite el escaneo y la visualización de 500 TiB de datos. Para escanear más de 500 TiB de datos,"instalar otro agente de consola" entonces"Implementar otra instancia de clasificación de datos" . + La interfaz de usuario de la consola muestra datos de un solo conector. Para obtener sugerencias sobre cómo ver datos de varios agentes de la consola, consulte"Trabajar con múltiples agentes de consola" .

Apoyo de la región gubernamental

La clasificación de datos se admite cuando el agente de consola se implementa en una región gubernamental (AWS GovCloud, Azure Gov o Azure DoD). Cuando se implementa de esta manera, la clasificación de datos tiene las siguientes restricciones:

"Obtenga información sobre cómo implementar el agente de consola en una región gubernamental.".

Prerrequisitos

Revise los siguientes requisitos previos para asegurarse de tener una configuración compatible antes de implementar la clasificación de datos en la nube. Cuando implementa la clasificación de datos en la nube, se ubica en la misma subred que el agente de consola.

Habilitar el acceso a Internet saliente desde la Clasificación de datos

La clasificación de datos requiere acceso a Internet saliente. Si su red virtual o física utiliza un servidor proxy para el acceso a Internet, asegúrese de que la instancia de clasificación de datos tenga acceso a Internet saliente para contactar los siguientes puntos finales. El proxy no debe ser transparente. Los servidores proxy transparentes no son compatibles actualmente.

Revise la tabla correspondiente a continuación según si está implementando la clasificación de datos en AWS, Azure o GCP.

Puntos finales necesarios para AWS

Puntos finales	Objetivo
\ https://api.console.netapp.com	Comunicación con el servicio de consola, que incluye cuentas de NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Comunicación con el sitio web de la consola para la autenticación centralizada de usuarios.
\ 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/	Proporciona acceso a imágenes de software, manifiestos y plantillas.
\ https://kinesis.us-east-1.amazonaws.com	Permite a NetApp transmitir datos desde registros de auditoría.
\ 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	Permite que la clasificación de datos acceda y descargue manifiestos y plantillas, y envíe registros y métricas.

Puntos de conexión necesarios para Azure

Puntos finales	Objetivo
\ https://api.console.netapp.com	Comunicación con el servicio de consola, que incluye cuentas de NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Comunicación con el sitio web de la consola para la autenticación centralizada de usuarios.
\ 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/	Proporciona acceso a imágenes de software, manifiestos, plantillas y permite enviar registros y métricas.
\ https://support.compliance.api.console.netapp.com/	Permite a NetApp transmitir datos desde registros de auditoría.

Puntos finales necesarios para GCP

Puntos finales	Objetivo
\ https://api.console.netapp.com	Comunicación con el servicio de consola, que incluye cuentas de NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Comunicación con el sitio web de la consola para la autenticación centralizada de usuarios.

Puntos finales	Objetivo
\ 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/	Proporciona acceso a imágenes de software, manifiestos, plantillas y permite enviar registros y métricas.
\ https://support.compliance.api.console.netapp.com/	Permite a NetApp transmitir datos desde registros de auditoría.

Asegúrese de que la clasificación de datos tenga los permisos necesarios

Asegúrese de que la clasificación de datos tenga permisos para implementar recursos y crear grupos de seguridad para la instancia de clasificación de datos.

- "Permisos de Google Cloud"
- "Permisos de AWS"
- "Permisos de Azure"

Asegúrese de que el agente de la consola pueda acceder a la clasificación de datos

Asegúrese de la conectividad entre el agente de la consola y la instancia de clasificación de datos. El grupo de seguridad del agente de consola debe permitir el tráfico entrante y saliente a través del puerto 443 hacia y desde la instancia de clasificación de datos. Esta conexión permite la implementación de la instancia de Clasificación de datos y le permite ver información en las pestañas Cumplimiento y Gobernanza. La clasificación de datos es compatible con las regiones gubernamentales en AWS y Azure.

Se requieren reglas de grupo de seguridad entrantes y salientes adicionales para las implementaciones de AWS y AWS GovCloud. Ver "Reglas para el agente de consola en AWS" Para más detalles.

Se requieren reglas de grupo de seguridad entrantes y salientes adicionales para las implementaciones de Azure y Azure Government. Ver "Reglas para el agente de consola en Azure" Para más detalles.

Asegúrese de poder mantener la clasificación de datos en funcionamiento

La instancia de Clasificación de datos debe permanecer activada para escanear continuamente sus datos.

Asegúrese de que el navegador web esté conectado a la clasificación de datos.

Una vez habilitada la clasificación de datos, asegúrese de que los usuarios accedan a la interfaz de la consola desde un host que tenga una conexión a la instancia de clasificación de datos.

La instancia de clasificación de datos utiliza una dirección IP privada para garantizar que los datos indexados no sean accesibles a través de Internet. Como resultado, el navegador web que utiliza para acceder a la Consola debe tener una conexión a esa dirección IP privada. Esa conexión puede provenir de una conexión directa a su proveedor de nube (por ejemplo, una VPN) o de un host que esté dentro de la misma red que la instancia de clasificación de datos.

Comprueba los límites de tu vCPU

Asegúrese de que el límite de vCPU de su proveedor de nube permita la implementación de una instancia con la cantidad necesaria de núcleos. Necesitará verificar el límite de vCPU para la familia de instancias relevante en la región donde se ejecuta la consola. "Ver los tipos de instancia requeridos" .

Consulte los siguientes enlaces para obtener más detalles sobre los límites de vCPU:

- "Documentación de AWS: Cuotas de servicio de Amazon EC2"
- "Documentación de Azure: Cuotas de vCPU de máquinas virtuales"
- "Documentación de Google Cloud: Cuotas de recursos"

Implementar la clasificación de datos en la nube

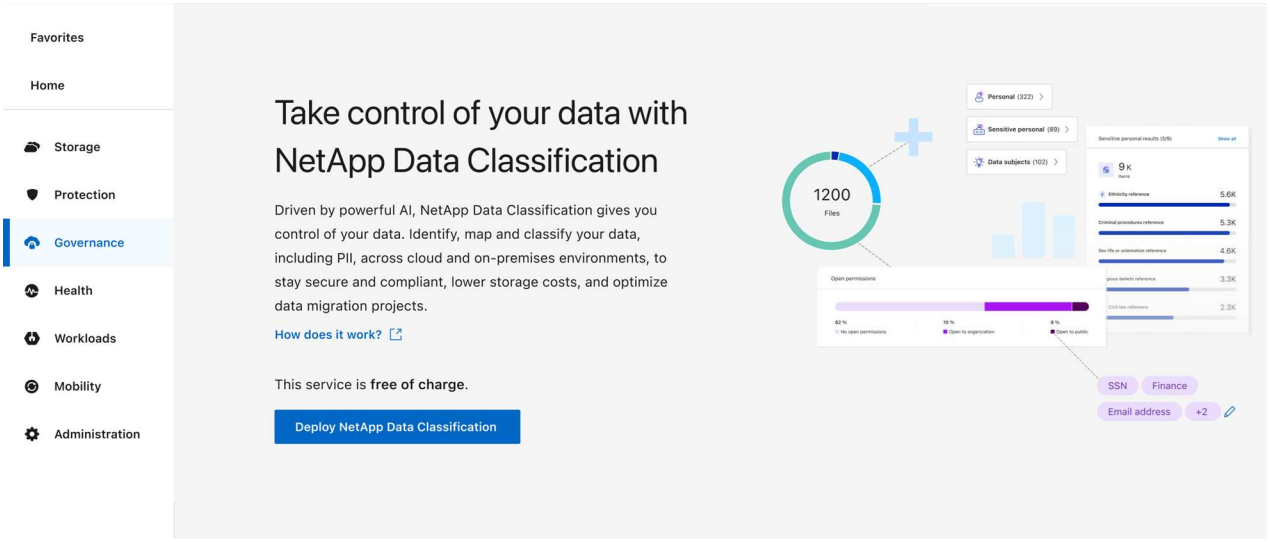
Siga estos pasos para implementar una instancia de Clasificación de datos en la nube. El agente de la consola implementará la instancia en la nube y luego instalará el software de clasificación de datos en esa instancia.

En las regiones donde el tipo de instancia predeterminado no está disponible, la clasificación de datos se ejecuta en un"tipo de instancia alternativo" .

Implementar en AWS

Pasos

1. Desde la página principal de Clasificación de datos, seleccione **Implementar clasificación en las instalaciones o en la nube**.

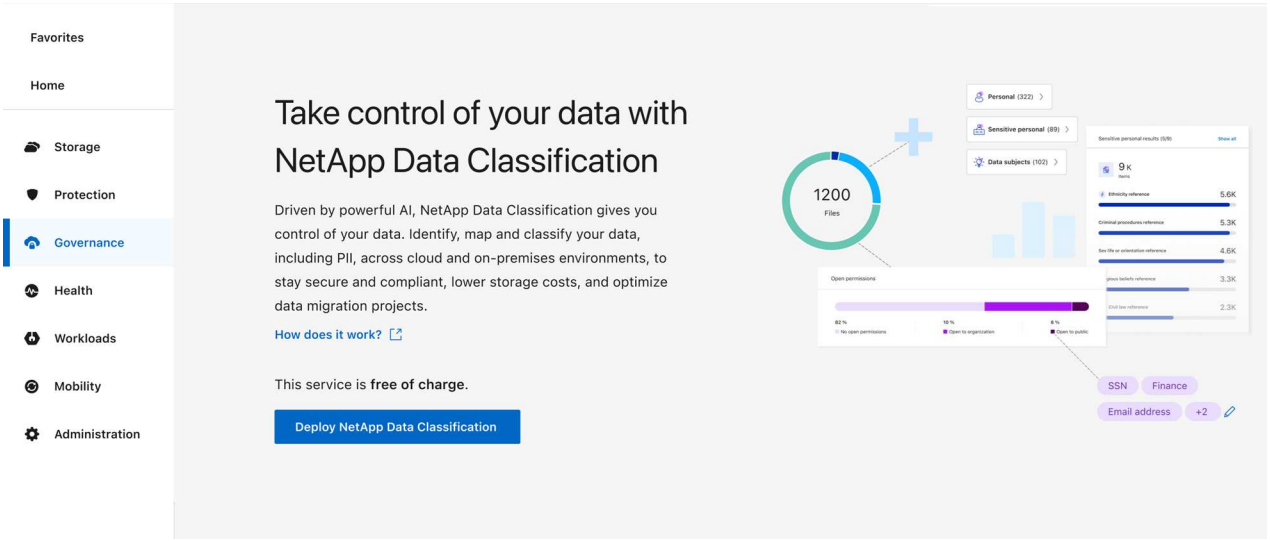


2. Desde la página *Instalación*, seleccione **Implementar > Implementar** para usar el tamaño de instancia "Grande" e iniciar el asistente de implementación en la nube.
3. El asistente muestra el progreso a medida que avanza en los pasos de implementación. Cuando se requieren entradas o si surgen problemas, se le solicitará información.
4. Cuando se implementa la instancia y se instala la Clasificación de datos, seleccione **Continuar con la configuración** para ir a la página *Configuración*.

Implementar en Azure

Pasos

1. Desde la página principal de Clasificación de datos, seleccione **Implementar clasificación en las instalaciones o en la nube**.



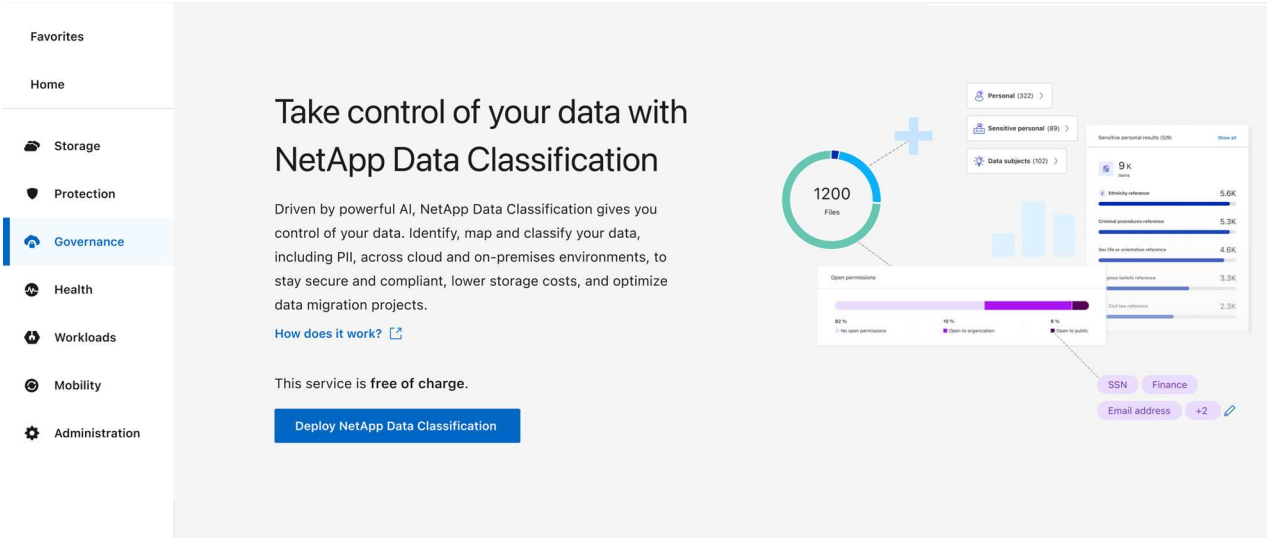
2. Seleccione **Implementar** para iniciar el asistente de implementación en la nube.

3. El asistente muestra el progreso a medida que avanza en los pasos de implementación. Se detendrá y solicitará información si surge algún problema.
4. Cuando se implementa la instancia y se instala la Clasificación de datos, seleccione **Continuar con la configuración** para ir a la página *Configuración*.

Implementar en Google Cloud

Pasos

1. Desde la página principal de Clasificación de datos, seleccione **Gobernanza > Clasificación**.
2. Seleccione **Implementar clasificación local o en la nube**.



3. Seleccione **Implementar** para iniciar el asistente de implementación en la nube.
4. El asistente muestra el progreso a medida que avanza en los pasos de implementación. Se detendrá y solicitará información si surge algún problema.
5. Cuando se implementa la instancia y se instala la Clasificación de datos, seleccione **Continuar con la configuración** para ir a la página *Configuración*.

Resultado

La consola implementa la instancia de clasificación de datos en su proveedor de nube.

Las actualizaciones del agente de consola y del software de clasificación de datos se automatizan siempre que las instancias tengan conectividad a Internet.

¿Qué sigue?

Desde la página de Configuración puede seleccionar las fuentes de datos que desea escanear.

Instalar NetApp Data Classification en un host que tenga acceso a Internet

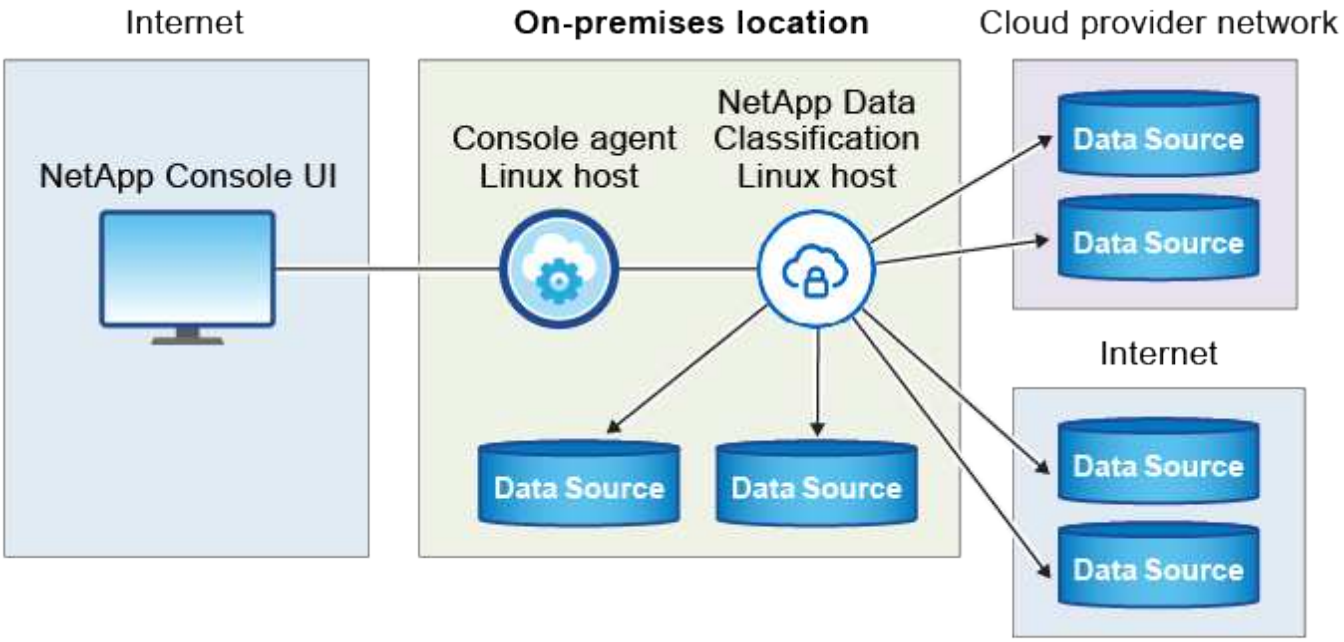
Para implementar NetApp Data Classification en un host Linux en su red o en un host Linux en la nube que tenga acceso a Internet, debe implementar el host Linux manualmente en su red o en la nube.

La instalación local es una buena opción si prefiere escanear sistemas ONTAP locales utilizando una instancia de clasificación de datos que también se encuentra en las instalaciones. Esto no es un requisito. El software

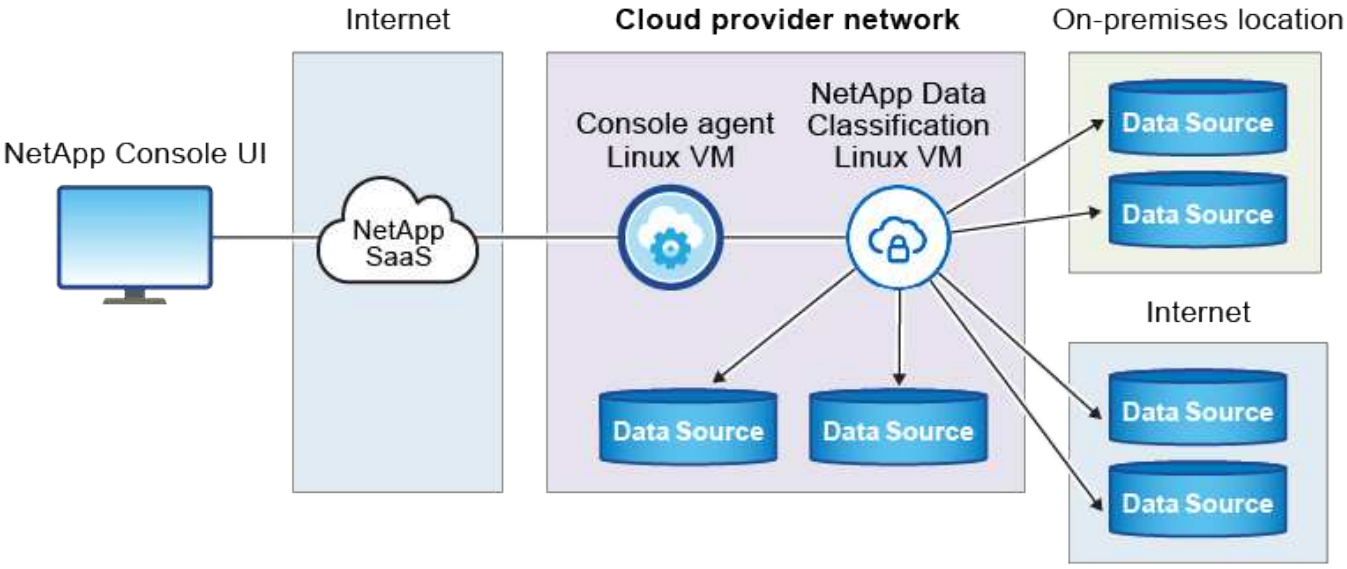
funciona de la misma manera independientemente del método de instalación que elija.

El script de instalación de Clasificación de datos comienza verificando si el sistema y el entorno cumplen los requisitos previos requeridos. Si se cumplen todos los requisitos previos, se inicia la instalación. Si desea verificar los requisitos previos independientemente de ejecutar la instalación de Clasificación de datos, hay un paquete de software separado que puede descargar que solo prueba los requisitos previos. "Vea cómo comprobar si su host Linux está listo para instalar la Clasificación de Datos" .

La instalación típica en un host Linux *en sus instalaciones* tiene los siguientes componentes y conexiones.



La instalación típica en un host Linux *en la nube* tiene los siguientes componentes y conexiones.



Inicio rápido

Comience rápidamente siguiendo estos pasos o desplácese hacia abajo hasta las secciones restantes para obtener detalles completos.

1 Crear un agente de consola

Si aún no tienes un agente de consola, "Implementar el agente de consola localmente" en un host Linux en su red o en un host Linux en la nube.

También puedes crear un agente de consola con tu proveedor de nube. Ver "Creación de un agente de consola en AWS" , "Creación de un agente de consola en Azure" , o "Creación de un agente de consola en GCP" .

2 Revisar los prerrequisitos

Asegúrese de que su entorno pueda cumplir con los requisitos previos. Esto incluye acceso a Internet saliente para la instancia, conectividad entre el agente de la consola y la clasificación de datos a través del puerto 443 y más. Ver la lista completa .

También necesitas un sistema Linux que cumpla con los requisitossiguientes requisitos .

3 Descargar e implementar la clasificación de datos

Descargue el software Cloud Data Classification del sitio de soporte de NetApp y copie el archivo de instalación en el host Linux que planea utilizar. Luego, inicie el asistente de instalación y siga las instrucciones para implementar la instancia de Clasificación de datos.

Crear un agente de consola

Se requiere un agente de consola antes de poder instalar y utilizar la clasificación de datos. En la mayoría de los casos, probablemente tendrá un agente de consola configurado antes de intentar activar la clasificación de datos porque la mayoría "Las funciones de la consola requieren un agente de consola" , pero habrá casos en los que necesitarás configurar uno ahora.

Para crear uno en su entorno de proveedor de nube, consulte "Creación de un agente de consola en AWS" , "Creación de un agente de consola en Azure" , o "Creación de un agente de consola en GCP" .

Hay algunos escenarios en los que es necesario utilizar un agente de consola implementado en un proveedor de nube específico:

- Al escanear datos en Cloud Volumes ONTAP en AWS o Amazon FSx para ONTAP, se utiliza un agente de consola en AWS.
- Al escanear datos en Cloud Volumes ONTAP en Azure o en Azure NetApp Files, se utiliza un agente de consola en Azure.

  Para Azure NetApp Files, debe implementarse en la misma región que los volúmenes que desea escanear.

- Al escanear datos en Cloud Volumes ONTAP en GCP, se utiliza un agente de consola en GCP.

Los sistemas ONTAP locales, los recursos compartidos de archivos de NetApp y las cuentas de bases de datos se pueden escanear utilizando cualquiera de estos agentes de consola en la nube.

Tenga en cuenta que también puede "Implementar el agente de consola localmente" en un host Linux en su red o en un host Linux en la nube. Algunos usuarios que planean instalar Data Classification en sus instalaciones también pueden optar por instalar el agente de consola en sus instalaciones.

Necesitará la dirección IP o el nombre de host del sistema del agente de consola al instalar Clasificación de datos. Tendrás esta información si instalaste el agente de consola en tus instalaciones. Si el agente de la consola está implementado en la nube, puede encontrar esta información en la consola: seleccione el ícono Ayuda, luego **Soporte** y luego **Agente de consola**.

Preparar el sistema host Linux

El software de clasificación de datos debe ejecutarse en un host que cumpla con los requisitos específicos del sistema operativo, requisitos de RAM, requisitos de software, etc. El host Linux puede estar en su red o en la nube.

Asegúrese de poder mantener la clasificación de datos en ejecución. La máquina de clasificación de datos debe permanecer encendida para escanear continuamente sus datos.

- La clasificación de datos debe realizarse en un host dedicado. El host no se puede compartir con otras aplicaciones o software de terceros, como antivirus.
- Elija el tamaño que se alinee con el conjunto de datos que planea escanear con Clasificación de datos.

Tamaño del sistema	UPC	RAM (la memoria de intercambio debe estar deshabilitada)	Disco
Extra grande	32 CPU	128 GB de RAM	• SSD de 1 TiB en /, o 100 GiB disponibles en /opt • 895 GiB disponibles en /var/lib/docker • 5 GiB en /tmp • **Para Podman, 30 GB en /var/tmp**
Grande	16 CPU	64 GB de RAM	• SSD de 500 GiB en /, o 100 GiB disponibles en /opt • 400 GiB disponibles en /var/lib/docker o para Podman /var/lib/containers • 5 GiB en /tmp • **Para Podman, 30 GB en /var/tmp**

- Al implementar una instancia de cómputo en la nube para su instalación de Clasificación de datos, se recomienda utilizar un sistema que cumpla con los requisitos del sistema "Grande" mencionados anteriormente:
  - **Tipo de instancia de Amazon Elastic Compute Cloud (Amazon EC2)**: "m6i.4xlarge". "Ver tipos de instancias de AWS adicionales" .
  - **Tamaño de máquina virtual de Azure**: "Standard_D16s_v3". "Ver tipos de instancias de Azure adicionales" .
  - **Tipo de máquina GCP**: "n2-standard-16". "Ver tipos de instancias de GCP adicionales" .
- **Permisos de carpeta UNIX**: Se requieren los siguientes permisos mínimos de UNIX:

Carpeta	Permisos mínimos
/tmp	`rwxrwxrwt`
/optar	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/sistema	`rwxr-xr-x`

- **Sistema operativo**:
  - Los siguientes sistemas operativos requieren el uso del motor de contenedores Docker:
    - Red Hat Enterprise Linux versión 7.8 y 7.9
    - Ubuntu 22.04 (requiere la versión 1.23 o superior de Data Classification)
    - Ubuntu 24.04 (requiere la versión 1.23 o superior de Data Classification)
  - Los siguientes sistemas operativos requieren el uso del motor de contenedores Podman y requieren la versión 1.30 o superior de Data Classification:
    - Red Hat Enterprise Linux versiones 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 y 9.6.
  - Las extensiones vectoriales avanzadas (AVX2) deben estar habilitadas en el sistema host.
- **Gestión de suscripciones de Red Hat**: el host debe estar registrado en Gestión de suscripciones de Red Hat. Si no está registrado, el sistema no puede acceder a los repositorios para actualizar el software de terceros requerido durante la instalación.
- **Software adicional**: Debe instalar el siguiente software en el host antes de instalar Data Classification:
  - Dependiendo del sistema operativo que estés usando, necesitas instalar uno de los motores de contenedores:
    - Docker Engine versión 19.3.1 o superior. "Ver instrucciones de instalación" .
    - Podman versión 4 o superior. Para instalar Podman, ingrese(`sudo yum install podman netavark -y`).
- Versión de Python 3.6 o superior. "Ver instrucciones de instalación" .
  - **Consideraciones sobre NTP**: NetApp recomienda configurar el sistema de clasificación de datos para utilizar un servicio de Protocolo de tiempo de red (NTP). La hora debe estar sincronizada entre el sistema de clasificación de datos y el sistema del agente de consola.
- **Consideraciones sobre Firewalld**: Si planea utilizar `firewalld` Le recomendamos que lo habilite antes de instalar Data Classification. Ejecute los siguientes comandos para configurar `firewalld` para que sea compatible con la Clasificación de Datos:

```
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 planea utilizar hosts de clasificación de datos adicionales como nodos de escáner, agregue estas reglas a su sistema principal en este momento:

```
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
```

Tenga en cuenta que debe reiniciar Docker o Podman cada vez que habilite o actualice `firewalld` ajustes.

La dirección IP del sistema host de clasificación de datos no se puede cambiar después de la instalación.

Habilitar el acceso a Internet saliente desde la Clasificación de datos

La clasificación de datos requiere acceso a Internet saliente. Si su red virtual o física utiliza un servidor proxy para el acceso a Internet, asegúrese de que la instancia de clasificación de datos tenga acceso a Internet saliente para contactar los siguientes puntos finales.

Puntos finales	Objetivo
\ https://api.console.netapp.com	Comunicación con la consola, que incluye cuentas de NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Comunicación con el sitio web de la consola para la autenticación centralizada de usuarios.
https://support.compliance.api.bluexp.netapp.com/ https://hub.docker.com https://auth.docker.io https://registry-1.docker.io https://index.docker.io/ https://dseasb33srnrn.cloudfront.net/ https://production.cloudflare.docker.com/	Proporciona acceso a imágenes de software, manifiestos, plantillas y permite enviar registros y métricas.
https://support.compliance.api.bluexp.netapp.com/	Permite a NetApp transmitir datos desde registros de auditoría.
\ https://github.com/docker \ https://download.docker.com	Proporciona paquetes de requisitos previos para la instalación de Docker.
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Proporciona paquetes de requisitos previos para la instalación de Ubuntu.

Verifique que todos los puertos requeridos estén habilitados

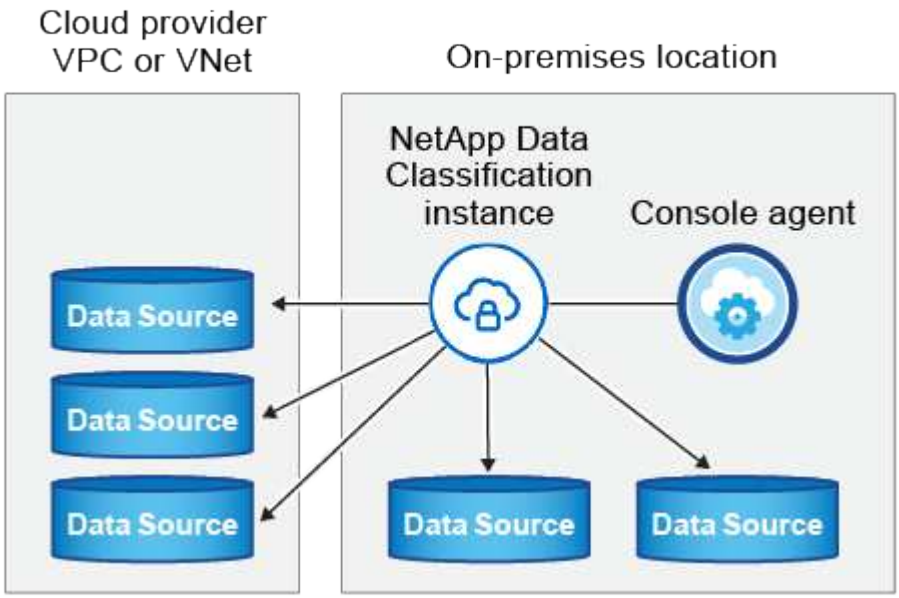
Debe asegurarse de que todos los puertos necesarios estén abiertos para la comunicación entre el agente de la consola, la clasificación de datos, Active Directory y sus fuentes de datos.

Tipo de conexión	Puertos	Descripción
Agente de consola <> Clasificación de datos	8080 (TCP), 443 (TCP) y 80. 9000	Las reglas de firewall o enrutamiento para el agente de la consola deben permitir el tráfico entrante y saliente a través del puerto 443 hacia y desde la instancia de clasificación de datos. Asegúrese de que el puerto 8080 esté abierto para que pueda ver el progreso de la instalación en la consola. Si se utiliza un firewall en el host Linux, se requiere el puerto 9000 para los procesos internos dentro de un servidor Ubuntu.
Agente de consola <> clúster ONTAP (NAS)	443 (TCP)	La consola descubre clústeres ONTAP mediante HTTPS. Si utiliza políticas de firewall personalizadas, deben cumplir los siguientes requisitos: • El host del agente de la consola debe permitir el acceso HTTPS saliente a través del puerto 443. Si el agente de la consola está en la nube, toda comunicación saliente está permitida por el firewall predefinido o las reglas de enrutamiento. • El clúster ONTAP debe permitir el acceso HTTPS entrante a través del puerto 443. La política de firewall predeterminada "mgmt" permite el acceso HTTPS entrante desde todas las direcciones IP. Si modificó esta política predeterminada o si creó su propia política de firewall, debe asociar el protocolo HTTPS con esa política y habilitar el acceso desde el host del agente de la Consola.
Clasificación de datos <> Clúster ONTAP	• Para NFS - 111 (TCP\UDP) y 2049 (TCP\UDP) • Para CIFS - 139 (TCP\UDP) y 445 (TCP\UDP)	La clasificación de datos necesita una conexión de red a cada subred de Cloud Volumes ONTAP o al sistema ONTAP local. Los firewalls o las reglas de enrutamiento para Cloud Volumes ONTAP deben permitir conexiones entrantes desde la instancia de clasificación de datos. Asegúrese de que estos puertos estén abiertos para la instancia de clasificación de datos: • Para NFS - 111 y 2049 • Para CIFS - 139 y 445 Las políticas de exportación de volumen NFS deben permitir el acceso desde la instancia de clasificación de datos.

Tipo de conexión	Puertos	Descripción
Clasificación de datos <> Active Directory	389 (TCP y UDP), 636 (TCP), 3268 (TCP) y 3269 (TCP)	Debe tener un Directorio Activo ya configurado para los usuarios de su empresa. Además, la clasificación de datos necesita credenciales de Active Directory para escanear volúmenes CIFS. Debes tener la información del Directorio Activo: • Dirección IP del servidor DNS o varias direcciones IP • Nombre de usuario y contraseña para el servidor • Nombre de dominio (nombre de Active Directory) • Ya sea que esté utilizando LDAP seguro (LDAPS) o no • Puerto del servidor LDAP (normalmente 389 para LDAP y 636 para LDAP seguro)

Instalar la clasificación de datos en el host Linux

Para configuraciones típicas, instalará el software en un solo sistema host. Vea esos pasos aquí .



VerPreparación del sistema host Linux yRevisión de prerrequisitos para obtener la lista completa de requisitos antes de implementar la clasificación de datos.

Las actualizaciones del software de clasificación de datos se automatizan siempre que la instancia tenga conectividad a Internet.

Actualmente, la clasificación de datos no puede escanear depósitos S3, Azure NetApp Files o FSx para ONTAP cuando el software está instalado en las instalaciones. En estos casos, necesitará implementar un agente de consola independiente y una instancia de clasificación de datos en la nube y "cambiar entre conectores" para sus diferentes fuentes de datos.

Instalación de un solo host para configuraciones típicas

Revise los requisitos y siga estos pasos al instalar el software de clasificación de datos en un solo host local.

"Mira este vídeo"para ver cómo instalar Clasificación de Datos.

Tenga en cuenta que todas las actividades de instalación se registran al instalar Data Classification. Si surge algún problema durante la instalación, puede ver el contenido del registro de auditoría de la instalación. Esta escrito para `/opt/netapp/install_logs/` .

Antes de empezar

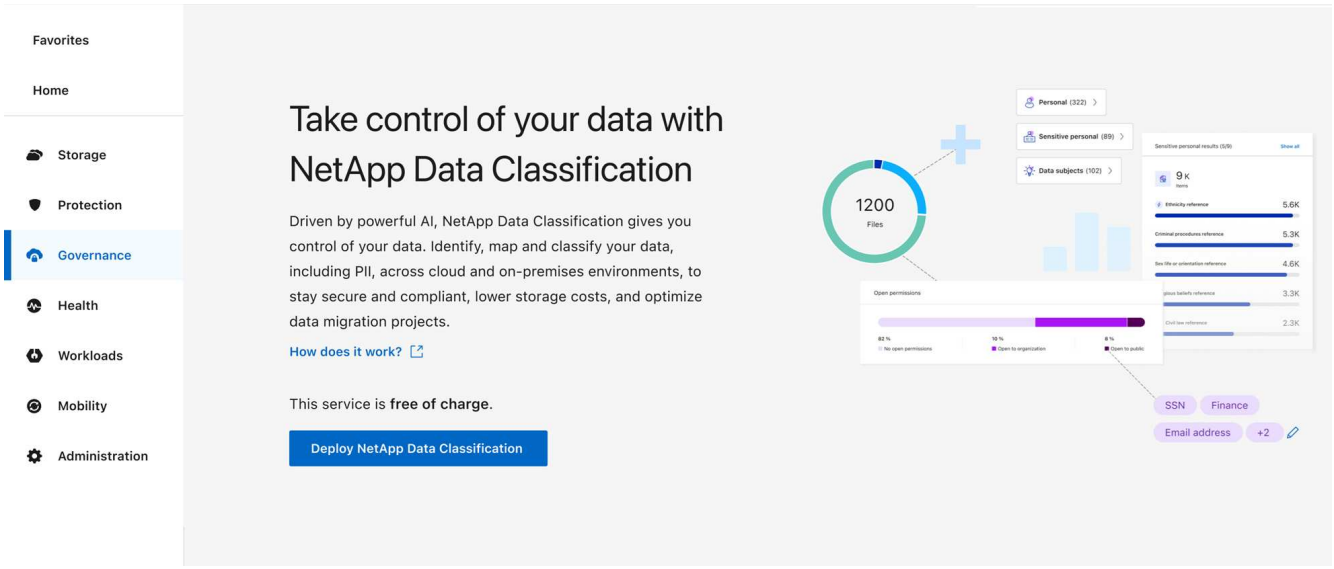
- Verifique que su sistema Linux cumpla con losrequisitos del anfitrión .
- Verifique que el sistema tenga instalados los dos paquetes de software necesarios (Docker Engine o Podman y Python 3).
- Asegúrese de tener privilegios de root en el sistema Linux.
- Si está utilizando un proxy para acceder a Internet:
  - Necesitará la información del servidor proxy (dirección IP o nombre de host, puerto de conexión, esquema de conexión: https o http, nombre de usuario y contraseña).
  - Si el proxy realiza la interceptación de TLS, necesitará saber la ruta en el sistema Linux de clasificación de datos donde se almacenan los certificados CA de TLS.
  - El proxy no debe ser transparente. Actualmente, la clasificación de datos no admite servidores proxy transparentes.
  - El usuario debe ser un usuario local. Los usuarios del dominio no son compatibles.
- Verifique que su entorno fuera de línea cumpla con los requisitospermisos y conectividad .

Pasos

1. Descargue el software de clasificación de datos desde "Sitio de soporte de NetApp" . El archivo que debe seleccionar se llama **DATASENSE-INSTALLER-<versión>.tar.gz**.
2. Copie el archivo de instalación en el host Linux que planea utilizar (usando `scp` o algún otro método).
3. Descomprima el archivo de instalación en la máquina host, por ejemplo:

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

4. En la consola, seleccione **Gobernanza > Clasificación**.
5. Seleccione **Implementar clasificación local o en la nube**.



6. Dependiendo de si está instalando Data Classification en una instancia que preparó en la nube o en una instancia que preparó en sus instalaciones, seleccione la opción **Implementar** adecuada para iniciar la instalación de Data Classification.
7. Se muestra el cuadro de diálogo *Implementar clasificación de datos en las instalaciones*. Copie el comando proporcionado (por ejemplo: `sudo ./install.sh -a 12345 -c 27AG75 -t 2198qq`) y pégalo en un archivo de texto para que puedas usarlo más tarde. Luego seleccione **Cerrar** para cerrar el cuadro de diálogo.
8. En la máquina host, ingrese el comando que copió y luego siga una serie de indicaciones, o puede proporcionar el comando completo incluidos todos los parámetros requeridos como argumentos de la línea de comando.

   Tenga en cuenta que el instalador realiza una verificación previa para asegurarse de que los requisitos del sistema y de la red estén cumplidos para una instalación exitosa. "Mira este vídeo" Para comprender los mensajes previos a la verificación y sus implicaciones.

Introduzca los parámetros según se le solicite:	Introduzca el comando completo:
a. Pegue el comando que copió del paso 7: `sudo ./install.sh -a <account_id> -c <client_id> -t <user_token>` Si está instalando en una instancia de nube (no en sus instalaciones), agregue `--manual-cloud-install <cloud_provider>`. b. Introduzca la dirección IP o el nombre de host de la máquina host de clasificación de datos para que el sistema del agente de la consola pueda acceder a ella. c. Ingrese la dirección IP o el nombre de host de la máquina host del agente de consola para que el sistema de clasificación de datos pueda acceder a ella. d. Introduzca los detalles del proxy cuando se le solicite. Si su agente de consola ya utiliza un proxy, no es necesario ingresar esta información nuevamente aquí ya que la clasificación de datos utilizará automáticamente el proxy utilizado por el agente de consola.	Alternativamente, puede crear todo el comando por adelantado, proporcionando los parámetros de host y proxy necesarios: `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>`

Valores variables:

- *account_id* = ID de cuenta de NetApp
- *client_id* = ID de cliente del agente de consola (agregue el sufijo "clients" al ID de cliente si aún no está allí)
- *user_token* = token de acceso de usuario JWT
- *ds_host* = dirección IP o nombre de host del sistema Linux de clasificación de datos.
- *cm_host* = dirección IP o nombre de host del sistema del agente de consola.
- *cloud_provider* = Al instalar en una instancia de nube, ingrese "AWS", "Azure" o "Gcp" según el proveedor de nube.
- *proxy_host* = IP o nombre de host del servidor proxy si el host está detrás de un servidor proxy.
- *proxy_port* = Puerto para conectarse al servidor proxy (predeterminado 80).
- *proxy_scheme* = Esquema de conexión: https o http (predeterminado http).
- *proxy_user* = Usuario autenticado para conectarse al servidor proxy, si se requiere autenticación básica. El usuario debe ser un usuario local (no se admiten usuarios de dominio).
- *proxy_password* = Contraseña para el nombre de usuario que usted especificó.
- *ca_cert_dir* = Ruta en el sistema Linux de clasificación de datos que contiene paquetes de certificados CA TLS adicionales. Solo es necesario si el proxy está realizando intercepción TLS.

Resultado

El instalador de Data Classification instala paquetes, registra la instalación e instala Data Classification. La instalación puede tardar entre 10 y 20 minutos.

Si hay conectividad a través del puerto 8080 entre la máquina host y la instancia del agente de la consola, verá el progreso de la instalación en la pestaña Clasificación de datos en la consola.

¿Qué sigue?

Desde la página de Configuración puede seleccionar las fuentes de datos que desea escanear.

Instalar NetApp Data Classification en un host Linux sin acceso a Internet

La instalación de NetApp Data Classification en un host Linux en un sitio local que no tiene acceso a Internet se conoce como *modo privado*. Este tipo de instalación, que utiliza un script de instalación, no tiene conectividad con la capa SaaS de la NetApp Console .

> El modo privado de BlueXP (interfaz BlueXP heredada) generalmente se usa con entornos locales que no tienen conexión a Internet y con regiones de nube seguras, que incluyen AWS Secret Cloud, AWS Top Secret Cloud y Azure IL6. NetApp continúa brindando soporte a estos entornos con la interfaz BlueXP heredada. Para obtener documentación del modo privado en la interfaz heredada de BlueXP , consulte"Documentación en PDF para el modo privado de BlueXP" .

Compruebe que su host Linux esté listo para instalar NetApp Data Classification

Antes de instalar NetApp Data Classification manualmente en un host Linux, opcionalmente ejecute un script en el host para verificar que todos los requisitos previos estén cumplidos para instalar Data Classification. Puede ejecutar este script en un host Linux en su red o en un host Linux en la nube. El host puede estar conectado a Internet o puede residir en un sitio que no tenga acceso a Internet (un *sitio oscuro*).

El script de instalación de Clasificación de datos incluye un script de prueba para garantizar que su entorno cumpla con los requisitos. Puede ejecutar este script por separado para verificar la preparación del host Linux antes de ejecutar el script de instalación.

Empezando

Realizarás las siguientes tareas:

- Opcionalmente, instale un agente de consola si aún no tiene uno instalado. Puede ejecutar el script de prueba sin tener un agente de consola instalado, pero el script verifica la conectividad entre el agente de consola y la máquina host de clasificación de datos, por lo que se recomienda que tenga un agente de consola.
- Prepare la máquina host y verifique que cumpla con todos los requisitos.
- Habilitar el acceso a Internet saliente desde la máquina host de clasificación de datos.
- Verifique que todos los puertos necesarios estén habilitados en todos los sistemas.
- Descargue y ejecute el script de prueba de prerrequisitos.

Crear un agente de consola

Se requiere un agente de consola antes de poder instalar y utilizar la clasificación de datos. Sin embargo, puede ejecutar el script de Requisitos previos sin un agente de consola.

Puede "Instalar el agente de consola local" en un host Linux en su red o en un host Linux en la nube. También puede instalar Clasificación de datos localmente si el agente de Consola está instalado localmente.

Para crear un agente de consola en su entorno de proveedor de nube, consulte:

- "Creación de un agente de consola en AWS"
- "Creación de un agente de consola en Azure"
- "Creación de un agente de consola en GCP"

Necesita la dirección IP o el nombre de host del sistema del agente de la consola al ejecutar el script de requisitos previos. Tienes esta información si instalaste el agente de consola en tus instalaciones. Si el agente de la Consola está implementado en la nube, puede encontrar esta información desde la Consola: seleccione el ícono Ayuda y luego **Soporte**; en la sección Agente y Auditoría, seleccione **Ir al agente**.

Verificar los requisitos del host

El software de clasificación de datos debe ejecutarse en un host que cumpla con los requisitos específicos del sistema operativo, requisitos de RAM y requisitos de software.

- La clasificación de datos debe realizarse en un host dedicado. El host no se puede compartir con otras aplicaciones o software de terceros, como antivirus.
- Elija el tamaño que se alinee con el conjunto de datos que planea escanear con Clasificación de datos.

Tamaño del sistema	UPC	RAM (la memoria de intercambio debe estar deshabilitada)	Disco
Extra grande	32 CPU	128 GB de RAM	• SSD de 1 TiB en /, o 100 GiB disponibles en /opt • 895 GiB disponibles en /var/lib/docker • 5 GiB en /tmp • **Para Podman, 30 GB en /var/tmp**
Grande	16 CPU	64 GB de RAM	• SSD de 500 GiB en /, o 100 GiB disponibles en /opt • 400 GiB disponibles en /var/lib/docker o para Podman /var/lib/containers • 5 GiB en /tmp • **Para Podman, 30 GB en /var/tmp**

- Al implementar una instancia de cómputo en la nube para su instalación de Clasificación de datos, se recomienda utilizar un sistema que cumpla con los requisitos del sistema "Grande" mencionados anteriormente:
  - **Tipo de instancia de Amazon Elastic Compute Cloud (Amazon EC2)**: "m6i.4xlarge". "Ver tipos de instancias de AWS adicionales" .

- **Tamaño de máquina virtual de Azure**: "Standard_D16s_v3". "Ver tipos de instancias de Azure adicionales" .
- **Tipo de máquina GCP**: "n2-standard-16". "Ver tipos de instancias de GCP adicionales" .

- **Permisos de carpeta UNIX**: Se requieren los siguientes permisos mínimos de UNIX:

Carpeta	Permisos mínimos
/tmp	`rwxrwxrwt`
/optar	`rwxr-xr-x`
/var/lib/docker	`rwx------`
/usr/lib/systemd/sistema	`rwxr-xr-x`

- **Sistema operativo**:
  - Los siguientes sistemas operativos requieren el uso del motor de contenedores Docker:
    - Red Hat Enterprise Linux versión 7.8 y 7.9
    - Ubuntu 22.04 (requiere la versión 1.23 o superior de Data Classification)
    - Ubuntu 24.04 (requiere la versión 1.23 o superior de Data Classification)
  - Los siguientes sistemas operativos requieren el uso del motor de contenedores Podman y requieren la versión 1.30 o superior de Data Classification:
    - Red Hat Enterprise Linux versiones 8.8, 8.10, 9.0, 9.1, 9.2, 9.3, 9.4, 9.5 y 9.6.
  - Las extensiones vectoriales avanzadas (AVX2) deben estar habilitadas en el sistema host.
- **Gestión de suscripciones de Red Hat**: el host debe estar registrado en Gestión de suscripciones de Red Hat. Si no está registrado, el sistema no puede acceder a los repositorios para actualizar el software de terceros requerido durante la instalación.
- **Software adicional**: Debe instalar el siguiente software en el host antes de instalar Data Classification:
  - Dependiendo del sistema operativo que estés usando, necesitas instalar uno de los motores de contenedores:
    - Docker Engine versión 19.3.1 o superior. "Ver instrucciones de instalación" .
    - Podman versión 4 o superior. Para instalar Podman, ingrese(`sudo yum install podman netavark -y`).
- Versión de Python 3.6 o superior. "Ver instrucciones de instalación" .
  - **Consideraciones sobre NTP**: NetApp recomienda configurar el sistema de clasificación de datos para utilizar un servicio de Protocolo de tiempo de red (NTP). La hora debe estar sincronizada entre el sistema de clasificación de datos y el sistema del agente de consola.
- **Consideraciones sobre Firewalld**: Si planea utilizar `firewalld` Le recomendamos que lo habilite antes de instalar Data Classification. Ejecute los siguientes comandos para configurar `firewalld` para que sea compatible con la Clasificación de Datos:

```
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 planea utilizar hosts de clasificación de datos adicionales como nodos de escáner (en un modelo distribuido), agregue estas reglas a su sistema principal en este momento:

```
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
```

Tenga en cuenta que debe reiniciar Docker o Podman cada vez que habilite o actualice `firewalld` ajustes.

Habilitar el acceso a Internet saliente desde la Clasificación de datos

La clasificación de datos requiere acceso a Internet saliente. Si su red virtual o física utiliza un servidor proxy para el acceso a Internet, asegúrese de que la instancia de clasificación de datos tenga acceso a Internet saliente para contactar los siguientes puntos finales.

Esta sección no es necesaria para los sistemas host instalados en sitios sin conectividad a Internet.

Puntos finales	Objetivo
\ https://api.console.netapp.com	Comunicación con el servicio de consola, que incluye cuentas de NetApp .
\ https://netapp-cloud-account.auth0.com \ https://auth0.com	Comunicación con el sitio web de la consola para la autenticación centralizada de usuarios.
\ 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/	Proporciona acceso a imágenes de software, manifiestos, plantillas y permite enviar registros y métricas.
\ https://support.compliance.api.console.netapp.com/	Permite a NetApp transmitir datos desde registros de auditoría.
\ https://github.com/docker \ https://download.docker.com	Proporciona paquetes de requisitos previos para la instalación de Docker.

Puntos finales	Objetivo
\ http://packages.ubuntu.com/ \ http://archive.ubuntu.com	Proporciona paquetes de requisitos previos para la instalación de Ubuntu.

Verifique que todos los puertos requeridos estén habilitados

Debe asegurarse de que todos los puertos necesarios estén abiertos para la comunicación entre el agente de la consola, la clasificación de datos, Active Directory y sus fuentes de datos.

Tipo de conexión	Puertos	Descripción
Agente de consola <> Clasificación de datos	8080 (TCP), 443 (TCP) y 80. 9000	Las reglas de firewall o enrutamiento para el agente de la consola deben permitir el tráfico entrante y saliente a través del puerto 443 hacia y desde la instancia de clasificación de datos. Asegúrese de que el puerto 8080 esté abierto para que pueda ver el progreso de la instalación en la consola. Si se utiliza un firewall en el host Linux, se requiere el puerto 9000 para los procesos internos dentro de un servidor Ubuntu.
Agente de consola <> clúster ONTAP (NAS)	443 (TCP)	La consola descubre clústeres ONTAP mediante HTTPS. Si utiliza políticas de firewall personalizadas, el host del agente de la consola debe permitir el acceso HTTPS saliente a través del puerto 443. Si el agente de la consola está en la nube, toda comunicación saliente está permitida por el firewall predefinido o las reglas de enrutamiento.

Ejecute el script de requisitos previos de clasificación de datos

Siga estos pasos para ejecutar el script de requisitos previos de clasificación de datos.

"Mira este vídeo"para ver cómo ejecutar el script de requisitos previos e interpretar los resultados.

Antes de empezar

- Verifique que su sistema Linux cumpla con losrequisitos del anfitrión .
- Verifique que el sistema tenga instalados los dos paquetes de software necesarios (Docker Engine o Podman y Python 3).
- Asegúrese de tener privilegios de root en el sistema Linux.

Pasos

1. Descargue el script de Requisitos previos de clasificación de datos desde "Sitio de soporte de NetApp" . El archivo que debe seleccionar se llama **standalone-pre-requisite-tester-<version>**.
2. Copie el archivo al host Linux que planea utilizar (usando `scp` o algún otro método).
3. Asignar permisos para ejecutar el script.

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

4. Ejecute el script utilizando el siguiente comando.

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

   Agregue la opción "--darksite" solo si está ejecutando el script en un host que no tiene acceso a Internet. Se omiten ciertas pruebas de requisitos previos cuando el host no está conectado a Internet.

5. El script le solicita la dirección IP de la máquina host de clasificación de datos.
   - Introduzca la dirección IP o el nombre de host.
6. El script le preguntará si tiene un agente de consola instalado.
   - Ingrese **N** si no tiene un agente de consola instalado.
   - Ingrese **Y** si tiene un agente de consola instalado. Y luego ingrese la dirección IP o el nombre de host del agente de la consola para que el script de prueba pueda probar esta conectividad.
7. El script ejecuta una variedad de pruebas en el sistema y muestra resultados a medida que avanza. Cuando termina, escribe un registro de la sesión en un archivo llamado `prerequisites-test-<timestamp>.log` en el directorio `/opt/netapp/install_logs` .

Resultado

Si todas las pruebas de requisitos previos se ejecutaron correctamente, puede instalar Data Classification en el host cuando esté listo.

Si se descubre algún problema, se clasifica como "Recomendado" o "Obligatorio" para su solución. Los problemas recomendados suelen ser elementos que harían que las tareas de categorización y escaneo de clasificación de datos se ejecuten más lentamente. No es necesario corregir estos elementos, pero es posible que quieras abordarlos.

Si tiene algún problema "Obligatorio", debe solucionarlo y ejecutar nuevamente el script de prueba de requisitos previos.

Activar el escaneo en sus fuentes de datos

Escanee fuentes de datos con NetApp Data Classification

NetApp Data Classification escanea los datos en los repositorios (los volúmenes, esquemas de bases de datos u otros datos de usuario) que seleccione para identificar datos personales y confidenciales. Luego, la clasificación de datos mapea los datos de su organización, categoriza cada archivo e identifica patrones predefinidos en los datos. El resultado del escaneo es un índice de información personal, información personal confidencial, categorías de datos y tipos de archivos.

Después del escaneo inicial, la clasificación de datos escanea continuamente sus datos de manera rotatoria para detectar cambios incrementales. Por eso es importante mantener la instancia en ejecución.

Puede habilitar y deshabilitar escaneos a nivel de volumen o a nivel de esquema de base de datos.

¿Cuál es la diferencia entre los escaneos de mapeo y clasificación?

Puede realizar dos tipos de escaneos en Clasificación de datos:

- Los escaneos de solo mapeo brindan únicamente una descripción general de alto nivel de sus datos y se realizan en fuentes de datos seleccionadas. Los escaneos de solo mapeo toman menos tiempo que los escaneos de mapas y clasificación porque no acceden a los archivos para ver los datos dentro. Es posible que desees hacer esto inicialmente para identificar áreas de investigación y luego realizar un escaneo de Mapa y Clasificación en esas áreas.
- **Los escaneos de mapas y clasificación** proporcionan un escaneo de nivel profundo de sus datos.

La siguiente tabla muestra algunas de las diferencias:

Característica	Mapear y clasificar escaneos	Escaneos de solo mapeo
Velocidad de escaneo	Lento	Rápido
Precios	Gratis	Gratis
Capacidad	Limitado a 500 TiB*	Limitado a 500 TiB*
Lista de tipos de archivos y capacidad utilizada	Sí	Sí
Número de archivos y capacidad utilizada	Sí	Sí
Edad y tamaño de los archivos	Sí	Sí
Capacidad para ejecutar un"Informe de mapeo de datos"	Sí	Sí
Página de investigación de datos para ver los detalles del archivo	Sí	No
Buscar nombres dentro de los archivos	Sí	No
Crear"consultas guardadas" que proporcionan resultados de búsqueda personalizados	Sí	No
Capacidad de ejecutar otros informes	Sí	No
Capacidad de ver metadatos de archivos**	No	Sí

* La clasificación de datos no impone un límite en la cantidad de datos que puede escanear. Cada agente de consola admite el escaneo y la visualización de 500 TiB de datos. Para escanear más de 500 TiB de datos,"instalar otro agente de consola" entonces"Implementar otra instancia de clasificación de datos" . + La interfaz de usuario de la consola muestra datos de un solo conector. Para obtener sugerencias sobre cómo ver datos de varios agentes de la consola, consulte"Trabajar con múltiples agentes de consola" .

** Los siguientes metadatos se extraen de los archivos durante los escaneos de mapeo:

- Sistema
- Tipo de sistema
- Repositorio de almacenamiento
- Tipo de archivo
- Capacidad utilizada
- Número de archivos

- Tamaño del archivo
- Creación de archivos
- Último acceso al archivo
- Archivo modificado por última vez
- Hora de descubrimiento del archivo
- Extracción de permisos

Diferencias en el panel de gobernanza:

Característica	Mapa y clasificación	Mapa
Datos obsoletos	Sí	Sí
Datos no comerciales	Sí	Sí
Archivos duplicados	Sí	Sí
Consultas guardadas predefinidas	Sí	No
Consultas guardadas predeterminadas	Sí	Sí
Informe de la DDA	Sí	Sí
Informe de mapeo	Sí	Sí
Detección del nivel de sensibilidad	Sí	No
Datos sensibles con amplios permisos	Sí	No
Permisos abiertos	Sí	Sí
La era de los datos	Sí	Sí
Tamaño de los datos	Sí	Sí
Categorías	Sí	No
Tipos de archivos	Sí	Sí

Diferencias en el panel de cumplimiento:

Característica	Mapa y clasificación	Mapa
Información personal	Sí	No
Información personal sensible	Sí	No
Informe de evaluación de riesgos de privacidad	Sí	No
Informe HIPAA	Sí	No
Informe PCI DSS	Sí	No

Diferencias en los filtros de investigación:

Característica	Mapa y clasificación	Mapa
Consultas guardadas	Sí	Sí
Tipo de sistema	Sí	Sí
Sistema	Sí	Sí
Repositorio de almacenamiento	Sí	Sí
Tipo de archivo	Sí	Sí
Tamaño del archivo	Sí	Sí
Hora de creación	Sí	Sí
Tiempo descubierto	Sí	Sí
Última modificación	Sí	Sí
Último acceso	Sí	Sí
Permisos abiertos	Sí	Sí
Ruta del directorio de archivos	Sí	Sí
Categoría	Sí	No
Nivel de sensibilidad	Sí	No
Número de identificadores	Sí	No
Datos personales	Sí	No
Datos personales sensibles	Sí	No
Titular de los datos	Sí	No
Duplicados	Sí	Sí
Estado de clasificación	Sí	El estado siempre es "Perspectivas limitadas"
Evento de análisis de escaneo	Sí	Sí
Hash de archivo	Sí	Sí
Número de usuarios con acceso	Sí	Sí
Permisos de usuario/grupo	Sí	Sí
Propietario del archivo	Sí	Sí
Tipo de directorio	Sí	Sí

Escanee Amazon FSx en busca de volúmenes ONTAP con la NetApp Data Classification

Complete unos pocos pasos para escanear Amazon FSx en busca de volúmenes ONTAP con NetApp Data Classification.

Antes de empezar

- Necesita un agente de consola activo en AWS para implementar y administrar la clasificación de datos.
- El grupo de seguridad que seleccionó al crear el sistema debe permitir el tráfico desde la instancia de Clasificación de datos. Puede encontrar el grupo de seguridad asociado utilizando el ENI conectado al sistema de archivos FSx para ONTAP y editarlo utilizando la Consola de administración de AWS.

  "Grupos de seguridad de AWS para instancias de Linux"

  "Grupos de seguridad de AWS para instancias de Windows"

  "Interfaces de red elásticas (ENI) de AWS"

- Asegúrese de que los siguientes puertos estén abiertos para la instancia de clasificación de datos:
  - Para NFS: puertos 111 y 2049.
  - Para CIFS: puertos 139 y 445.

Implementar la instancia de clasificación de datos

"Implementar la clasificación de datos"si aún no hay una instancia implementada.

Debe implementar la clasificación de datos en la misma red de AWS que el agente de consola para AWS y los volúmenes FSx que desea escanear.

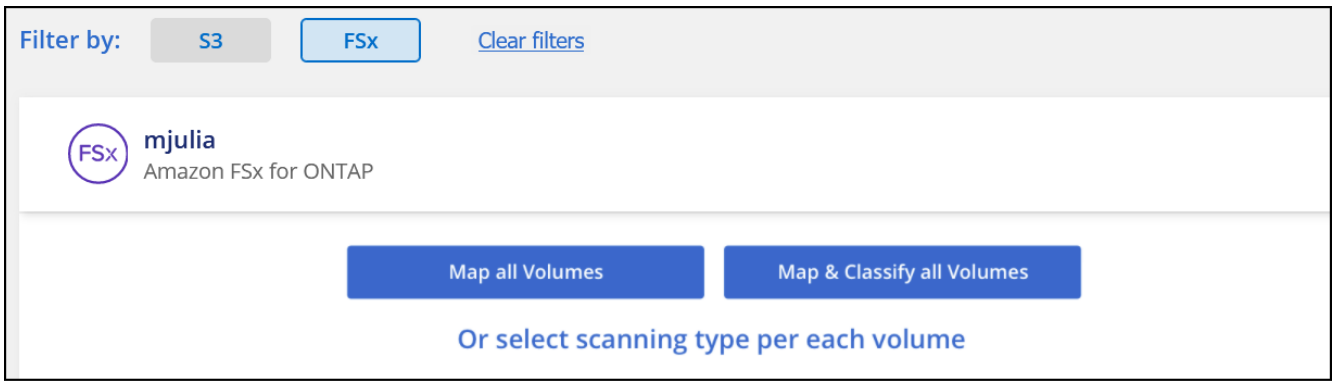
Nota: Actualmente, no se admite la implementación de la clasificación de datos en una ubicación local al escanear volúmenes FSx.

Las actualizaciones del software de clasificación de datos se automatizan siempre que la instancia tenga conectividad a Internet.

Habilite la clasificación de datos en sus sistemas

Puede habilitar la clasificación de datos para FSx para volúmenes ONTAP .

1. Desde la NetApp Console, **Gobernanza > Clasificación**.
2. Desde el menú Clasificación de datos, seleccione **Configuración**.



3. Seleccione cómo desea escanear los volúmenes en cada sistema. "Aprenda sobre escaneos de mapeo y clasificación":
   - Para mapear todos los volúmenes, seleccione **Mape todos los volúmenes**.

- Para mapear y clasificar todos los volúmenes, seleccione **Mapear y clasificar todos los volúmenes**.
- Para personalizar el escaneo para cada volumen, seleccione **O seleccione el tipo de escaneo para cada volumen** y luego elija los volúmenes que desea mapear y/o clasificar.

4. En el cuadro de diálogo de confirmación, seleccione **Aprobar** para que la Clasificación de datos comience a escanear sus volúmenes.

Resultado

La clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados estarán disponibles en el panel de Cumplimiento tan pronto como la Clasificación de Datos finalice los escaneos iniciales. El tiempo que lleva depende de la cantidad de datos: pueden ser unos minutos u horas. Puede seguir el progreso del escaneo inicial navegando al menú **Configuración** y luego seleccionando **Configuración del sistema**. Realice un seguimiento del progreso de cada escaneo en la barra de progreso; puede pasar el cursor sobre la barra de progreso para ver la cantidad de archivos escaneados en relación con el total de archivos en el volumen.

- De forma predeterminada, si la clasificación de datos no tiene permisos de atributos de escritura en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos en sus volúmenes porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, seleccione **O seleccione el tipo de escaneo para cada volumen**. La página resultante tiene una configuración que puede habilitar para que la Clasificación de datos escanee los volúmenes independientemente de los permisos.
- La clasificación de datos escanea solo un recurso compartido de archivos bajo un volumen. Si tiene varias acciones en sus volúmenes, deberá escanear esas otras acciones por separado como un grupo de acciones. "Ver más detalles sobre esta limitación de clasificación de datos" .

Verificar que la clasificación de datos tenga acceso a los volúmenes

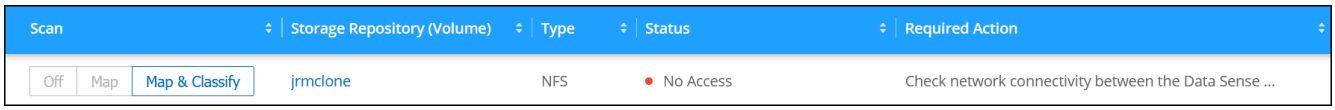
Asegúrese de que la clasificación de datos pueda acceder a los volúmenes verificando su red, grupos de seguridad y políticas de exportación.

Necesitará proporcionar a Data Classification las credenciales CIFS para que pueda acceder a los volúmenes CIFS.

Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. En la página de Configuración, seleccione **Ver detalles** para revisar el estado y corregir cualquier error.

   Por ejemplo, la siguiente imagen muestra un volumen que la clasificación de datos no puede escanear debido a problemas de conectividad de red entre la instancia de clasificación de datos y el volumen.



3. Asegúrese de que haya una conexión de red entre la instancia de clasificación de datos y cada red que incluya volúmenes para FSx para ONTAP.

   Para FSx para ONTAP, la clasificación de datos puede escanear volúmenes solo en la misma región que la consola.

4. Asegúrese de que las políticas de exportación de volumen NFS incluyan la dirección IP de la instancia de clasificación de datos para que pueda acceder a los datos de cada volumen.
5. Si utiliza CIFS, proporcione a Clasificación de datos credenciales de Active Directory para que pueda escanear volúmenes CIFS.

    a. Desde el menú Clasificación de datos, seleccione **Configuración**.

    b. Para cada sistema, seleccione **Editar credenciales CIFS** e ingrese el nombre de usuario y la contraseña que la clasificación de datos necesita para acceder a los volúmenes CIFS en el sistema.

        Las credenciales pueden ser de solo lectura, pero proporcionar credenciales de administrador garantiza que Data Classification pueda leer cualquier dato que requiera permisos elevados. Las credenciales se almacenan en la instancia de clasificación de datos.

        Si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de datos, se recomienda que el usuario tenga permisos de atributos de escritura en CIFS o permisos de escritura en NFS. Si es posible, configure el usuario de Active Directory como parte de un grupo principal en la organización que tenga permisos para todos los archivos.

    Después de ingresar las credenciales, debería ver un mensaje que indica que todos los volúmenes CIFS se autenticaron correctamente.

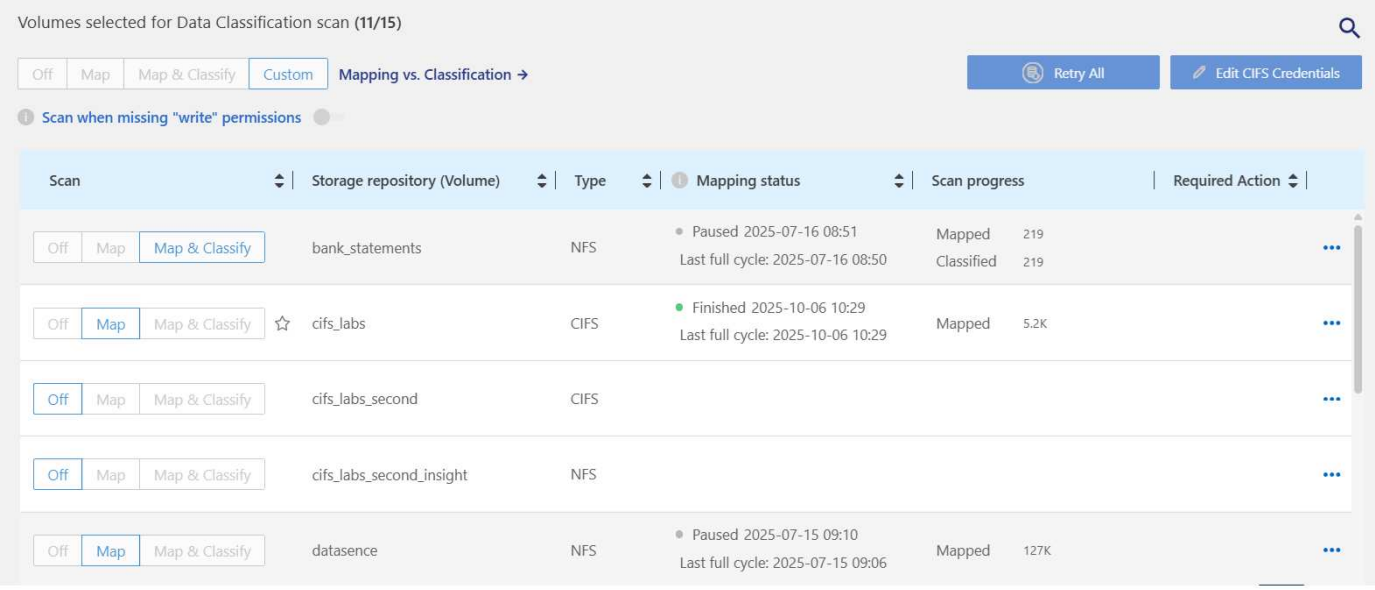
Habilitar y deshabilitar escaneos en volúmenes

Puede iniciar o detener análisis en cualquier sistema en cualquier momento desde la página de Configuración. También puede cambiar los escaneos de solo mapeo a escaneos de mapeo y clasificación, y viceversa. Se recomienda escanear todos los volúmenes de un sistema.

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha seleccionado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando lo configure en **Personalizado** o **Desactivado** en el área de encabezado, deberá activar el mapeo y/o el escaneo completo en cada nuevo volumen que agregue al sistema.

El interruptor en la parte superior de la página para **Escanear cuando faltan permisos de "escritura"** está deshabilitado de manera predeterminada. Esto significa que si la clasificación de datos no tiene permisos de escritura de atributos en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, active el interruptor y se escanearán todos los archivos independientemente de los permisos. "Más información".

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha configurado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando la configuración para todos los volúmenes es **Personalizada** o **Desactivada**, deberá activar el escaneo manualmente para cada nuevo volumen que agregue.



Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione un sistema y luego seleccione **Configuración**.
3. Para habilitar o deshabilitar los escaneos para todos los volúmenes, seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** en el encabezado sobre todos los volúmenes.

   Para habilitar o deshabilitar los escaneos de volúmenes individuales, busque los volúmenes en la lista y luego seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** junto al nombre del volumen.

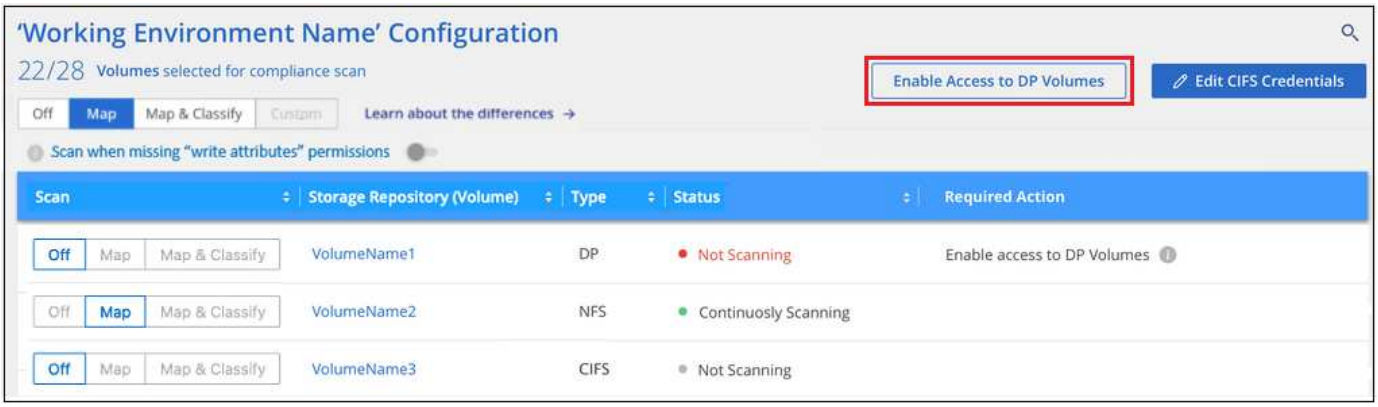
Resultado

Cuando habilita el escaneo, la clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados comienzan a aparecer en el panel de Cumplimiento tan pronto como la Clasificación de datos inicia el escaneo. El tiempo de finalización del escaneo depende de la cantidad de datos y puede variar entre minutos y horas.

Escanear volúmenes de protección de datos

De forma predeterminada, los volúmenes de protección de datos (DP) no se escanean porque no están expuestos externamente y la clasificación de datos no puede acceder a ellos. Estos son los volúmenes de destino para las operaciones de SnapMirror desde un sistema de archivos FSx para ONTAP .

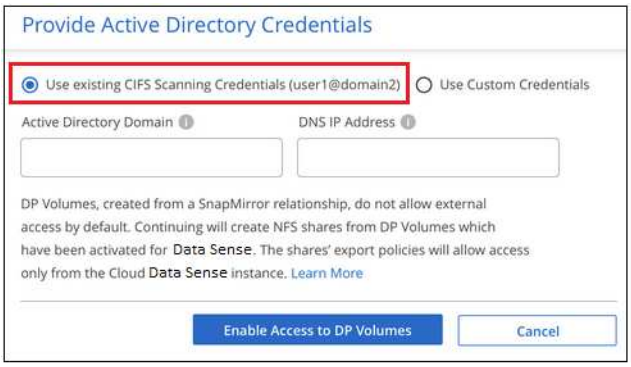
Inicialmente, la lista de volúmenes identifica estos volúmenes como *Tipo* **DP** con el *Estado* **No escaneando** y la *Acción requerida* **Habilitar acceso a volúmenes DP**.

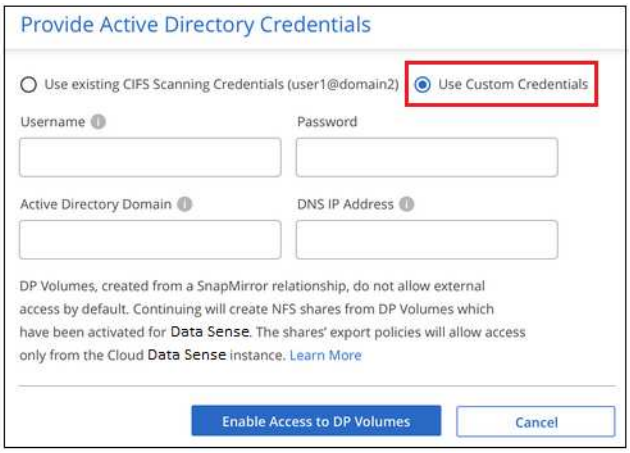


Pasos

Si desea escanear estos volúmenes de protección de datos:

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione **Habilitar acceso a volúmenes DP** en la parte superior de la página.
3. Revise el mensaje de confirmación y seleccione **Habilitar acceso a volúmenes DP** nuevamente.
   - Se habilitan los volúmenes que se crearon inicialmente como volúmenes NFS en el sistema de archivos de origen FSx para ONTAP .
   - Los volúmenes que se crearon inicialmente como volúmenes CIFS en el sistema de archivos de origen FSx para ONTAP requieren que ingrese credenciales CIFS para escanear esos volúmenes DP. Si ya ingresó las credenciales de Active Directory para que la Clasificación de datos pueda escanear volúmenes CIFS, puede usar esas credenciales o puede especificar un conjunto diferente de credenciales de administrador.





4. Active cada volumen DP que desee escanear.

Resultado

Una vez habilitada, la clasificación de datos crea un recurso compartido NFS de cada volumen DP que se activó para el escaneo. Las políticas de exportación de acciones solo permiten el acceso desde la instancia de Clasificación de datos.

Si no tenía volúmenes de protección de datos CIFS cuando habilitó inicialmente el acceso a los volúmenes DP y luego agregó algunos, el botón **Habilitar acceso a CIFS DP** aparece en la parte superior de la página de Configuración. Seleccione este botón y agregue credenciales CIFS para habilitar el acceso a estos volúmenes DP CIFS.

Las credenciales de Active Directory se registran solo en la VM de almacenamiento del primer volumen DP CIFS, por lo que se escanearán todos los volúmenes DP en esa SVM. Cualquier volumen que resida en otras SVM no tendrá las credenciales de Active Directory registradas, por lo que esos volúmenes de DP no se escanearán.

Escanee volúmenes de Azure NetApp Files con NetApp Data Classification

Complete unos pocos pasos para comenzar a utilizar NetApp Data Classification para Azure NetApp Files.

Descubra el sistema de Azure NetApp Files que desea escanear

Si el sistema de Azure NetApp Files que desea escanear aún no se encuentra en la NetApp Console como sistema,"agréguelo en la página de Sistemas" .

Implementar la instancia de clasificación de datos

"Implementar la clasificación de datos"si aún no hay una instancia implementada.

La clasificación de datos debe implementarse en la nube al escanear volúmenes de Azure NetApp Files y debe implementarse en la misma región que los volúmenes que desea escanear.

Nota: Actualmente, no se admite la implementación de la clasificación de datos en una ubicación local al escanear volúmenes de Azure NetApp Files .

Habilite la clasificación de datos en sus sistemas

Puede habilitar la clasificación de datos en sus volúmenes de Azure NetApp Files .

1. Desde el menú Clasificación de datos, seleccione **Configuración**.

2. Seleccione cómo desea escanear los volúmenes en cada sistema. "Aprenda sobre escaneos de mapeo y clasificación":
   - Para mapear todos los volúmenes, seleccione **Mape todos los volúmenes**.
   - Para mapear y clasificar todos los volúmenes, seleccione **Mapear y clasificar todos los volúmenes**.
   - Para personalizar el escaneo de cada volumen, seleccione **O seleccione el tipo de escaneo para cada volumen** y luego elija los volúmenes que desea mapear o mapear y clasificar.

     VerHabilitar o deshabilitar escaneos en volúmenes Para más detalles.

3. En el cuadro de diálogo de confirmación, seleccione **Aprobar**.

Resultado

La clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados están disponibles en el panel de Cumplimiento tan pronto como la Clasificación de datos finaliza los escaneos iniciales. El tiempo que lleva depende de la cantidad de datos: pueden ser unos minutos u horas. Puede seguir el progreso del escaneo inicial navegando al menú **Configuración** y luego seleccionando **Configuración del sistema**. La clasificación de datos muestra una barra de progreso para cada escaneo. Puede pasar el cursor sobre la barra de progreso para ver la cantidad de archivos escaneados en relación con la cantidad total de archivos en el volumen.

- De forma predeterminada, si la clasificación de datos no tiene permisos de atributos de escritura en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos en sus volúmenes porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, seleccione **O seleccione el tipo de escaneo para cada volumen**. La página resultante tiene una configuración que puede habilitar para que la Clasificación de datos escanee los volúmenes independientemente de los permisos.
- La clasificación de datos escanea solo un recurso compartido de archivos bajo un volumen. Si tiene varias acciones en sus volúmenes, deberá escanear esas otras acciones por separado como un grupo de acciones. "Conozca esta limitación de clasificación de datos" .

Verificar que la clasificación de datos tenga acceso a los volúmenes

Asegúrese de que la clasificación de datos pueda acceder a los volúmenes verificando su red, grupos de seguridad y políticas de exportación. Debe proporcionar a la clasificación de datos credenciales CIFS para que pueda acceder a los volúmenes CIFS.

Para Azure NetApp Files, la clasificación de datos solo puede escanear volúmenes en la misma región que la consola.

Lista de verificación

- Asegúrese de que haya una conexión de red entre la instancia de Clasificación de datos y cada red que incluya volúmenes para Azure NetApp Files.
- Asegúrese de que los siguientes puertos estén abiertos para la instancia de clasificación de datos:
  - Para NFS: puertos 111 y 2049.
  - Para CIFS: puertos 139 y 445.
- Asegúrese de que las políticas de exportación de volumen NFS incluyan la dirección IP de la instancia de clasificación de datos para que pueda acceder a los datos de cada volumen.

Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.

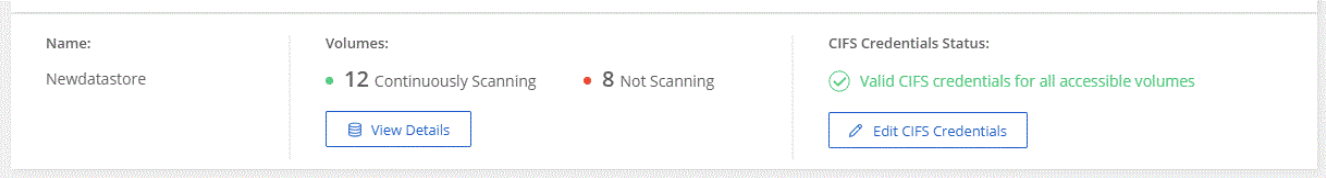
   a. Si utiliza CIFS (SMB), asegúrese de que las credenciales de Active Directory sean correctas. Para cada sistema, seleccione **Editar credenciales CIFS** y luego ingrese el nombre de usuario y la contraseña que la clasificación de datos necesita para acceder a los volúmenes CIFS en el sistema.

      Las credenciales pueden ser de solo lectura; proporcionar credenciales de administrador garantiza que Data Classification pueda leer cualquier dato que requiera permisos elevados. Las credenciales se almacenan en la instancia de clasificación de datos.

      Si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de datos, se recomienda que el usuario tenga permisos de atributos de escritura en CIFS o permisos de escritura en NFS. Si es posible, configure el usuario de

Active Directory como parte de un grupo principal en la organización que tenga permisos para todos los archivos.

Después de ingresar las credenciales, debería ver un mensaje que indica que todos los volúmenes CIFS se autenticaron correctamente.



2. En la página de Configuración, seleccione **Ver detalles** para revisar el estado de cada volumen CIFS y NFS. Si es necesario, corrija cualquier error como problemas de conectividad de red.

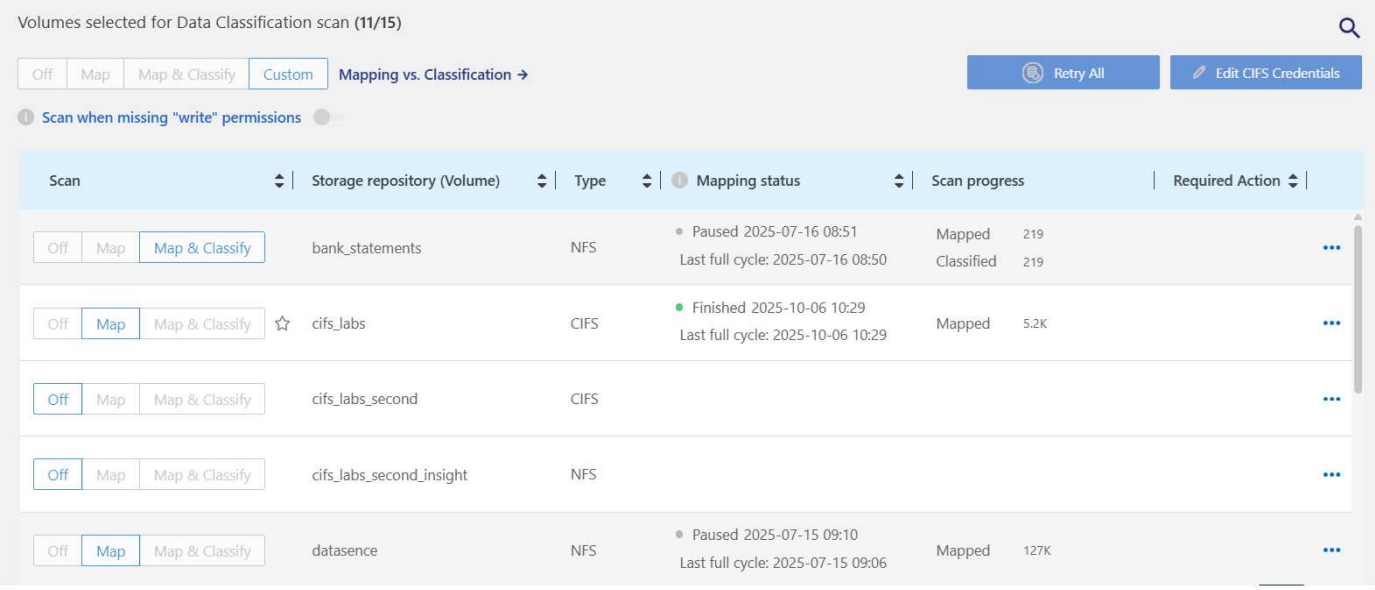
Habilitar o deshabilitar escaneos en volúmenes

Puede iniciar o detener análisis en cualquier sistema en cualquier momento desde la página de Configuración. También puede cambiar los escaneos de solo mapeo a escaneos de mapeo y clasificación, y viceversa. Se recomienda escanear todos los volúmenes de un sistema.

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha seleccionado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando lo configure en **Personalizado** o **Desactivado** en el área de encabezado, deberá activar el mapeo y/o el escaneo completo en cada nuevo volumen que agregue al sistema.

El interruptor en la parte superior de la página para **Escanear cuando faltan permisos de "escritura"** está deshabilitado de manera predeterminada. Esto significa que si la clasificación de datos no tiene permisos de escritura de atributos en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, active el interruptor y se escanearán todos los archivos independientemente de los permisos. "Más información".

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha configurado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando la configuración para todos los volúmenes es **Personalizada** o **Desactivada**, deberá activar el escaneo manualmente para cada nuevo volumen que agregue.



Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione un sistema y luego seleccione **Configuración**.
3. Para habilitar o deshabilitar los escaneos para todos los volúmenes, seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** en el encabezado sobre todos los volúmenes.

   Para habilitar o deshabilitar los escaneos de volúmenes individuales, busque los volúmenes en la lista y luego seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** junto al nombre del volumen.

Resultado

Cuando habilita el escaneo, la clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados comienzan a aparecer en el panel de Cumplimiento tan pronto como la Clasificación de datos inicia el escaneo. El tiempo de finalización del escaneo depende de la cantidad de datos y puede variar entre minutos y horas.

Escanee Cloud Volumes ONTAP y volúmenes ONTAP locales con NetApp Data Classification

Complete unos pocos pasos para comenzar a escanear sus Cloud Volumes ONTAP y volúmenes ONTAP locales utilizando NetApp Data Classification.

Prerrequisitos

Antes de habilitar la Clasificación de datos, asegúrese de tener una configuración compatible.

- Si está escaneando Cloud Volumes ONTAP y sistemas ONTAP locales a los que se puede acceder a través de Internet, puede"Implementar la clasificación de datos en la nube" o"En una ubicación local que tenga acceso a Internet" .
- Si está escaneando sistemas ONTAP locales que se han instalado en un sitio oscuro que no tiene acceso a Internet, debe"Implementar la clasificación de datos en la misma ubicación local que no tiene acceso a Internet" . Esto requiere que el agente de consola se implemente en la misma ubicación local.

Verificar que la clasificación de datos tenga acceso a los volúmenes

Asegúrese de que la clasificación de datos pueda acceder a los volúmenes verificando su red, grupos de seguridad y políticas de exportación. Necesitará proporcionar a Data Classification las credenciales CIFS para que pueda acceder a los volúmenes CIFS.

Lista de verificación

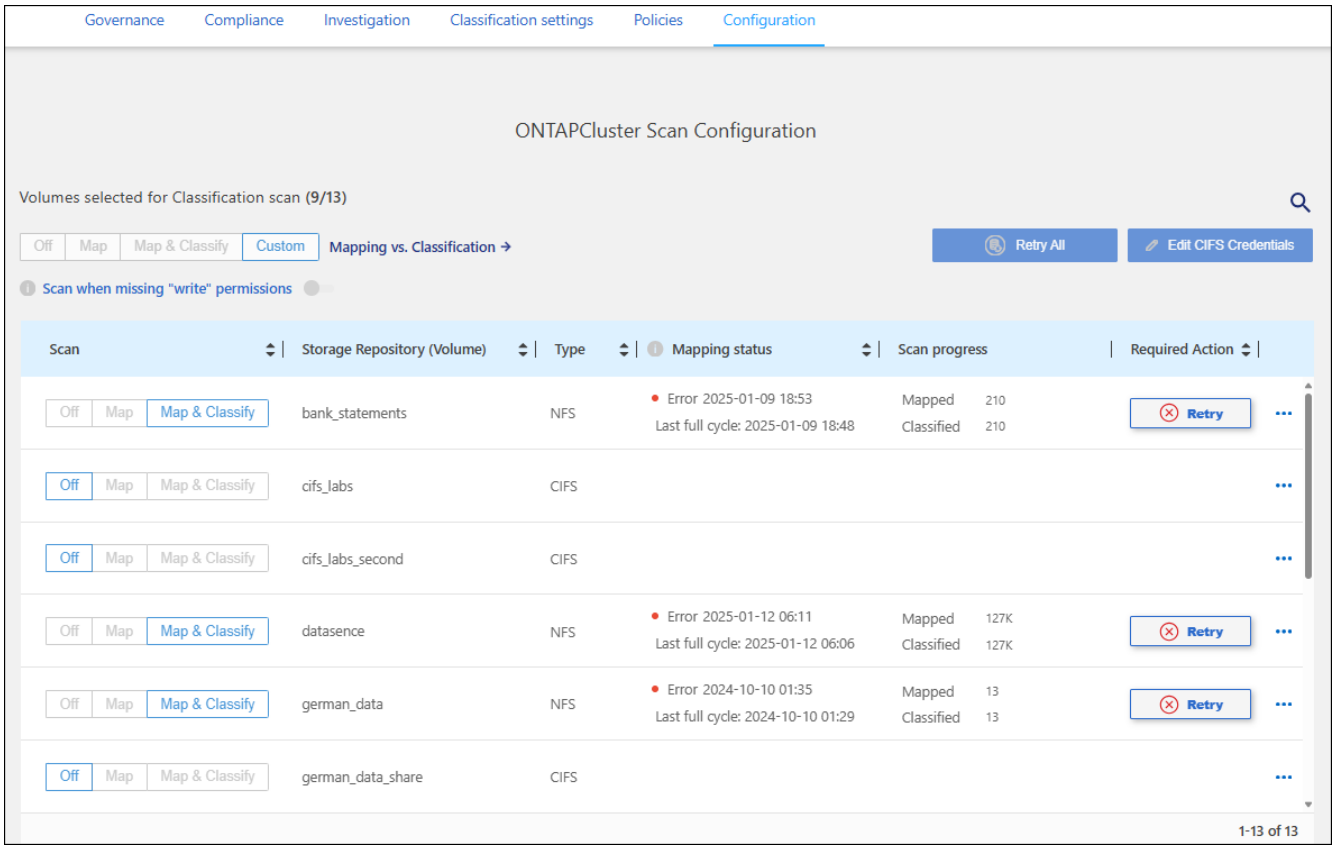
- Asegúrese de que haya una conexión de red entre la instancia de clasificación de datos y cada red que incluya volúmenes para Cloud Volumes ONTAP o clústeres ONTAP locales.
- Asegúrese de que el grupo de seguridad de Cloud Volumes ONTAP permita el tráfico entrante desde la instancia de clasificación de datos.

  Puede abrir el grupo de seguridad para el tráfico desde la dirección IP de la instancia de clasificación de datos o puede abrir el grupo de seguridad para todo el tráfico desde dentro de la red virtual.

- Asegúrese de que las políticas de exportación de volumen NFS incluyan la dirección IP de la instancia de clasificación de datos para que pueda acceder a los datos de cada volumen.

Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.



2. Si utiliza CIFS, proporcione a Clasificación de datos credenciales de Active Directory para que pueda escanear volúmenes CIFS. Para cada sistema, seleccione **Editar credenciales CIFS** e ingrese el nombre de usuario y la contraseña que la clasificación de datos necesita para acceder a los volúmenes CIFS en el sistema.

   Las credenciales pueden ser de solo lectura, pero proporcionar credenciales de administrador garantiza que Data Classification pueda leer cualquier dato que requiera permisos elevados. Las credenciales se almacenan en la instancia de clasificación de datos.

Si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de datos, se recomienda que el usuario tenga permisos de atributos de escritura en CIFS o permisos de escritura en NFS. Si es posible, configure el usuario de Active Directory como parte de un grupo principal en la organización que tenga permisos para todos los archivos.

Si ingresó las credenciales correctamente, un mensaje confirmará que todos los volúmenes CIFS se autenticaron exitosamente.

3. En la página Configuración, seleccione **Configuración** para revisar el estado de cada volumen CIFS y NFS y corregir cualquier error.

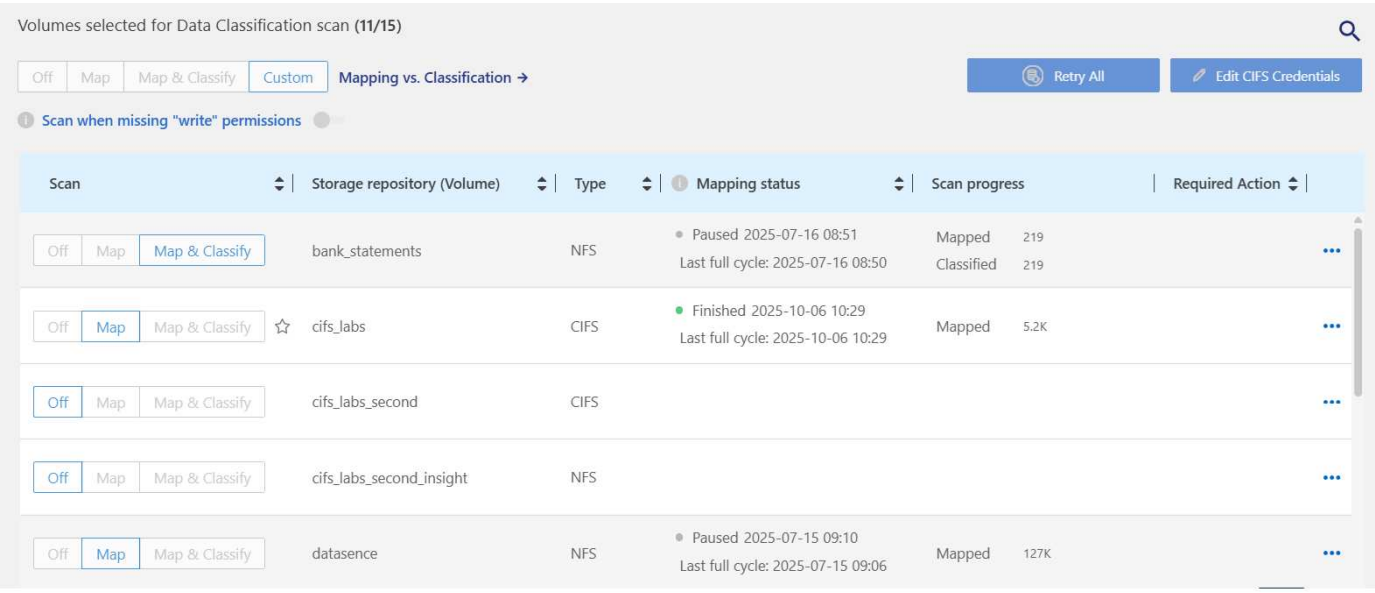
Habilitar o deshabilitar escaneos en volúmenes

Puede iniciar o detener análisis en cualquier sistema en cualquier momento desde la página de Configuración. También puede cambiar los escaneos de solo mapeo a escaneos de mapeo y clasificación, y viceversa. Se recomienda escanear todos los volúmenes de un sistema.

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha seleccionado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando lo configure en **Personalizado** o **Desactivado** en el área de encabezado, deberá activar el mapeo y/o el escaneo completo en cada nuevo volumen que agregue al sistema.

El interruptor en la parte superior de la página para **Escanear cuando faltan permisos de "escritura"** está deshabilitado de manera predeterminada. Esto significa que si la clasificación de datos no tiene permisos de escritura de atributos en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, active el interruptor y se escanearán todos los archivos independientemente de los permisos. "Más información".

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha configurado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando la configuración para todos los volúmenes es **Personalizada** o **Desactivada**, deberá activar el escaneo manualmente para cada nuevo volumen que agregue.



Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione un sistema y luego seleccione **Configuración**.
3. Para habilitar o deshabilitar los escaneos para todos los volúmenes, seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** en el encabezado sobre todos los volúmenes.

   Para habilitar o deshabilitar los escaneos de volúmenes individuales, busque los volúmenes en la lista y luego seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** junto al nombre del volumen.

Resultado

Cuando habilita el escaneo, la clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados comienzan a aparecer en el panel de Cumplimiento tan pronto como la Clasificación de datos inicia el escaneo. El tiempo de finalización del escaneo depende de la cantidad de datos y puede variar entre minutos y horas.

La clasificación de datos escanea solo un recurso compartido de archivos bajo un volumen. Si tiene varias acciones en sus volúmenes, deberá escanear esas otras acciones por separado como un grupo de acciones. "Ver más detalles sobre esta limitación de clasificación de datos" .

Escanee esquemas de bases de datos con NetApp Data Classification

Complete unos pocos pasos para comenzar a escanear sus esquemas de base de datos con NetApp Data Classification.

Revisar los prerrequisitos

Revise los siguientes requisitos previos para asegurarse de tener una configuración compatible antes de habilitar la Clasificación de datos.

Bases de datos compatibles

La clasificación de datos puede escanear esquemas de las siguientes bases de datos:

- Servicio de base de datos relacional de Amazon (Amazon RDS)
- MongoDB
- MySQL
- Oráculo
- PostgreSQL
- SAP HANA
- Servidor SQL (MSSQL)

La función de recopilación de estadísticas **debe estar habilitada** en la base de datos.

Requisitos de la base de datos

Se puede escanear cualquier base de datos con conectividad a la instancia de Clasificación de Datos, independientemente de dónde esté alojada. Solo necesitas la siguiente información para conectarte a la base de datos:

- Dirección IP o nombre de host
- Puerto
- Nombre del servicio (sólo para acceder a bases de datos Oracle)
- Credenciales que permiten acceso de lectura a los esquemas

  Al elegir un nombre de usuario y una contraseña, es importante elegir uno que tenga permisos de lectura completos para todos los esquemas y tablas que desea escanear. Le recomendamos que cree un usuario dedicado para el sistema de clasificación de datos con todos los permisos necesarios.

Para MongoDB, se requiere un rol de administrador de solo lectura.

Implementar la instancia de clasificación de datos

Implementar la clasificación de datos si aún no hay una instancia implementada.

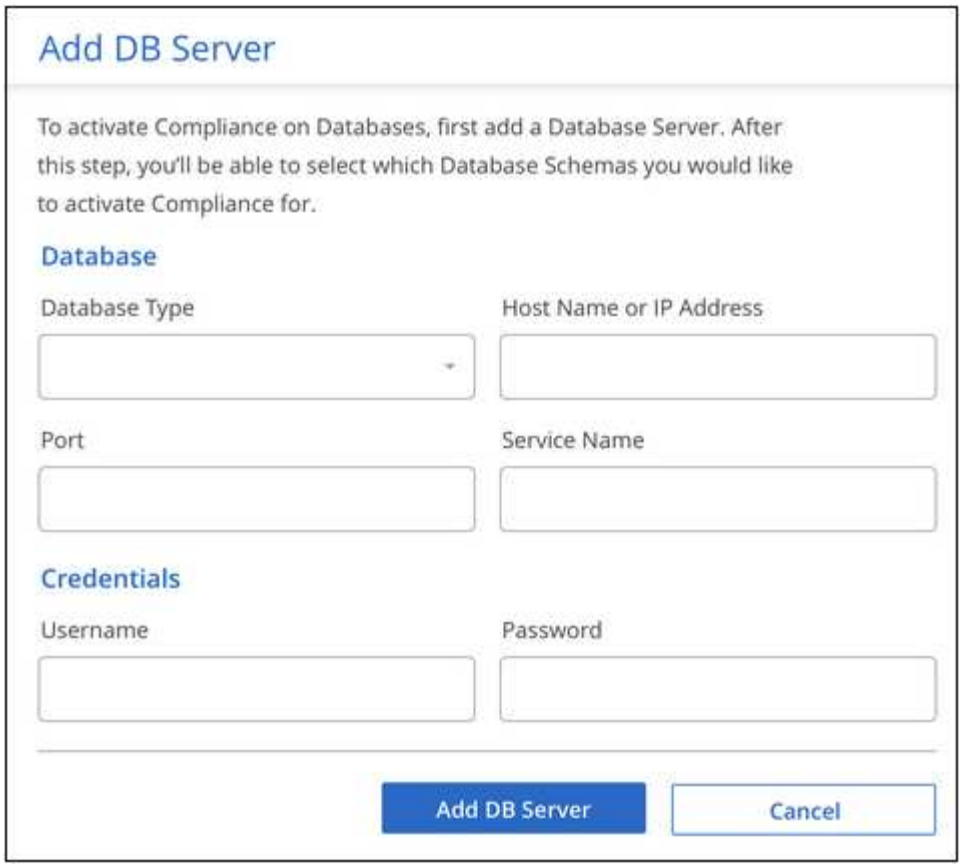
Si está escaneando esquemas de bases de datos a los que se puede acceder a través de Internet, puede"Implementar la clasificación de datos en la nube" o"Implementar la clasificación de datos en una ubicación local que tenga acceso a Internet" .

Si está escaneando esquemas de bases de datos que se han instalado en un sitio oscuro que no tiene acceso a Internet, debe"Implementar la clasificación de datos en la misma ubicación local que no tiene acceso a Internet" . Esto también requiere que el agente de consola esté implementado en esa misma ubicación local.

Agregar el servidor de base de datos

Agregue el servidor de base de datos donde residen los esquemas.

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Desde la página de Configuración, seleccione **Agregar sistema** > **Agregar servidor de base de datos**.
3. Ingrese la información requerida para identificar el servidor de base de datos.
   a. Seleccione el tipo de base de datos.
   b. Introduzca el puerto y el nombre de host o dirección IP para conectarse a la base de datos.
   c. Para las bases de datos Oracle, ingrese el nombre del servicio.
   d. Introduzca las credenciales para que Clasificación de Datos pueda acceder al servidor.
   e. Seleccione **Agregar servidor de base de datos**.



La base de datos se agrega a la lista de sistemas.

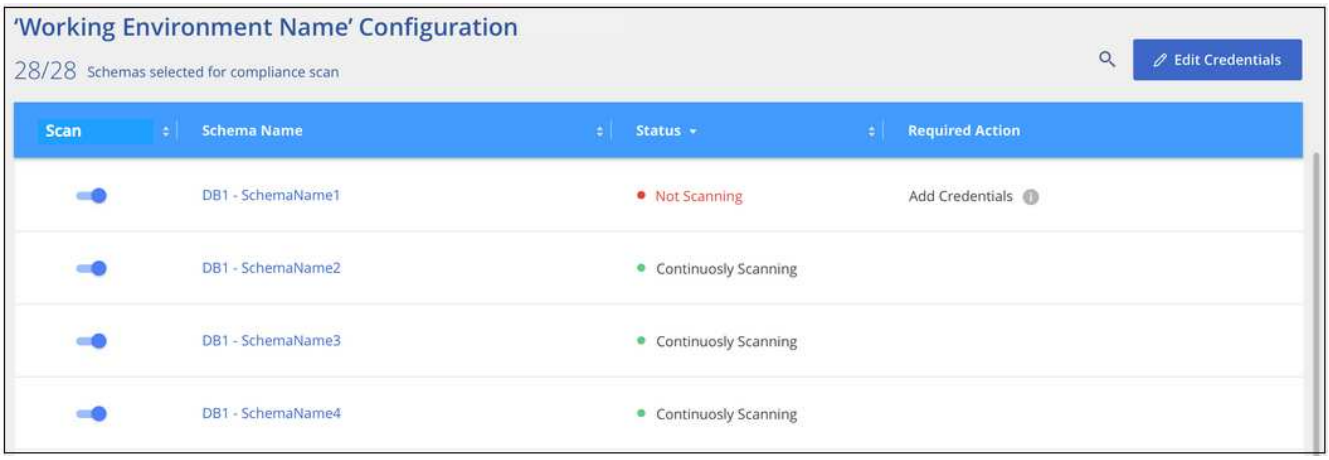
Habilitar y deshabilitar escaneos en esquemas de bases de datos

Puede detener o iniciar el escaneo completo de sus esquemas en cualquier momento.

No existe ninguna opción para seleccionar escaneos de solo mapeo para esquemas de base de datos.

1. Desde la página de Configuración, seleccione el botón **Configuración** para la base de datos que desea configurar.

2. Seleccione los esquemas que desea escanear moviendo el control deslizante hacia la derecha.



Resultado

La clasificación de datos comienza a escanear los esquemas de base de datos que usted habilitó. Puede seguir el progreso del escaneo inicial navegando al menú **Configuración** y luego seleccionando **Configuración del sistema**. El progreso de cada escaneo se muestra como una barra de progreso. También puede pasar el cursor sobre la barra de progreso para ver la cantidad de archivos escaneados en relación con la cantidad total de archivos en el volumen. Si hay algún error, aparecerá en la columna Estado, junto con las acciones necesarias para solucionarlo.

La clasificación de datos escanea sus bases de datos una vez al día; las bases de datos no se escanean continuamente como otras fuentes de datos.

Escanee Google Cloud NetApp Volumes con la NetApp Data Classification

NetApp Data Classification admite Google Cloud NetApp Volumes como sistema. Aprenda a escanear su sistema Google Cloud NetApp Volumes .

Descubra el sistema Google Cloud NetApp Volumes que desea escanear

Si el sistema de Google Cloud NetApp Volumes que desea escanear aún no se encuentra en la NetApp Console como sistema,"agréguelo a la página de Sistemas" .

Implementar la instancia de clasificación de datos

"Implementar la clasificación de datos"si aún no hay una instancia implementada.

La clasificación de datos debe implementarse en la nube al escanear Google Cloud NetApp Volumes y debe implementarse en la misma región que los volúmenes que desea escanear.

Nota: Actualmente, no se admite la implementación de la clasificación de datos en una ubicación local al escanear Google Cloud NetApp Volumes.

Habilite la clasificación de datos en sus sistemas

Puede habilitar la clasificación de datos en su sistema Google Cloud NetApp Volumes .

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione cómo desea escanear los volúmenes en cada sistema. "Aprenda sobre escaneos de mapeo y clasificación":

- Para mapear todos los volúmenes, seleccione **Mape todos los volúmenes**.
- Para mapear y clasificar todos los volúmenes, seleccione **Mapear y clasificar todos los volúmenes**.
- Para personalizar el escaneo para cada volumen, seleccione **O seleccione el tipo de escaneo para cada volumen** y luego elija los volúmenes que desea mapear y/o clasificar.

  VerHabilitar y deshabilitar escaneos en volúmenes Para más detalles.

3. En el cuadro de diálogo de confirmación, seleccione **Aprobar**.

Resultado

La clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados están disponibles en el panel de Cumplimiento tan pronto como la Clasificación de datos finaliza los escaneos iniciales. El tiempo que tarda depende de la cantidad de datos: desde unos minutos hasta unas horas. Puede seguir el progreso del escaneo inicial en la sección **Configuración del sistema** del menú **Configuración**. La clasificación de datos muestra una barra de progreso para cada escaneo. También puede pasar el cursor sobre la barra de progreso para ver la cantidad de archivos escaneados en relación con el total de archivos en el volumen.

- De forma predeterminada, si la clasificación de datos no tiene permisos de atributos de escritura en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos en sus volúmenes porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, seleccione **O seleccione el tipo de escaneo para cada volumen**. La página resultante tiene una configuración que puede habilitar para que la Clasificación de datos escanee los volúmenes independientemente de los permisos.
- La clasificación de datos escanea solo un recurso compartido de archivos bajo un volumen. Si tiene varias acciones en sus volúmenes, deberá escanear esas otras acciones por separado como un grupo de acciones. "Conozca esta limitación de clasificación de datos".

Verificar que la clasificación de datos tenga acceso a los volúmenes

Asegúrese de que la clasificación de datos pueda acceder a los volúmenes verificando su red, grupos de seguridad y políticas de exportación. Para los volúmenes CIFS, debe proporcionar Clasificación de datos con credenciales CIFS.

Para los Google Cloud NetApp Volumes, la clasificación de datos solo puede escanear volúmenes en la misma región que la consola.

Lista de verificación

- Asegúrese de que haya una conexión de red entre la instancia de Clasificación de datos y cada red que incluya volúmenes para Google Cloud NetApp Volumes.
- Asegúrese de que los siguientes puertos estén abiertos para la instancia de clasificación de datos:
  - Para NFS: puertos 111 y 2049.
  - Para CIFS: puertos 139 y 445.
- Asegúrese de que las políticas de exportación de volumen NFS incluyan la dirección IP de la instancia de clasificación de datos para que pueda acceder a los datos de cada volumen.

Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.

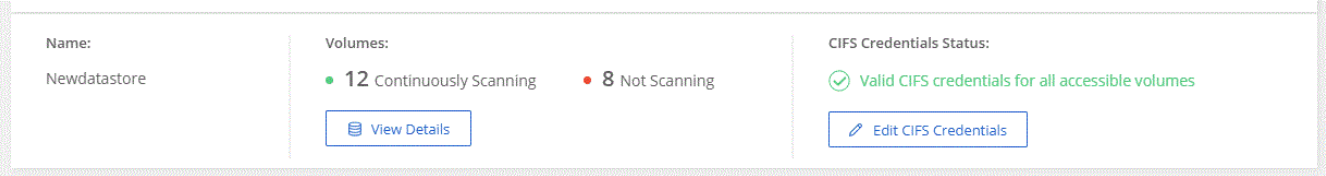
   a. Si utiliza CIFS (SMB), asegúrese de que las credenciales de Active Directory sean correctas. Para

cada sistema, seleccione **Editar credenciales CIFS** y luego ingrese el nombre de usuario y la contraseña que la clasificación de datos necesita para acceder a los volúmenes CIFS en el sistema.

Las credenciales pueden ser de solo lectura, pero proporcionar credenciales de administrador garantiza que Data Classification pueda leer cualquier dato que requiera permisos elevados. Las credenciales se almacenan en la instancia de clasificación de datos.

Si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de datos, se recomienda que el usuario tenga permisos de atributos de escritura en CIFS o permisos de escritura en NFS. Si es posible, configure el usuario de Active Directory como parte de un grupo principal en la organización que tenga permisos para todos los archivos.

Después de ingresar las credenciales, debería ver un mensaje que indica que todos los volúmenes CIFS se autenticaron correctamente.



2. En la página de Configuración, seleccione **Ver detalles** para revisar el estado de cada volumen CIFS y NFS y corregir cualquier error.

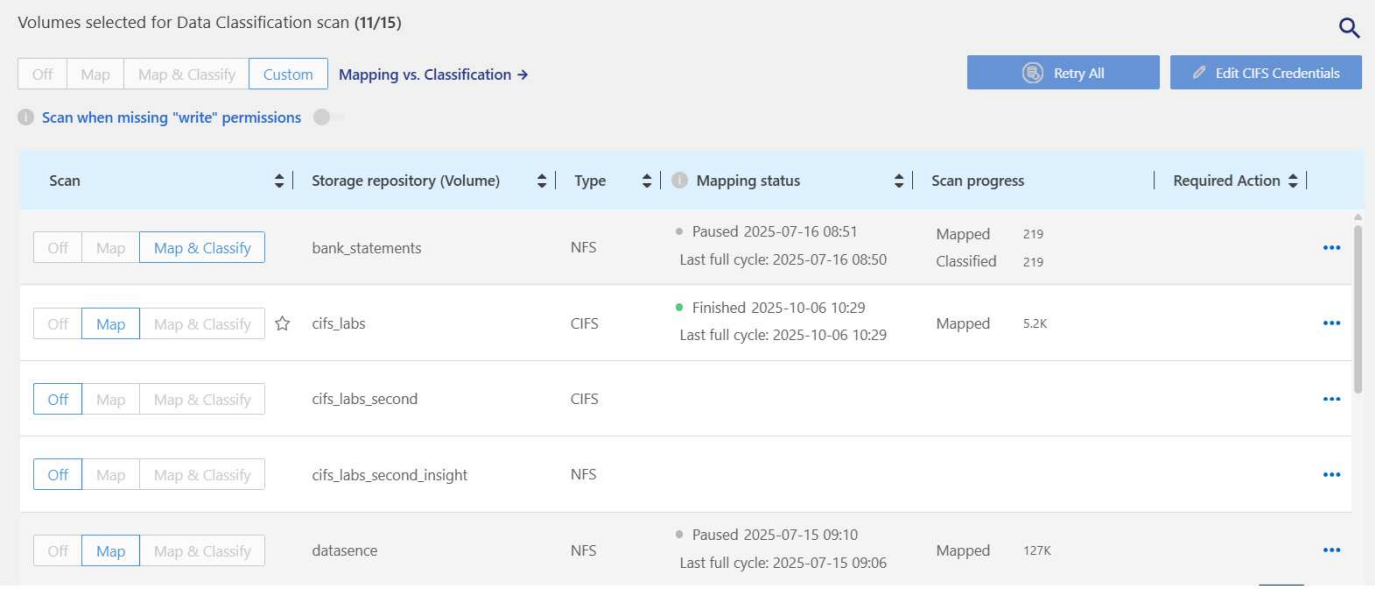
Habilitar y deshabilitar escaneos en volúmenes

Puede iniciar o detener análisis en cualquier sistema en cualquier momento desde la página de Configuración. También puede cambiar los escaneos de solo mapeo a escaneos de mapeo y clasificación, y viceversa. Se recomienda escanear todos los volúmenes de un sistema.

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha seleccionado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando lo configure en **Personalizado** o **Desactivado** en el área de encabezado, deberá activar el mapeo y/o el escaneo completo en cada nuevo volumen que agregue al sistema.

El interruptor en la parte superior de la página para **Escanear cuando faltan permisos de "escritura"** está deshabilitado de manera predeterminada. Esto significa que si la clasificación de datos no tiene permisos de escritura de atributos en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. Si no le importa si se restablece el último tiempo de acceso, active el interruptor y se escanearán todos los archivos independientemente de los permisos. "Más información".

Los nuevos volúmenes agregados al sistema se escanean automáticamente solo cuando usted ha configurado la configuración **Mapa** o **Mapa y clasificar** en el área de encabezado. Cuando la configuración para todos los volúmenes es **Personalizada** o **Desactivada**, deberá activar el escaneo manualmente para cada nuevo volumen que agregue.



Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione un sistema y luego seleccione **Configuración**.
3. Para habilitar o deshabilitar los escaneos para todos los volúmenes, seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** en el encabezado sobre todos los volúmenes.

   Para habilitar o deshabilitar los escaneos de volúmenes individuales, busque los volúmenes en la lista y luego seleccione **Mapa**, **Mapa y clasificar** o **Desactivado** junto al nombre del volumen.

Resultado

Cuando habilita el escaneo, la clasificación de datos comienza a escanear los volúmenes que seleccionó en el sistema. Los resultados comienzan a aparecer en el panel de Cumplimiento tan pronto como la Clasificación de datos inicia el escaneo. El tiempo de finalización del escaneo depende de la cantidad de datos y puede variar entre minutos y horas.

Escanee recursos compartidos de archivos con NetApp Data Classification

Para escanear recursos compartidos de archivos, primero debe crear un grupo de recursos compartidos de archivos en NetApp Data Classification. Los grupos de recursos compartidos de archivos son para recursos compartidos NFS o CIFS (SMB) alojados localmente o en la nube.

La versión principal de Clasificación de datos no admite el escaneo de datos de recursos compartidos de archivos que no sean de NetApp .

Prerrequisitos

Revise los siguientes requisitos previos para asegurarse de tener una configuración compatible antes de habilitar la Clasificación de datos.

- Las acciones se pueden alojar en cualquier lugar, incluso en la nube o en las instalaciones locales. Los recursos compartidos CIFS de sistemas de almacenamiento NetApp 7-Mode más antiguos se pueden escanear como recursos compartidos de archivos.

- La clasificación de datos no puede extraer permisos ni la "hora del último acceso" de los sistemas 7-Mode.
- Debido a un problema conocido entre algunas versiones de Linux y los recursos compartidos CIFS en sistemas 7-Mode, debe configurar el recurso compartido para usar solo SMBv1 con la autenticación NTLM habilitada.

- Debe haber conectividad de red entre la instancia de clasificación de datos y los recursos compartidos.
- Puede agregar un recurso compartido DFS (sistema de archivos distribuido) como un recurso compartido CIFS normal. Debido a que la clasificación de datos no sabe que el recurso compartido está construido sobre múltiples servidores/volúmenes combinados como un único recurso compartido CIFS, es posible que reciba errores de permiso o conectividad acerca del recurso compartido cuando el mensaje en realidad solo se aplica a una de las carpetas/recursos compartidos que se encuentra en un servidor/volumen diferente.
- Para los recursos compartidos CIFS (SMB), asegúrese de tener credenciales de Active Directory que proporcionen acceso de lectura a los recursos compartidos. Se prefieren las credenciales de administrador en caso de que la clasificación de datos necesite escanear datos que requieran permisos elevados.

  Si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de datos, se recomienda que el usuario tenga permisos de atributos de escritura en CIFS o permisos de escritura en NFS. Si es posible, configure el usuario de Active Directory como parte de un grupo principal en la organización que tenga permisos para todos los archivos.

- Todos los recursos compartidos de archivos CIFS de un grupo deben utilizar las mismas credenciales de Active Directory.
- Puede combinar recursos compartidos NFS y CIFS (utilizando Kerberos o NTLM). Debes agregar las acciones al grupo por separado. Es decir, debes completar el proceso dos veces: una por protocolo.
  - No se puede crear un grupo de recursos compartidos de archivos que combine tipos de autenticación CIFS (Kerberos y NTLM).
- Si utiliza CIFS con autenticación Kerberos, asegúrese de que la dirección IP proporcionada sea accesible para la clasificación de datos. No se pueden agregar archivos compartidos si la dirección IP no es accesible.

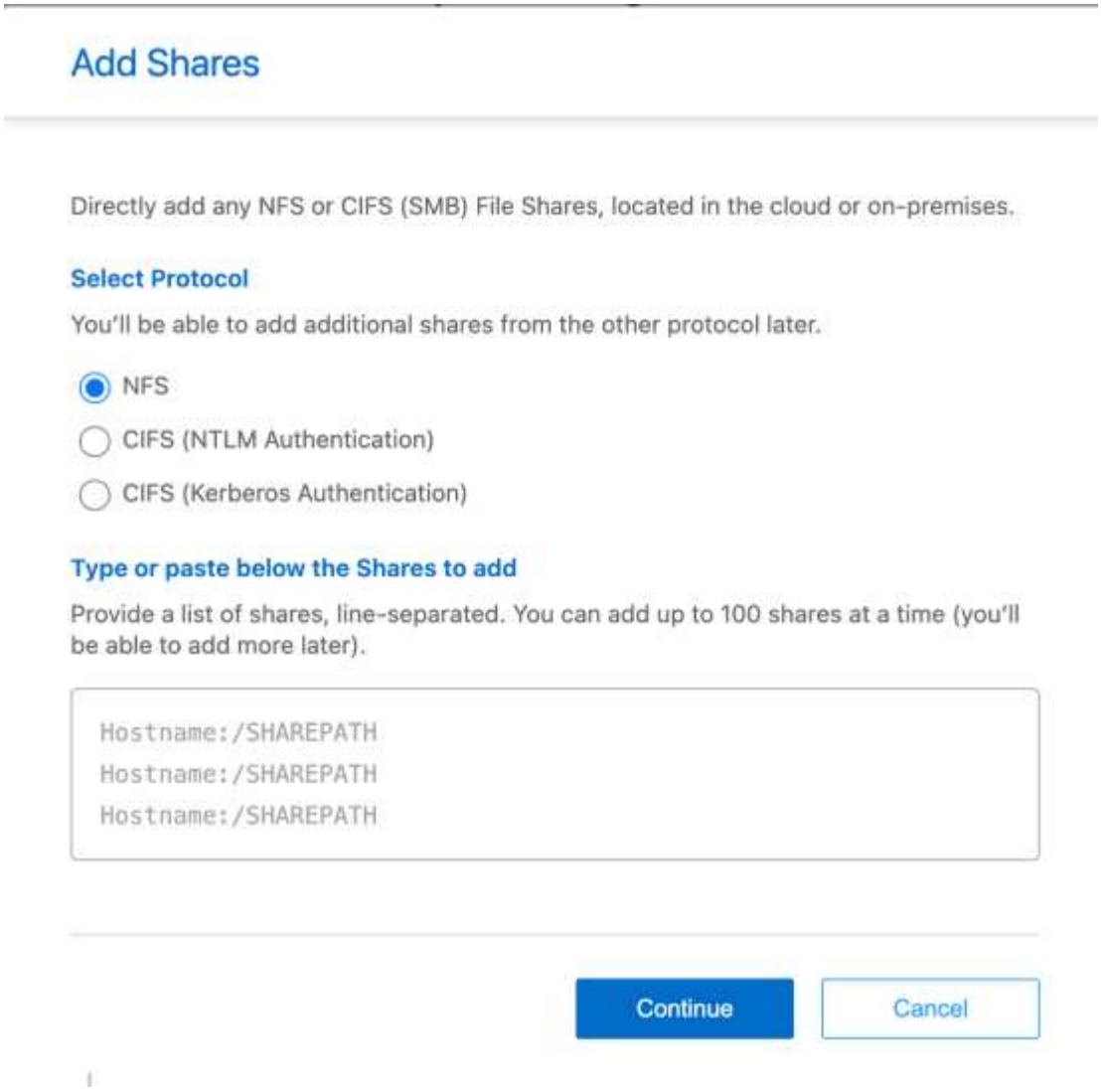
Crear un grupo de recursos compartidos de archivos

Cuando agregue recursos compartidos de archivos al grupo, debe utilizar el formato `<host_name>:/<share_path>`.

Puede agregar recursos compartidos de archivos individualmente o puede ingresar una lista separada por líneas de los recursos compartidos de archivos que desea escanear. Puedes agregar hasta 100 acciones a la vez.

Pasos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Desde la página de Configuración, seleccione **Agregar sistema** > **Agregar grupo de recursos compartidos de archivos**.
3. En el cuadro de diálogo Agregar grupo de recursos compartidos de archivos, ingrese el nombre del grupo de recursos compartidos y luego seleccione **Continuar**.
4. Seleccione el protocolo para los recursos compartidos de archivos que está agregando.



a. Si está agregando recursos compartidos CIFS con autenticación NTLM, ingrese las credenciales de Active Directory para acceder a los volúmenes CIFS. Aunque se admiten credenciales de solo lectura, se recomienda proporcionar acceso completo con credenciales de administrador. Seleccione **Guardar**.

5. Agregue los recursos compartidos de archivos que desea escanear (un recurso compartido de archivos por línea). Luego seleccione **Continuar**.
6. Un cuadro de diálogo de confirmación muestra la cantidad de acciones que se agregaron.

   Si el cuadro de diálogo enumera recursos compartidos que no se pudieron agregar, capture esta información para poder resolver el problema. Si el problema está relacionado con una convención de nomenclatura, puede volver a agregar el recurso compartido con un nombre corregido.

7. Configurar el escaneo en el volumen:
   - Para habilitar escaneos de solo mapeo en recursos compartidos de archivos, seleccione **Mapa**.
   - Para habilitar escaneos completos en recursos compartidos de archivos, seleccione **Mapear y clasificar**.
   - Para deshabilitar el escaneo en recursos compartidos de archivos, seleccione **Desactivado**.

El interruptor en la parte superior de la página para **Escanear cuando faltan permisos de "atributos de escritura"** está deshabilitado de manera predeterminada. Esto significa que si la clasificación de datos no tiene permisos de escritura de atributos en CIFS o permisos de escritura en NFS, el sistema no escaneará los archivos porque la clasificación de datos no puede revertir la "hora del último acceso" a la marca de tiempo original. + Si cambia **Escanear cuando faltan permisos de "atributos de escritura"** a **Activado**, el escaneo restablece la última hora de acceso y escanea todos los archivos independientemente de los permisos. + Para obtener más información sobre la marca de tiempo del último acceso, consulte"Metadatos recopilados de fuentes de datos en la clasificación de datos" .

Resultado

La clasificación de datos comienza a escanear los archivos en los recursos compartidos de archivos que agregó. PuedeSeguimiento del progreso del escaneo y ver los resultados del escaneo en el **Panel de Control**.

Si el escaneo no se completa exitosamente para una configuración CIFS con autenticación Kerberos, verifique la pestaña **Configuración** para ver si hay errores.

Editar un grupo de recursos compartidos de archivos

Después de crear un grupo de recursos compartidos de archivos, puede editar el protocolo CIFS o agregar y eliminar recursos compartidos de archivos.

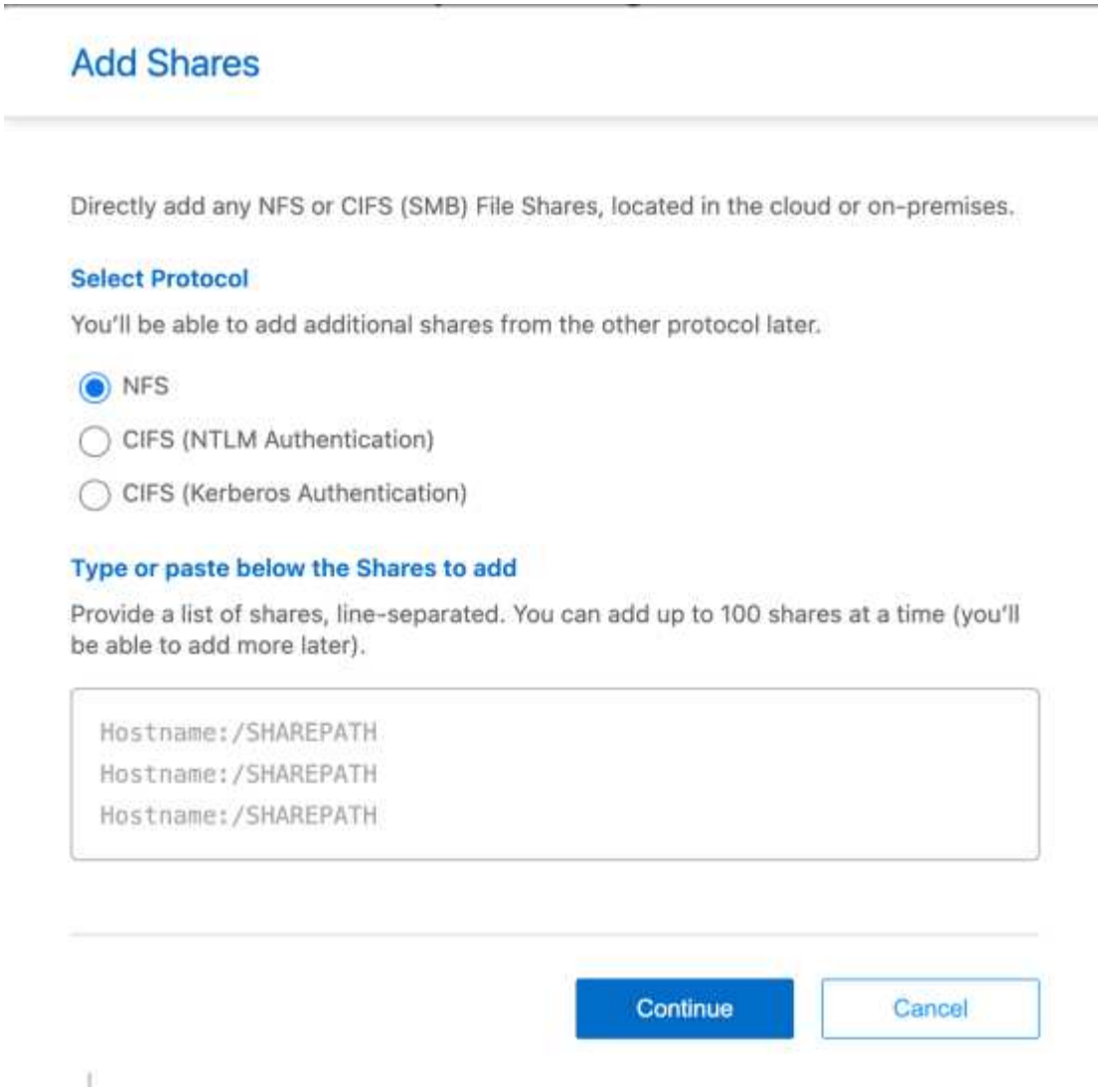
Editar la configuración del protocolo CIFS

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Desde la página de Configuración, seleccione el grupo de recursos compartidos de archivos que desea modificar.
3. Seleccione **Editar credenciales CIFS**.

4. Elija el método de autenticación: **NTLM** o **Kerberos**.
5. Ingrese el **nombre de usuario** y la **contraseña** del Directorio Activo.
6. Seleccione **Guardar** para completar el proceso.

Agregar recursos compartidos de archivos a los escaneos

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Desde la página de Configuración, seleccione el grupo de recursos compartidos de archivos que desea modificar.
3. Seleccione **+ Agregar acciones**.
4. Seleccione el protocolo para los recursos compartidos de archivos que está agregando.



Si está agregando recursos compartidos de archivos a un protocolo que ya ha configurado, no se requieren cambios.

Si está agregando recursos compartidos de archivos con un segundo protocolo, asegúrese de haber configurado correctamente la autenticación como se detalla en"prerrequisitos" .

5. Agregue los recursos compartidos de archivos que desea escanear (un recurso compartido de archivos por línea) utilizando el formato `<host_name>:/<share_path>` .
6. Seleccione **Continuar** para completar la adición de los recursos compartidos de archivos.

Eliminar un recurso compartido de archivos de los análisis

1. Desde el menú Clasificación de datos, seleccione **Configuración**.
2. Seleccione el sistema del cual desea eliminar los recursos compartidos de archivos.
3. Seleccione **Configuración**.
4. Desde la página de Configuración, seleccione Acciones ••• para el recurso compartido de archivos que desea eliminar.
5. En el menú Acciones, seleccione **Eliminar recurso compartido**.

Seguimiento del progreso del escaneo

Puede realizar un seguimiento del progreso del escaneo inicial.

1. Seleccione el menú **Configuración**.
2. Seleccione la **Configuración del sistema**.
3. Para el repositorio de almacenamiento, consulte la columna Progreso del escaneo para ver su estado.

Escanee datos de StorageGRID con la NetApp Data Classification

Complete unos pocos pasos para comenzar a escanear datos dentro de StorageGRID directamente con NetApp Data Classification.

Revisar los requisitos de StorageGRID

Revise los siguientes requisitos previos para asegurarse de tener una configuración compatible antes de habilitar la Clasificación de datos.

- Debe tener la URL del punto final para conectarse con el servicio de almacenamiento de objetos.
- Debe tener la clave de acceso y la clave secreta de StorageGRID para que la clasificación de datos pueda acceder a los depósitos.

Implementar la instancia de clasificación de datos

Implementar la clasificación de datos si aún no hay una instancia implementada.

Si está escaneando datos de StorageGRID a los que se puede acceder a través de Internet, puede"Implementar la clasificación de datos en la nube" o"Implementar la clasificación de datos en una ubicación local que tenga acceso a Internet" .

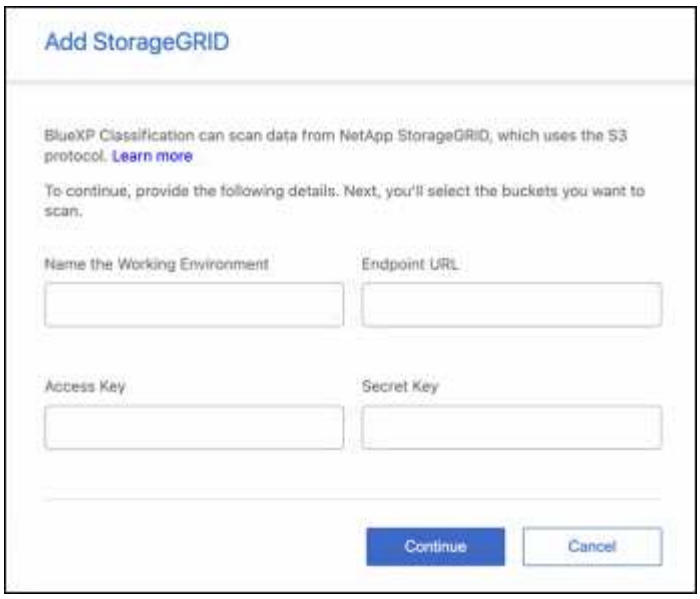
Si está escaneando datos de StorageGRID que se instaló en un sitio oscuro que no tiene acceso a Internet, debe"Implementar la clasificación de datos en la misma ubicación local que no tiene acceso a Internet" . Esto también requiere que el agente de consola esté implementado en esa misma ubicación local.

Agregue el servicio StorageGRID a la clasificación de datos

Agregue el servicio StorageGRID .

Pasos

1. Desde el menú Clasificación de datos, seleccione la opción **Configuración**.
2. Desde la página de Configuración, seleccione **Agregar sistema** > **Agregar StorageGRID**.
3. En el cuadro de diálogo Agregar servicio StorageGRID , ingrese los detalles del servicio StorageGRID y seleccione **Continuar**.
   a. Introduzca el nombre que desea utilizar para el Sistema. Este nombre debe reflejar el nombre del servicio StorageGRID al que se está conectando.
   b. Introduzca la URL del punto final para acceder al servicio de almacenamiento de objetos.
   c. Ingrese la clave de acceso y la clave secreta para que la clasificación de datos pueda acceder a los depósitos en StorageGRID.



Resultado

StorageGRID se agrega a la lista de sistemas.

Habilitar y deshabilitar escaneos en depósitos StorageGRID

Después de habilitar la Clasificación de datos en StorageGRID, el siguiente paso es configurar los depósitos que desea escanear. La clasificación de datos descubre esos grupos y los muestra en el sistema que usted creó.

Pasos

1. En la página de Configuración, busque el sistema StorageGRID .
2. En el mosaico del sistema StorageGRID , seleccione **Configuración**.
3. Complete uno de los siguientes pasos para habilitar o deshabilitar el escaneo:
   - Para habilitar escaneos de solo mapeo en un bucket, seleccione **Mapa**.
   - Para habilitar escaneos completos en un bucket, seleccione **Mapear y clasificar**.
   - Para deshabilitar el escaneo en un depósito, seleccione **Desactivado**.

Resultado

La clasificación de datos comienza a escanear los grupos que usted habilitó. Puede seguir el progreso del escaneo inicial navegando al menú **Configuración** y luego seleccionando **Configuración del sistema**. El progreso de cada escaneo se muestra como una barra de progreso. También puede pasar el cursor sobre la barra de progreso para ver la cantidad de archivos escaneados en relación con el total de archivos en el volumen. Si hay algún error, aparecerá en la columna Estado, junto con la acción necesaria para solucionarlo.

Integre su Active Directory con NetApp Data Classification

Puede integrar un Active Directory global con NetApp Data Classification para mejorar los resultados que Data Classification informa sobre los propietarios de archivos y qué usuarios y grupos tienen acceso a sus archivos.

Cuando configura ciertas fuentes de datos (enumeradas a continuación), debe ingresar las credenciales de Active Directory para que la clasificación de datos escanee los volúmenes CIFS. Esta integración proporciona

clasificación de datos con detalles del propietario del archivo y permisos para los datos que residen en esas fuentes de datos. El Active Directory ingresado para esas fuentes de datos puede ser diferente de las credenciales de Active Directory globales que ingrese aquí. La clasificación de datos buscará detalles de usuarios y permisos en todos los directorios activos integrados.

Esta integración proporciona información adicional en las siguientes ubicaciones en Clasificación de datos:

- Puedes utilizar el "Propietario del archivo""filtrar" y ver los resultados en los metadatos del archivo en el panel Investigación. En lugar de que el propietario del archivo contenga el SID (identificador de seguridad), se completa con el nombre de usuario real.

  También puede ver más detalles sobre el propietario del archivo: nombre de la cuenta, dirección de correo electrónico y nombre de la cuenta SAM, o ver los elementos que pertenecen a ese usuario.

- Ya puedes ver"permisos de archivo completos" para cada archivo y directorio cuando hace clic en el botón "Ver todos los permisos".
- En el"Panel de gobernanza" , el panel Permisos abiertos mostrará un mayor nivel de detalle sobre sus datos.

> Los SID de usuarios locales y los SID de dominios desconocidos no se traducen al nombre de usuario real.

Fuentes de datos compatibles

Una integración de Active Directory con clasificación de datos puede identificar datos de las siguientes fuentes de datos:

- Sistemas ONTAP locales
- Cloud Volumes ONTAP
- Azure NetApp Files
- FSx para ONTAP

Conéctese a su servidor de Active Directory

Una vez que haya implementado la Clasificación de datos y haya activado el escaneo en sus fuentes de datos, puede integrar la Clasificación de datos con su Active Directory. Se puede acceder a Active Directory mediante una dirección IP de servidor DNS o una dirección IP de servidor LDAP.

Las credenciales de Active Directory pueden ser de solo lectura, pero proporcionar credenciales de administrador garantiza que Data Classification pueda leer cualquier dato que requiera permisos elevados. Las credenciales se almacenan en la instancia de clasificación de datos.

Para volúmenes CIFS/recursos compartidos de archivos, si desea asegurarse de que los "últimos tiempos de acceso" de sus archivos no se modifiquen mediante los análisis de clasificación de clasificación de datos, el usuario debe tener permiso de escritura de atributos. Si es posible, recomendamos hacer que el usuario configurado de Active Directory sea parte de un grupo principal en la organización que tenga permisos para todos los archivos.

Requisitos

- Debe tener un Directorio Activo ya configurado para los usuarios de su empresa.
- Debes tener la información del Directorio Activo:

- Dirección IP del servidor DNS o varias direcciones IP

  o

  Dirección IP del servidor LDAP o varias direcciones IP

- Nombre de usuario y contraseña para acceder al servidor
- Nombre de dominio (nombre de Active Directory)
- Ya sea que esté utilizando LDAP seguro (LDAPS) o no
- Puerto del servidor LDAP (normalmente 389 para LDAP y 636 para LDAP seguro)

- Los siguientes puertos deben estar abiertos para la comunicación saliente por parte de la instancia de clasificación de datos:

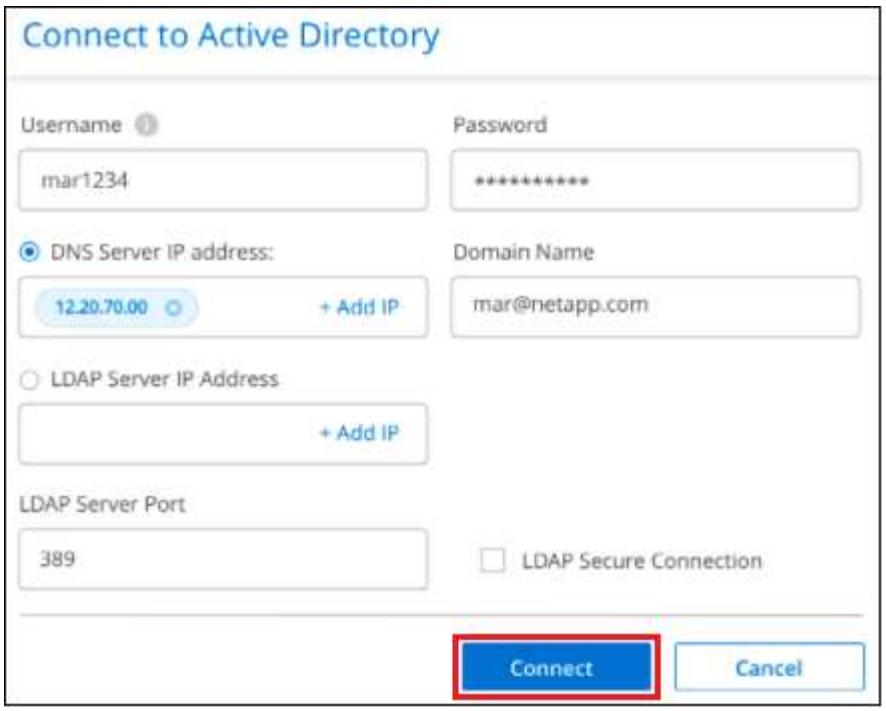
Protocolo	Puerto	Destino	Objetivo
TCP y UDP	389	Directorio activo	LDAP
TCP	636	Directorio activo	LDAP sobre SSL
TCP	3268	Directorio activo	Catálogo global
TCP	3269	Directorio activo	Catálogo global sobre SSL

Pasos

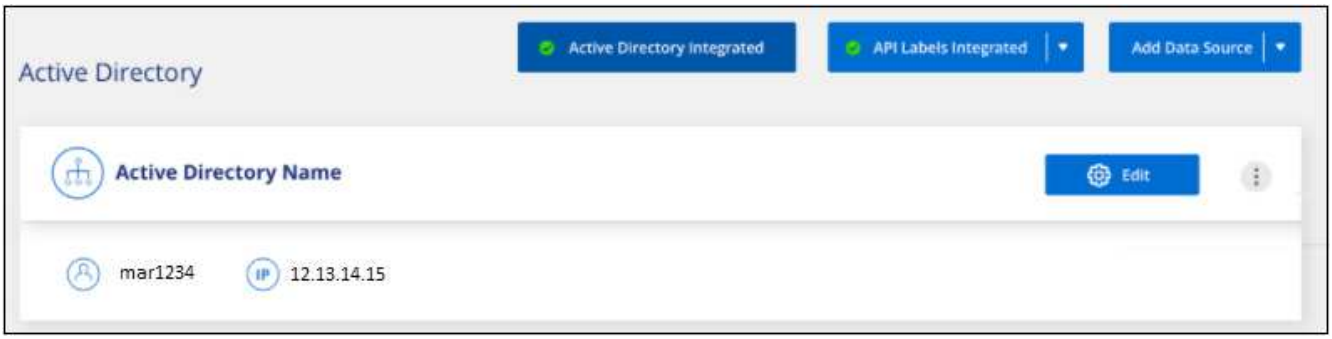
1. Desde la página Configuración de clasificación de datos, haga clic en **Agregar Active Directory**.

2. En el cuadro de diálogo Conectar a Active Directory, ingrese los detalles de Active Directory y haga clic en **Conectar**.

   Puede agregar varias direcciones IP, si es necesario, seleccionando **Agregar IP**.



La clasificación de datos se integra al Directorio Activo y se agrega una nueva sección a la página de Configuración.



Administre su integración de Active Directory

Si necesita modificar algún valor en su integración de Active Directory, haga clic en el botón **Editar** y realice los cambios.

También puedes eliminar la integración seleccionando la opción botón y luego **Eliminar Active Directory**.