Administrar la clasificación de datos

NetApp Data Classification

NetApp
February 11, 2026

Tabla de contenidos

Administrar la clasificación de datos

Excluir directorios específicos de los análisis de NetApp Data Classification

Si desea que NetApp Data Classification excluya directorios específicos de los análisis, puede agregar estos nombres de directorio a un archivo de configuración. Después de aplicar este cambio, el motor de clasificación de datos excluye esos directorios de los análisis.

De forma predeterminada, los análisis de clasificación de datos excluyen los datos de instantáneas de volumen, que son idénticos a su origen en el volumen.

Fuentes de datos compatibles

La exclusión de directorios específicos de los análisis de clasificación de datos es compatible con recursos compartidos NFS y CIFS en las siguientes fuentes de datos:

- ONTAP local
- Cloud Volumes ONTAP
- Amazon FSx for NetApp ONTAP
- Azure NetApp Files
- Recursos compartidos de archivos generales

Define los directorios que se excluirán del escaneo

Antes de poder excluir directorios del escaneo de clasificación, debe iniciar sesión en el sistema de clasificación de datos para poder editar un archivo de configuración y ejecutar un script. Vea cómo"Iniciar sesión en el sistema de clasificación de datos" dependiendo de si instaló manualmente el software en una máquina Linux o si implementó la instancia en la nube.

Consideraciones

- Puede excluir un máximo de 50 rutas de directorio por sistema de clasificación de datos.
- La exclusión de rutas de directorio puede afectar los tiempos de escaneo.

Pasos

1. En el sistema de clasificación de datos, vaya a "/opt/netapp/config/custom_configuration" y luego abra el archivo `data_provider.yaml` .
2. En la sección "data_providers", debajo de la línea "exclude:", ingrese las rutas de directorio que desea excluir. Por ejemplo:

```
exclude:
- "folder1"
- "folder2"
```

No modifique nada más en este archivo.

3. Guarde los cambios en el archivo.
4. Vaya a "/opt/netapp/Datasense/tools/customer_configuration/data_providers" y ejecute el siguiente script:

```
update_data_providers_from_config_file.sh
```

+ Este comando envía los directorios que se excluirán del escaneo al motor de clasificación.

Resultado

Todos los escaneos posteriores de sus datos excluirán el escaneo de aquellos directorios especificados.

Puede agregar, editar o eliminar elementos de la lista de exclusión siguiendo estos mismos pasos. La lista de exclusión revisada se actualizará después de ejecutar el script para confirmar los cambios.

Ejemplos

Configuración 1:

Cualquier carpeta que contenga "carpeta1" en cualquier parte del nombre será excluida de todas las fuentes de datos.

```
data_providers:
   exclude:
   - "folder1"
```

Resultados esperados para las rutas que se excluirán:

- /CVO1/carpeta1
- /CVO1/nombrecarpeta1
- /CVO1/carpeta10
- /CVO1/*carpeta1
- /CVO1/+nombrecarpeta1
- /CVO1/nocarpeta10
- /CVO22/carpeta1
- /CVO22/nombrecarpeta1
- /CVO22/carpeta10

Ejemplos de rutas que no se excluirán:

- /CVO1/*carpeta
- /CVO1/nombredecarpeta
- /CVO22/*carpeta20

Configuración 2:

Se excluirá cualquier carpeta que contenga "*folder1" sólo al comienzo del nombre.

```
data_providers:
   exclude:
   - "\\*folder1"
```

Resultados esperados para las rutas que se excluirán:

- /CVO/*carpeta1
- /CVO/*nombredecarpeta1
- /CVO/*carpeta10

Ejemplos de rutas que no se excluirán:

- /CVO/carpeta1
- /CVO/nombrecarpeta1
- /CVO/no*carpeta10

Configuración 3:

Se excluirá cualquier carpeta en la fuente de datos "CVO22" que contenga "carpeta1" en cualquier parte del nombre.

```
data_providers:
   exclude:
   - "CVO22/folder1"
```

Resultados esperados para las rutas que se excluirán:

- /CVO22/carpeta1
- /CVO22/nombrecarpeta1
- /CVO22/carpeta10

Ejemplos de rutas que no se excluirán:

- /CVO1/carpeta1
- /CVO1/nombrecarpeta1
- /CVO1/carpeta10

Cómo escapar caracteres especiales en los nombres de carpetas

Si tiene un nombre de carpeta que contiene uno de los siguientes caracteres especiales y desea excluir los datos de esa carpeta del análisis, deberá utilizar la secuencia de escape \\ antes del nombre de la carpeta.

```
 ., +, *, ?, ^, $, (, ), [, ], {, }, |
Por ejemplo:
```

Ruta en la fuente: `/project/*not_to_scan`

Sintaxis en archivo de exclusión: `"\\*not_to_scan"`

Ver la lista de exclusiones actual

Es posible que el contenido de la `data_provider.yaml` archivo de configuración para que sea diferente de lo que realmente se confirmó después de ejecutar el `update_data_providers_from_config_file.sh` guion. Para ver la lista actual de directorios que ha excluido del análisis de clasificación de datos, ejecute el siguiente comando desde "/opt/netapp/Datasense/tools/customer_configuration/data_providers":

```
get_data_providers_configuration.sh
```

Definir identificadores de grupo adicionales como abiertos a la organización en la NetApp Data Classification

Cuando los identificadores de grupo (GID) se adjuntan a archivos o carpetas en recursos compartidos de archivos NFS, definen los permisos para el archivo o la carpeta; por ejemplo, si están "abiertos para la organización". Si algunos GID no están configurados inicialmente con el nivel de permiso "Abierto a la organización", puede agregar ese permiso al GID para que todos los archivos y carpetas que tengan ese GID adjunto se consideren "abiertos a la organización".

Después de realizar este cambio y NetApp Data Classification vuelva a escanear sus archivos y carpetas, todos los archivos y carpetas que tengan estos ID de grupo adjuntos mostrarán este permiso en la página Detalles de la investigación y también aparecerán en los informes donde se muestren los permisos de archivos.

Para activar esta funcionalidad, debe iniciar sesión en el sistema de clasificación de datos para poder editar un archivo de configuración y ejecutar un script. Vea cómo"Iniciar sesión en el sistema de clasificación de datos" dependiendo de si instaló manualmente el software en una máquina Linux o si implementó la instancia en la nube.

Agregue el permiso "abierto a la organización" a los ID de grupo

Debe tener los números de identificación de grupo (GID) antes de comenzar esta tarea.

Pasos

1. En el sistema de clasificación de datos, vaya a "/opt/netapp/config/custom_configuration" y abra el archivo `data_provider.yaml` .
2. En la línea "organization_group_ids: []" agregue los ID del grupo. Por ejemplo:

   ```
   organization_group_ids: [1014, 1015, 21, 2021, 1013, 2020, 1018, 1019]
   ```

   No cambie nada más en este archivo.
3. Guarde los cambios en el archivo.
4. Vaya a "/opt/netapp/Datasense/tools/customer_configuration/data_providers" y ejecute el siguiente script:

```
update_data_providers_from_config_file.sh
```

Este comando envía los permisos de ID de grupo revisados al motor de clasificación.

Resultado

Todos los análisis posteriores de sus datos identificarán los archivos o carpetas que tengan estos ID de grupo adjuntos como "abiertos a la organización".

Puede editar la lista de ID de grupo y eliminar cualquier ID de grupo que haya agregado en el pasado siguiendo estos mismos pasos. La lista revisada de ID de grupo se actualizará después de ejecutar el script para confirmar los cambios.

Ver la lista actual de ID de grupo

Es posible que el contenido de la `data_provider.yaml` archivo de configuración para que difiera de lo que realmente se ha confirmado después de ejecutar el `update_data_providers_from_config_file.sh` guion. Para ver la lista actual de ID de grupo que ha agregado a Clasificación de datos, ejecute el siguiente comando desde "/opt/netapp/Datasense/tools/customer_configuration/data_providers":

```
get_data_providers_configuration.sh
```

Personalice la definición de datos obsoletos en NetApp Data Classification

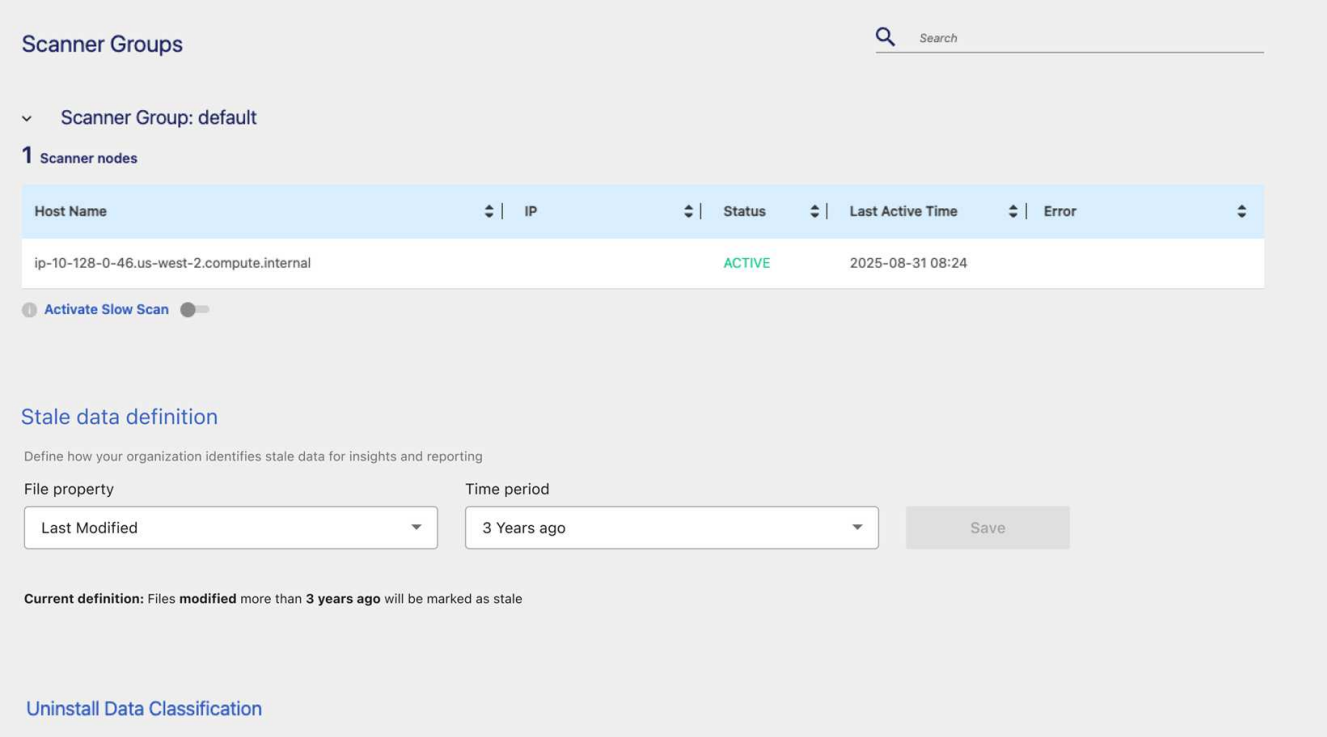
La NetApp Data Classification identifica datos obsoletos para ayudarlo a identificar oportunidades de ahorro y riesgos de gobernanza. Debido a que la definición de datos obsoletos puede variar según los diferentes contextos organizacionales, puede personalizar el modo en que la Clasificación de datos define los datos obsoletos.

Los datos obsoletos se pueden definir en función de cuándo se *accedió a ellos por última vez* o cuándo se *modificó por última vez*. Las selecciones del período de tiempo varían desde hace 6 meses hasta hace 10 años.

De forma predeterminada, los datos se consideran obsoletos si se modificaron por última vez hace tres años.

Definir datos obsoletos

1. En Ransomware Resilience, seleccione **Configuración**.
2. En la página de Configuración, desplácese hasta el encabezado **Definición de datos obsoletos**.
3. En el menú desplegable **Propiedades de archivo**, elija si desea definir datos obsoletos según cuándo se accedió por última vez o se modificó por última vez.
4. Seleccione el período de tiempo para la definición de datos obsoletos.



5. Seleccione **Guardar**.

Eliminar fuentes de datos de NetApp Data Classification

Si es necesario, puede detener que NetApp Data Classification escanee uno o más sistemas, bases de datos o grupos de recursos compartidos de archivos.

Desactivar los análisis de un sistema

Cuando desactiva los escaneos, la Clasificación de datos ya no escanea los datos en el sistema y elimina la información indexada de la instancia de Clasificación de datos. Los datos del propio sistema no se eliminan.

1. Desde la página *Configuración*, seleccione la botón en la fila del sistema y luego **Desactivar clasificación de datos**.

> También puede deshabilitar los análisis de un sistema desde el panel Servicios cuando selecciona el sistema.

Eliminar una base de datos de Clasificación de datos

Si ya no necesita escanear una determinada base de datos, puede eliminarla de la interfaz de Clasificación de datos y detener todos los escaneos.

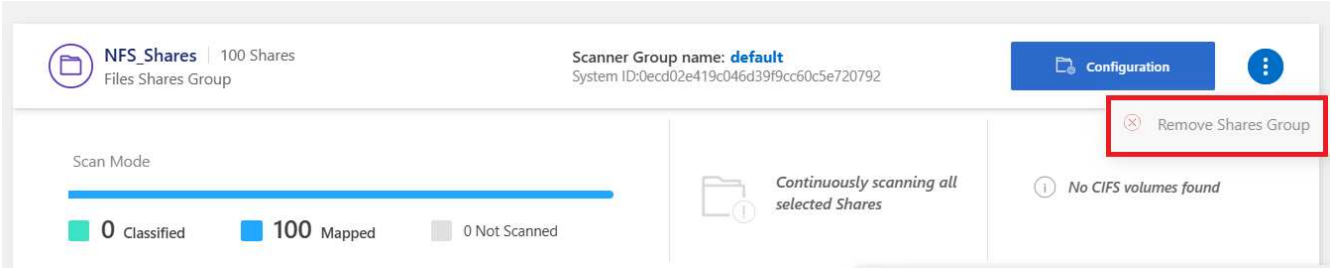
1. Desde la página *Configuración*, seleccione la botón en la fila de la base de datos y luego **Eliminar servidor de base de datos**.

Eliminar un grupo de recursos compartidos de archivos de la Clasificación de datos

Si ya no desea escanear archivos de usuario de un grupo de recursos compartidos de archivos, puede eliminar el grupo de recursos compartidos de archivos de la interfaz de Clasificación de datos y detener todos los escaneos.

Pasos

1. Desde la página *Configuración*, seleccione la botón en la fila del grupo de recursos compartidos de archivos y luego **Eliminar grupo de recursos compartidos de archivos**.



2. Seleccione **Eliminar grupo de acciones** en el cuadro de diálogo de confirmación.

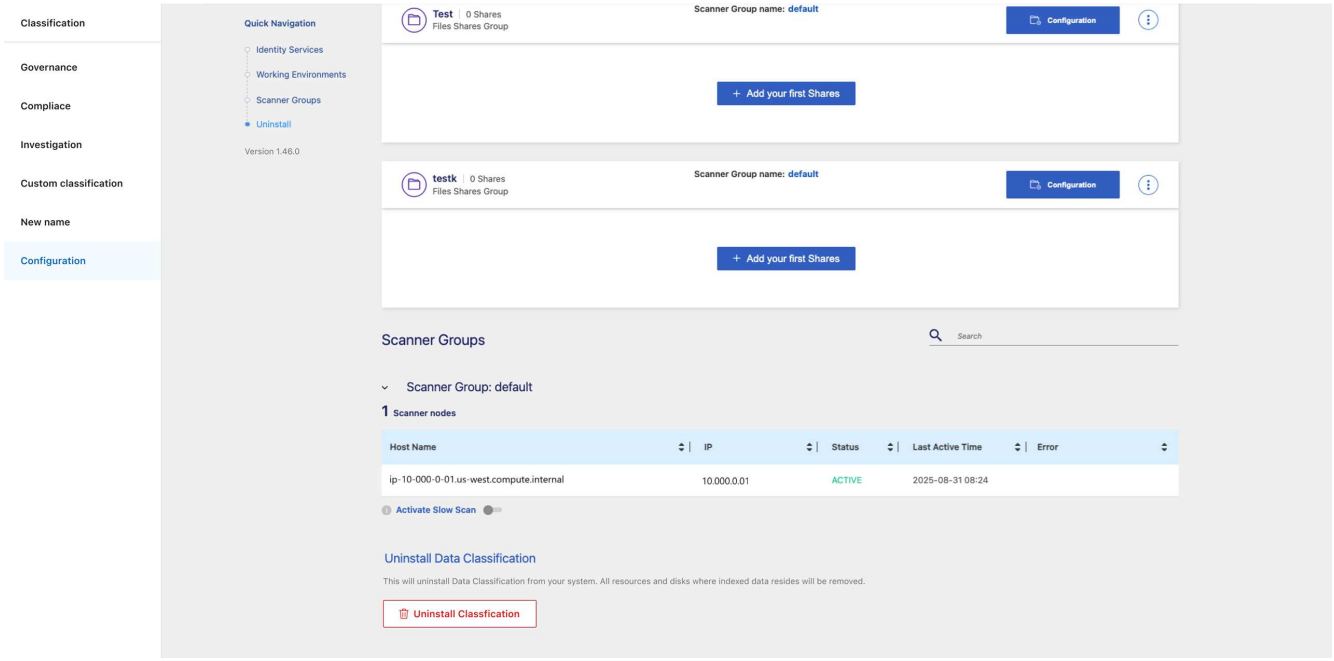
Desinstalar NetApp Data Classification

Puede desinstalar NetApp Data Classification para solucionar problemas o eliminar permanentemente el software del host. Al eliminar la instancia también se eliminan los discos asociados donde residen los datos indexados, lo que significa que toda la información que Data Classification ha escaneado se eliminará de forma permanente.

Los pasos que debes seguir dependen de si implementaste la clasificación de datos en la nube o en un host local.

Desinstalar la clasificación de datos de un proveedor de nube

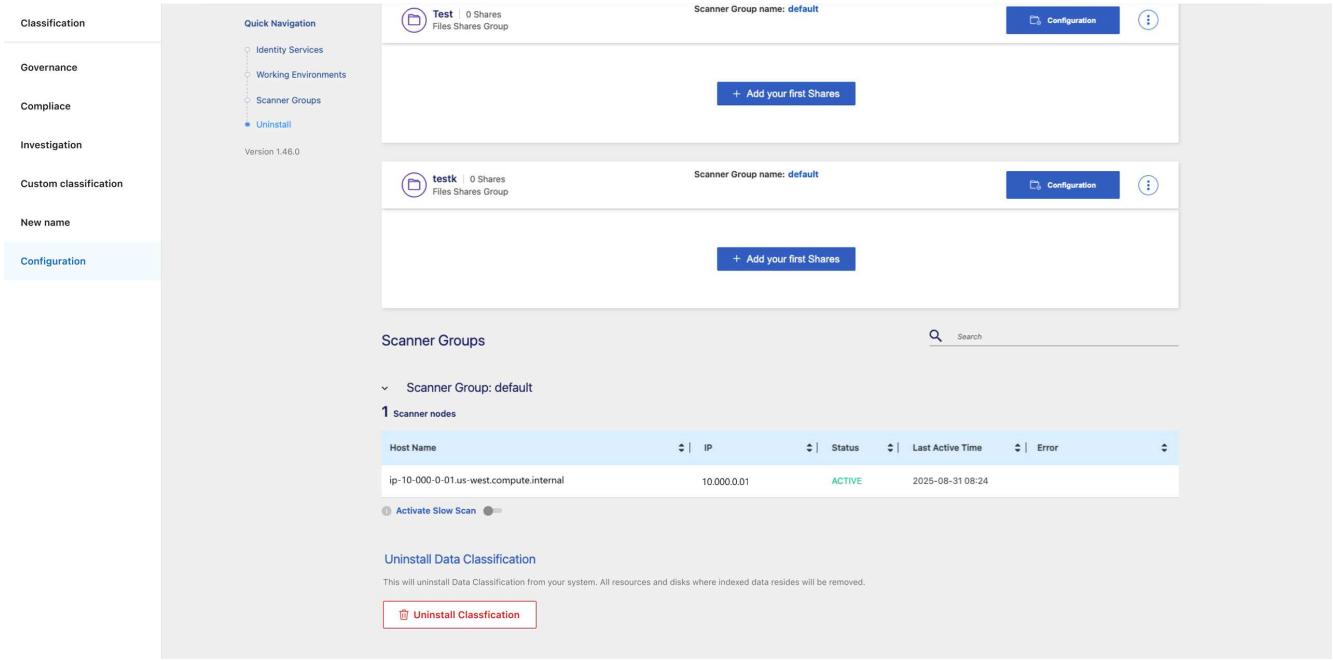
1. Desde Clasificación de datos, seleccione **Configuración**.
2. En la parte inferior de la página de configuración, seleccione **Desinstalar clasificación**.



3. En el cuadro de diálogo, ingrese "desinstalar" para continuar con la desconexión de la instancia de Clasificación de datos del agente de Consola. Seleccione **Desinstalar** para confirmar.
4. En el cuadro de diálogo *Desinstalar clasificación*, escriba **uninstall** para confirmar que desea desconectar la instancia de Clasificación de datos del agente de Consola y luego seleccione **Desinstalar**.
5. Para finalizar el proceso de desinstalación, vaya a la consola de su proveedor de nube y elimine la instancia de Clasificación de datos. La instancia se llama *CloudCompliance* con un hash generado (UUID) concatenado a ella. Por ejemplo: *CloudCompliance-16bb6564-38ad-4080-9a92-36f5fd2f71c7*

Desinstalar la clasificación de datos de una implementación local

1. Desde Clasificación de datos, seleccione **Configuración**.
2. En la parte inferior de la página de configuración, seleccione **Desinstalar clasificación**.



3. En el cuadro de diálogo, ingrese "desinstalar" para continuar con la desconexión de la instancia de Clasificación de datos del agente de Consola. Seleccione **Desinstalar** para confirmar.
4. Para desinstalar el software del host, ejecute el `cleanup.sh` script en la máquina host de clasificación de datos, por ejemplo:

```
cleanup.sh
```

El guión se encuentra en el `/install/light_probe/onprem_installer/cleanup.sh` directorio. Vea cómo"Inicie sesión en la máquina host de clasificación de datos" .

Información de copyright

Información de la marca comercial