Guías de herramientas y aplicaciones

How to enable StorageGRID in your environment

NetApp
April 26, 2024

Tabla de contenidos

Guías de herramientas y aplicaciones

Utilice el conector Hadoop S3A de Cloudera con StorageGRID

Hadoop ha sido el favorito de los científicos de datos desde hace algún tiempo. Hadoop permite el procesamiento distribuido de grandes conjuntos de datos a través de clusters de ordenadores utilizando marcos de programación simples. Hadoop se diseñó para escalar verticalmente de un solo servidor a miles de máquinas, mientras que cada máquina poseía almacenamiento e informática local.

¿Por qué utilizar S3A para los flujos de trabajo de Hadoop?

A medida que el volumen de datos crece con el tiempo, el método de añadir nuevos equipos con sus propios recursos informáticos y de almacenamiento resulta ineficiente. Escalar de forma lineal crea retos para usar los recursos de forma eficiente y gestionar la infraestructura.

Para abordar estos retos, el cliente Hadoop S3A ofrece I/o de alto rendimiento frente al almacenamiento de objetos S3. Implementar un flujo de trabajo de Hadoop con S3A le ayuda a aprovechar el almacenamiento de objetos como repositorio de datos y le permite separar los recursos informáticos y de almacenamiento, lo que, a su vez, le permite escalar la computación y el almacenamiento de forma independiente. La disociación de la computación y el almacenamiento también le permite dedicar la cantidad adecuada de recursos a sus tareas informáticas y proporcionar capacidad en función del tamaño del conjunto de datos. Por lo tanto, es posible reducir el TCO general para los flujos de trabajo de Hadoop.

Configurar el conector S3A para usar StorageGRID

Requisitos previos

- Una URL de extremo de StorageGRID S3, una clave de acceso de inquilino s3 y una clave secreta para las pruebas de conexión Hadoop S3A.
- Un clúster Cloudera y permiso root o sudo para cada host del clúster para instalar el paquete Java.

A partir de abril de 2022, se había probado Java 11.0.14 con Cloudera 7.1.7 frente a StorageGRID 11.5 y 11.6. Sin embargo, el número de versión de Java puede ser diferente en el momento de una instalación nueva.

Instale el paquete Java

1. Compruebe la "Matriz de compatibilidad de Cloudera" Para la versión de JDK compatible.
2. Descargue el "Paquete Java 11.x." Que coincidan con el sistema operativo del clúster Cloudera. Copie este paquete en cada host del clúster. En este ejemplo, el paquete rpm se utiliza para CentOS.
3. Inicie sesión en cada host como raíz o utilice una cuenta con permiso sudo. Realice los siguientes pasos en cada host:
   a. Instale el paquete:

```
$ sudo rpm -Uvh jdk-11.0.14_linux-x64_bin.rpm
```

b. Compruebe dónde está instalado Java. Si se instalan varias versiones, establezca la versión recién instalada como predeterminada:

```
alternatives --config java

There are 2 programs which provide 'java'.

  Selection    Command
-----------------------------------------------
 +1           /usr/java/jre1.8.0_291-amd64/bin/java
  2           /usr/java/jdk-11.0.14/bin/java

Enter to keep the current selection[+], or type selection number: 2
```

c. Agregue esta línea al final de `/etc/profile`. La ruta debe coincidir con la ruta de la selección anterior:

```
export JAVA_HOME=/usr/java/jdk-11.0.14
```

d. Ejecute el siguiente comando para que el perfil surta efecto:

```
source /etc/profile
```

Configuración de Cloudera HDFS S3A

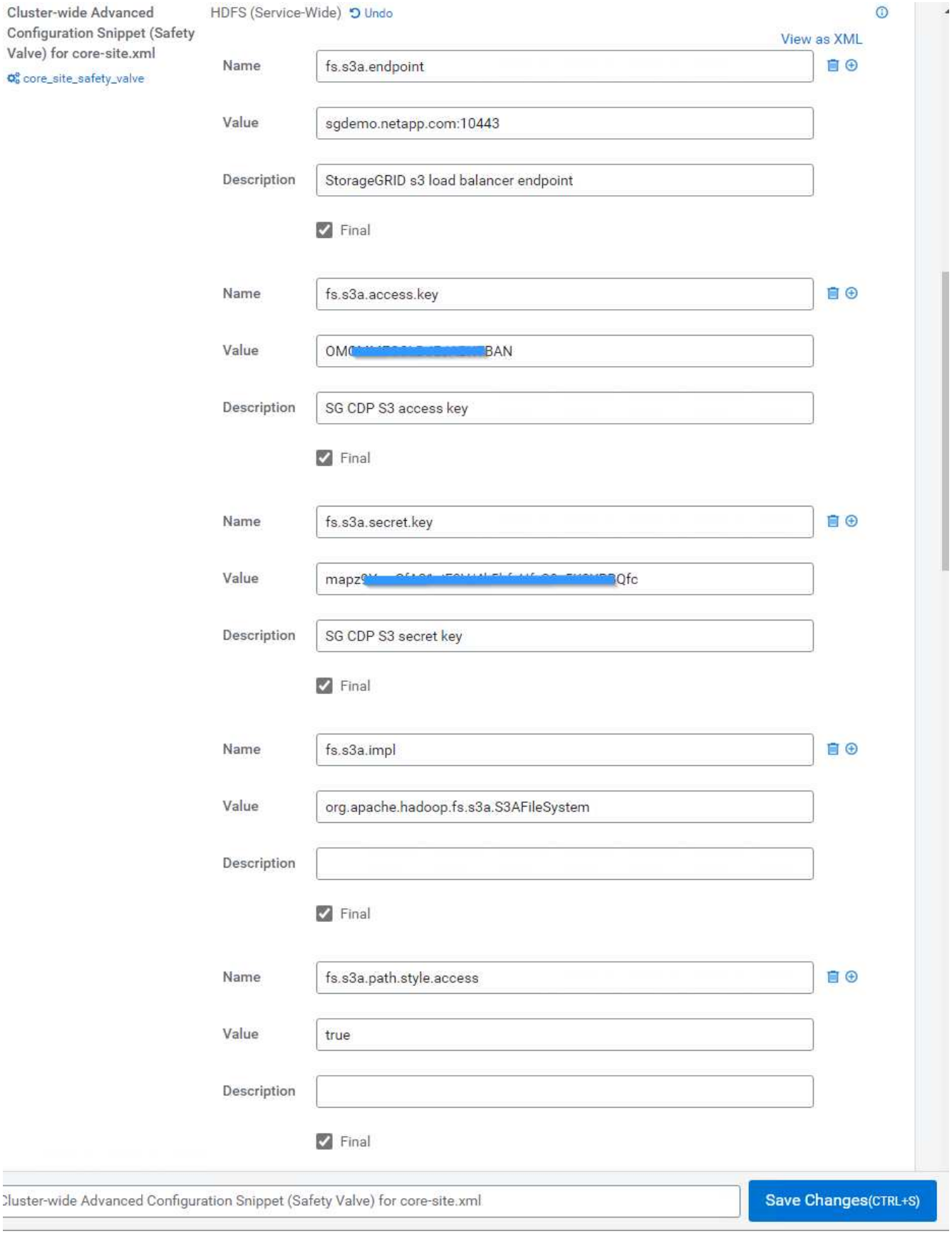
- Pasos*

1. En la interfaz gráfica de usuario de Cloudera Manager, seleccione Clusters > HDFS y seleccione Configuration.
2. En CATEGORÍA, seleccione Avanzado y desplácese hacia abajo para localizar `Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml`.
3. Haga clic en el signo (+) y agregue los siguientes pares de valor.

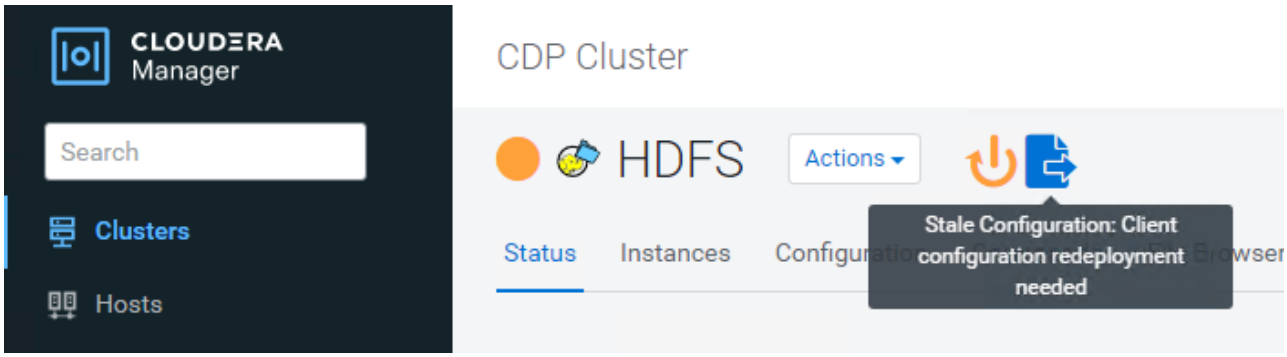
Nombre	Valor
fs.s3a.access.key	*<clave de acceso de inquilino s3 de StorageGRID>*
fs.s3a.secret.key	*<clave secreta de inquilino s3 de StorageGRID>*
fs.s3a.connection.ssl.enabled	[verdadero o falso] (el valor predeterminado es https si falta esta entrada)
fs.s3a.endpoint	*<StorageGRID S3 endpoint:Port>*

Nombre	Valor
fs.s3a.impl	Org.apache.hadoop.fs.s3a.S3AFileSystem
fs.s3a.path.style.access	[verdadero o falso] (el estilo de host virtual es el predeterminado si falta esta entrada)

Captura de pantalla de ejemplo



1. Haga clic en el botón Save Changes. Seleccione el icono Configuración obsoleta en la barra de menús de HDFS, seleccione Reiniciar servicios obsoletos en la página siguiente y seleccione Reiniciar ahora.



Probar la conexión S3A a StorageGRID

Realizar una prueba de conexión básica

Inicie sesión en uno de los hosts del clúster Cloudera y escriba `hadoop fs -ls s3a://<bucket-name>/`.

En el siguiente ejemplo se utiliza el sistema de ruta con un cubo de prueba hdfs preexistente y un objeto de prueba.

```
[root@ce-n1 ~]# hadoop fs -ls s3a://hdfs-test/
22/02/15 18:24:37 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/02/15 18:24:37 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/02/15 18:24:37 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/02/15 18:24:37 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
Found 1 items
-rw-rw-rw-   1 root root       1679 2022-02-14 16:03 s3a://hdfs-test/test
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/02/15 18:24:38 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
```

Resolución de problemas

Situación 1

Utilice una conexión HTTPS a StorageGRID y obtenga una `handshake_failure` error tras un tiempo de espera de 15 minutos.

Razón: Versión antigua de JRE/JDK utilizando un conjunto de cifrado TLS obsoleto o no compatible para la conexión con StorageGRID.

Mensaje de error de muestra

```
[root@ce-n1 ~]# hadoop fs -ls s3a://hdfs-test/
22/02/15 18:52:34 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/02/15 18:52:34 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/02/15 18:52:34 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/02/15 18:52:35 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/02/15 19:04:51 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
22/02/15 19:04:51 WARN fs.FileSystem: Failed to initialize fileystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClientIOException:
doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: Received fatal alert: handshake_failure: Unable to
execute HTTP request: Received fatal alert: handshake_failure
ls: doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: Received fatal alert: handshake_failure: Unable to
execute HTTP request: Received fatal alert: handshake_failure
```

Resolución: Asegúrese de que JDK 11.x o posterior esté instalado y establecido en la biblioteca Java predeterminada. Consulte la Instale el paquete Java para obtener más información.

Situación 2:

Error al conectarse a StorageGRID con mensaje de error `Unable to find valid certification path to requested target`.

Razón: el programa Java no confía en el certificado del servidor de extremo StorageGRID S3.

Mensaje de error de muestra:

```
[root@hdp6 ~]# hadoop fs -ls s3a://hdfs-test/
22/03/11 20:58:12 WARN impl.MetricsConfig: Cannot locate configuration:
tried hadoop-metrics2-s3a-file-system.properties,hadoop-
metrics2.properties
22/03/11 20:58:13 INFO impl.MetricsSystemImpl: Scheduled Metric snapshot
period at 10 second(s).
22/03/11 20:58:13 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system started
22/03/11 20:58:13 INFO Configuration.deprecation: No unit for
fs.s3a.connection.request.timeout(0) assuming SECONDS
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: Stopping s3a-file-system
metrics system...
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system stopped.
22/03/11 21:12:25 INFO impl.MetricsSystemImpl: s3a-file-system metrics
system shutdown complete.
22/03/11 21:12:25 WARN fs.FileSystem: Failed to initialize fileystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClientIOException:
doesBucketExistV2 on hdfs: com.amazonaws.SdkClientException: Unable to
execute HTTP request: PKIX path building failed:
sun.security.provider.certpath.SunCertPathBuilderException: unable to find
valid certification path to requested target: Unable to execute HTTP
request: PKIX path building failed:
sun.security.provider.certpath.SunCertPathBuilderException: unable to find
valid certification path to requested target
```

Resolución: NetApp recomienda el uso de un certificado de servidor emitido por una autoridad pública de firma de certificación conocida para garantizar la seguridad de la autenticación. También puede agregar un certificado de servidor o CA personalizado al almacén de confianza de Java.

Complete los siguientes pasos para agregar una CA personalizada de StorageGRID o un certificado de servidor al almacén de confianza de Java.

1. Realice una copia de seguridad del archivo Cacits de Java predeterminado existente.

```
cp -ap $JAVA_HOME/lib/security/cacerts
$JAVA_HOME/lib/security/cacerts.orig
```

2. Importe el certificado de extremo de StorageGRID S3 al almacén de confianza de Java.

```
keytool -import -trustcacerts -keystore $JAVA_HOME/lib/security/cacerts
-storepass changeit -noprompt -alias sg-lb -file <StorageGRID CA or
server cert in pem format>
```

Consejos para la solución de problemas

1. Aumente el nivel de registro de hadoop para DEPURAR.

   ```
   export HADOOP_ROOT_LOGGER=hadoop.root.logger=DEBUG,console
   ```

2. Ejecute el comando y dirija los mensajes del registro a error.log.

   ```
   hadoop fs -ls s3a://<bucket-name>/ &>error.log
   ```

Por Angela Cheng

Use S3cmd para probar y demostrar el acceso S3 en StorageGRID

S3cmd es una herramienta de línea de comandos gratuita y un cliente para las operaciones S3. Puede usar s3cmd para probar y demostrar el acceso s3 en StorageGRID.

Instale y configure S3cmd

Para instalar S3cmd en una estación de trabajo o servidor, descárguelo desde "Línea de comandos del cliente S3". S3cmd está preinstalado en cada nodo StorageGRID como herramienta que ayuda a resolver problemas.

Pasos de configuración inicial

1. s3cmd --configure
2. Proporcione solo access_key y Secret_Key, para que el resto conserve los valores predeterminados.
3. ¿Probar el acceso con las credenciales proporcionadas? [Y/n]: n (omitir la prueba ya que fallará)
4. ¿Desea guardar la configuración? [S/N] y.
   a. Configuración guardada en '/root/.s3cfg'
5. En .s3cfg, haga que los campos host_base y host_bucket estén vacíos después del signo "=" :
   a. host_base =
   b. host_bucket =

> Si especifica host_base y host_bucket en el paso 4, no es necesario especificar un extremo con --host en la CLI. Ejemplo:

```
host_base = 192.168.1.91:8082
host_bucket = bucketX.192.168.1.91:8082
s3cmd ls s3://bucketX --no-check-certificate
```

Ejemplos de comandos básicos

- * Crear un cubo:*

```
s3cmd mb s3://s3cmdbucket --host=<endpoint>:<port> --no-check-certificate
```

- **Lista todos los cucharones:**

```
s3cmd ls --host=<endpoint>:<port> --no-check-certificate
```

- **Lista todos los cucharones y su contenido:**

```
s3cmd la --host=<endpoint>:<port> --no-check-certificate
```

- **Lista objetos en un cubo específico:**

```
s3cmd ls s3://<bucket> --host=<endpoint>:<port> --no-check-certificate
```

- **Eliminar un cucharón:**

```
s3cmd rb s3://s3cmdbucket --host=<endpoint>:<port> --no-check-certificate
```

- * Poner un objeto:*

```
s3cmd put <file> s3://<bucket> --host=<endpoint>:<port> --no-check-certificate
```

- **Obtener un objeto:**

```
s3cmd get s3://<bucket>/<object> <file> --host=<endpoint>:<port> --no-check
-certificate
```

- **Eliminar un objeto:**

```
s3cmd del s3://<bucket>/<object> --host=<endpoint>:<port> --no-check
-certificate
```

Por Aron Klein

Vertica Eon mode Database utilizando NetApp StorageGRID como almacenamiento comunitario

En esta guía se describe el procedimiento para crear una base de datos de Vertica Eon Mode con almacenamiento compartido en StorageGRID de NetApp.

Introducción

Vertica es un software de gestión de bases de datos analíticas. Se trata de una plataforma de almacenamiento en columnas diseñada para gestionar grandes volúmenes de datos, lo que permite un rendimiento de consultas muy rápido en una situación tradicionalmente intensiva. Una base de datos Vertica se ejecuta en uno de los dos modos: Eon o Enterprise. Puede poner en marcha ambos modos en las instalaciones o en el cloud.

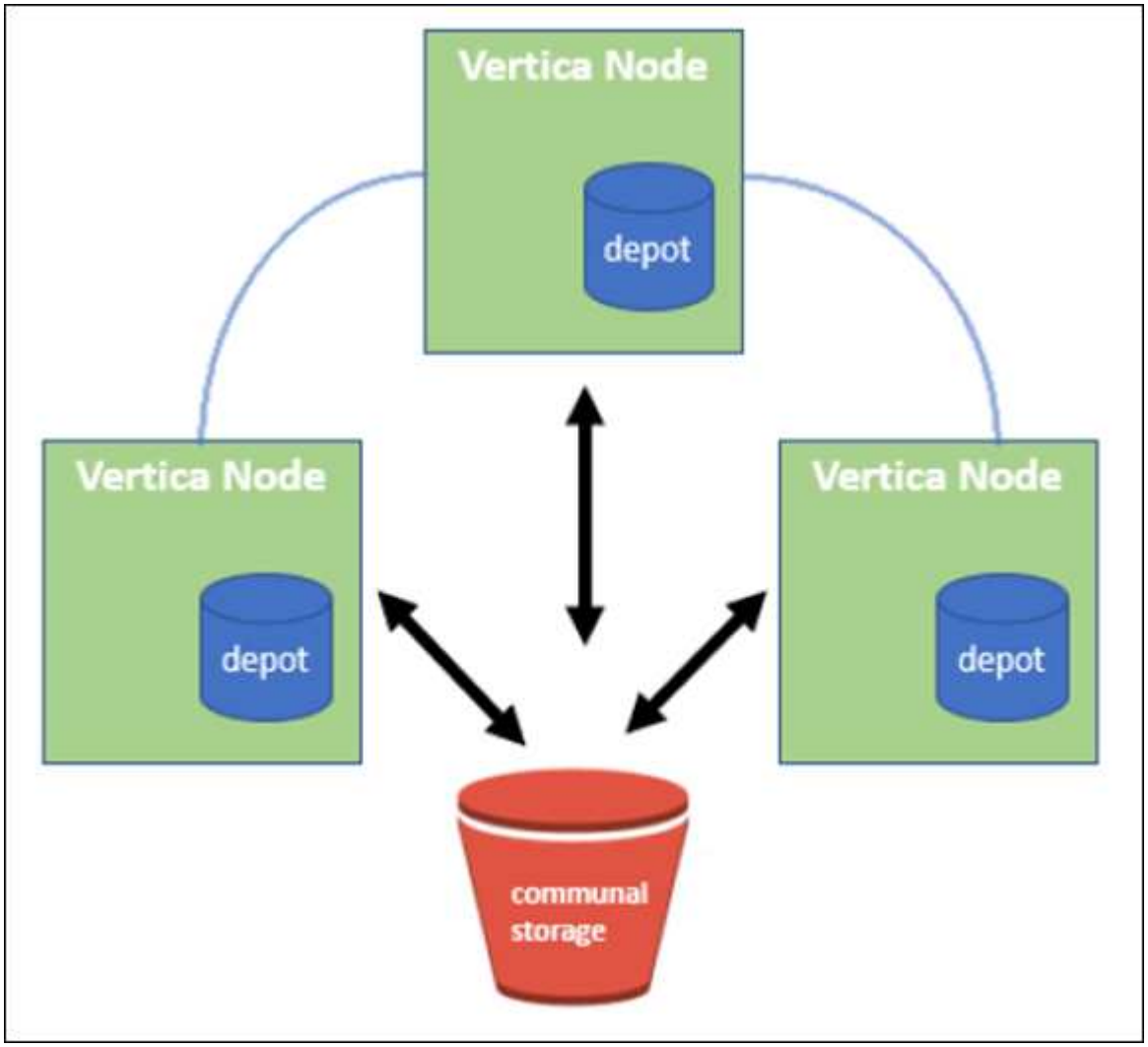
Los modos Eon y Enterprise difieren principalmente en el lugar en el que almacenan datos:

- Las bases de datos de modo Eon utilizan almacenamiento común para sus datos. Esto es recomendado por Vertica.
- Las bases de datos de Enterprise Mode almacenan los datos localmente en el sistema de archivos de nodos que componen la base de datos.

Arquitectura de modo Eon

Eon Mode separa los recursos computacionales de la capa de almacenamiento comunal de la base de datos, que permite que la computación y el almacenamiento se escalen por separado. Vertica en modo Eon está optimizada para trabajar con cargas de trabajo variables y aislarlas entre sí mediante el uso de recursos informáticos y de almacenamiento independientes.

Eon Mode almacena datos en un almacén de objetos compartidos denominado almacenamiento comunitario, un bucket de S3, que está alojado en las instalaciones o en Amazon S3.



Almacenamiento comunitario

En lugar de almacenar datos localmente, Eon Mode utiliza una única ubicación de almacenamiento comunal para todos los datos y el catálogo (metadatos). El almacenamiento comunitario es la ubicación de almacenamiento centralizada de la base de datos, compartida entre los nodos de la base de datos.

El almacenamiento comunitario tiene las siguientes propiedades:

- El almacenamiento común en el cloud o el almacenamiento de objetos en las instalaciones es más resiliente y menos susceptible a la pérdida de datos debido a fallos del almacenamiento que el almacenamiento en disco en máquinas individuales.
- Cualquier dato puede ser leído por cualquier nodo con la misma ruta.
- La capacidad no está limitada por el espacio en disco de los nodos.
- Debido a que los datos se almacenan de una manera comunitaria, puede ampliar el clúster con flexibilidad para satisfacer las demandas cambiantes. Si los datos estuvieran almacenados localmente en los nodos, añadir o quitar nodos requeriría mover cantidades considerables de datos entre nodos o bien fuera de los nodos que se van a quitar o a los nodos recién creados.

El almacén

Un inconveniente del almacenamiento común es su velocidad. Tener acceso a los datos desde una ubicación en cloud compartida es más lento que leerlos desde el disco local. Asimismo, la conexión al almacenamiento comunitario puede convertirse en un cuello de botella si muchos nodos leen datos de él a la vez. Para aumentar la velocidad del acceso a los datos, los nodos de una base de datos de Eon Mode mantienen una caché de datos de disco local denominada almacén. Al ejecutar una consulta, los nodos comprueban en primer lugar si los datos que necesitan están en el almacén. Si es así, termina la consulta utilizando la copia local de los datos. Si los datos no están en el almacén, el nodo recupera los datos del almacenamiento comunitario y guarda una copia en el almacén.

Recomendaciones de StorageGRID de NetApp

Vertica almacena datos de bases de datos en el almacenamiento de objetos como miles (o millones) de objetos comprimidos (el tamaño observado es de 200 a 500 MB por objeto. Cuando un usuario ejecuta consultas de base de datos, Vertica recupera el intervalo de datos seleccionado de estos objetos comprimidos en paralelo mediante la llamada GET del intervalo de bytes. Cada BYTE-range GET es aproximadamente de 8 KB.

Durante la prueba de 10 TB del almacén de bases de datos fuera de las consultas de los usuarios, se enviaron a la cuadrícula entre 4,000 y 10,000 solicitudes GET (byte-range GET) por segundo. Cuando se ejecuta esta prueba con dispositivos SG6060, aunque el porcentaje de uso de CPU por nodo del dispositivo es bajo (entre el 20 % y el 30 %), 2/3 veces el tiempo de la CPU espera para I/O. Se observa un porcentaje muy pequeño (0% a 0.5%) de espera de E/S en el SGF6024.

Debido a la alta demanda de IOPS pequeñas con requisitos de latencia muy bajos (la media debe ser inferior a 0.01 segundos), NetApp recomienda utilizar SFG6024 para los servicios de almacenamiento de objetos. Si el SG6060 se necesita para tamaños de base de datos muy grandes, el cliente debe trabajar con el equipo de cuentas de Vertica en el ajuste de tamaño del almacén para admitir el conjunto de datos que se consulta activamente.

Para el nodo Admin y el nodo API Gateway, el cliente puede utilizar el SG100 o SG1000. La elección depende del número de solicitudes de consulta de los usuarios en paralelo y del tamaño de la base de datos. Si el cliente prefiere utilizar un equilibrador de carga de terceros, NetApp recomienda un equilibrador de carga dedicado para la carga de trabajo con demanda de alto rendimiento. Para obtener información sobre cómo ajustar el tamaño de StorageGRID, consulte al equipo de cuentas de NetApp.

Otras recomendaciones de configuración de StorageGRID incluyen:

- **Topología de cuadrícula**. No mezcle el SGF6024 con otros modelos de dispositivos de almacenamiento en el mismo sitio de la red. Si prefiere utilizar el SG6060 para la protección de archivado a largo plazo,

mantenga el SGF6024 con un equilibrador de carga de grid dedicado en su propio sitio de grid (sitio físico o lógico) para una base de datos activa con el fin de mejorar el rendimiento. La mezcla de diferentes modelos de dispositivo en el mismo sitio reduce el rendimiento general in situ.

- **Protección de datos**. Use copias replicadas para proteger. No utilice código de borrado para una base de datos activa. El cliente puede utilizar códigos de borrado para proteger bases de datos inactivas a largo plazo.
- **No active la compresión de red**. Vertica comprime objetos antes de almacenarlos en almacenamiento de objetos. Habilitar la compresión de grid no ahorra aún más el uso del almacenamiento y reduce de forma significativa el rendimiento A nivel de byte.
- **HTTP frente a HTTPS S3 Endpoint connection**. Durante las pruebas de prueba de rendimiento, observamos una mejora del rendimiento del 5% al utilizar una conexión HTTP S3 del clúster Vertica al extremo del equilibrador de carga de StorageGRID. Esta opción debe basarse en los requisitos de seguridad del cliente.

Las recomendaciones para una configuración de Vertica incluyen:

- **La configuración predeterminada del almacén de la base de datos Vertica está activada (valor = 1) para las operaciones de lectura y escritura**. NetApp recomienda mantener la configuración de estos almacenes habilitada para mejorar el rendimiento.
- **Desactive las limitaciones de streaming**. Para obtener información detallada sobre la configuración, consulte la sección Desactivación de las limitaciones de la transmisión.

Instalación del modo Eon en las instalaciones con almacenamiento común en StorageGRID

En las siguientes secciones se describe el procedimiento para instalar el modo Eon en las instalaciones con almacenamiento común en StorageGRID. El procedimiento para configurar un almacenamiento de objetos compatible con simple Storage Service (S3) en las instalaciones es similar al procedimiento descrito en la guía de Vertica, "Instale una base de datos Eon Mode en las instalaciones".

Para la prueba funcional se utilizó la siguiente configuración:

- StorageGRID 11.4.0.4
- Vertica 10.1.0
- Tres máquinas virtuales (VM) con CentOS 7.x OS para nodos Vertica y formar un clúster. Esta configuración es sólo para la prueba funcional, no para el clúster de base de datos de producción Vertica.

Estos tres nodos se configuran con una clave Secure Shell (SSH) para permitir a SSH sin una contraseña entre los nodos del clúster.

Información necesaria de StorageGRID de NetApp

Para instalar Eon Mode en las instalaciones con almacenamiento comunitario en StorageGRID, debe tener la siguiente información de requisitos previos.

- La dirección IP o el nombre de dominio completo (FQDN) y el número de puerto del extremo de StorageGRID S3. Si utiliza HTTPS, utilice una entidad de certificación (CA) personalizada o un certificado SSL autofirmado implementado en el extremo de StorageGRID S3.
- Nombre del bloque. Debe existir previamente y estar vacío.
- El ID de clave de acceso y la clave de acceso secreta con acceso de lectura y escritura al bloque.

Creación de un archivo de autorización para acceder al extremo de S3

Los siguientes requisitos previos se aplican al crear un archivo de autorización para acceder al extremo de S3:

- Vertica está instalada.
- Un clúster está configurado, configurado y listo para la creación de bases de datos.

Para crear un archivo de autorización para acceder al extremo de S3, siga estos pasos:

1. Inicie sesión en el nodo Vertica donde se ejecutará `admintools` Para crear la base de datos Eon Mode.

   El usuario predeterminado es `dbadmin`, Creado durante la instalación del clúster Vertica.

2. Utilice un editor de texto para crear un archivo en la `/home/dbadmin` directorio. El nombre del archivo puede ser cualquier cosa que desee, por ejemplo, `sg_auth.conf`.
3. Si el extremo de S3 utiliza un puerto HTTP 80 o un puerto HTTPS 443 estándar, omita el número del puerto. Para utilizar HTTPS, configure los siguientes valores:

   ◦ `awsenablehttps = 1`, de lo contrario, establezca el valor en `0`.

   ◦ `awsauth = <s3 access key ID>:<secret access key>`

   ◦ `awsendpoint = <StorageGRID s3 endpoint>:<port>`

   Para usar una CA personalizada o un certificado SSL autofirmado para la conexión HTTPS de extremo StorageGRID S3, especifique la ruta de archivo completa y el nombre de archivo del certificado. Este archivo debe estar en la misma ubicación de cada nodo Vertica y tener permiso de lectura para todos los usuarios. Omita este paso si la CA conocida públicamente firma del certificado SSL de extremo de StorageGRID S3.

   - `awscafile = <filepath/filename>`

   Por ejemplo, consulte el siguiente archivo de ejemplo:

```
awsauth = MNVU4OYFAY2xyz123:03vuO4M4KmdfwffT8nqnBmnMVTr78Gu9wANabcxyz
awsendpoint = s3.england.connectlab.io:10443
awsenablehttps = 1
awscafile = /etc/custom-cert/grid.pem
```

+

> En un entorno de producción, el cliente debe implementar un certificado de servidor firmado por una CA conocida públicamente en un extremo de equilibrador de carga de StorageGRID S3.

Elegir una ruta de almacén en todos los nodos de Vertica

Seleccione o cree un directorio en cada nodo para la ruta de almacenamiento del almacén. El directorio que suministre para el parámetro de ruta de almacenamiento del almacén debe tener lo siguiente:

- La misma ruta en todos los nodos del clúster (por ejemplo, `/home/dbadmin/depot`)

- El usuario dbadmin puede leer y escribir
- Almacenamiento suficiente

  De forma predeterminada, Vertica utiliza el 60% del espacio del sistema de archivos que contiene el directorio para el almacenamiento del almacén. Puede limitar el tamaño del almacén mediante el `--depot-size` en el `create_db` comando. Consulte "Ajuste de tamaño del clúster Vertica para una base de datos en modo Eon" artículo para las pautas generales de ajuste de tamaño de Vertica o consulte con su gestor de cuentas de Vertica.

  La `admintools create_db` la herramienta intenta crear la ruta del almacén para usted si no existe.

Creación de la base de datos Eon en las instalaciones

Para crear la base de datos Eon en las instalaciones, siga estos pasos:

1. Para crear la base de datos, utilice `admintools create_db` herramienta.

   La siguiente lista proporciona una breve explicación de los argumentos utilizados en este ejemplo. Consulte el documento Vertica para obtener una explicación detallada de todos los argumentos necesarios y opcionales.

   - -x <path/filename of authorization file created in “Creación de un archivo de autorización para acceder al extremo de S3” >.

     Los detalles de autorización se almacenan dentro de la base de datos después de haber creado correctamente. Puede eliminar este archivo para evitar exponer la clave secreta de S3.

   - --comunal-almacenamiento-ubicación <s3://storagegrid bucketname>
   - -S <comma-separated list of Vertica nodes to be used for this database>
   - -d <name of database to be created>
   - -p <password to be set for this new database>. Por ejemplo, consulte el siguiente comando de ejemplo:

     ```
     admintools -t create_db -x sg_auth.conf --communal-storage
     -location=s3://vertica --depot-path=/home/dbadmin/depot --shard
     -count=6 -s vertica-vm1,vertica-vm2,vertica-vm3 -d vmart -p
     '<password>'
     ```

     La creación de una nueva base de datos tarda varios minutos en función del número de nodos de la base de datos. Al crear la base de datos por primera vez, se le solicitará que acepte el Contrato de licencia.

Por ejemplo, consulte el siguiente archivo de autorización de ejemplo y. `create db` comando:

```
[dbadmin@vertica-vm1 ~]$ cat sg_auth.conf
awsauth = MNVU4OYFAY2CPKVXVxxxx:03vuO4M4KmdfwffT8nqnBmnMVTr78Gu9wAN+xxxx
awsendpoint = s3.england.connectlab.io:10445
awsenablehttps = 1
```

```
[dbadmin@vertica-vm1 ~]$ admintools -t create_db -x sg_auth.conf
--communal-storage-location=s3://vertica --depot-path=/home/dbadmin/depot
--shard-count=6 -s vertica-vm1,vertica-vm2,vertica-vm3 -d vmart -p
'xxxxxxxx'
Default depot size in use
Distributing changes to cluster.
    Creating database vmart
    Starting bootstrap node v_vmart_node0007 (10.45.74.19)
    Starting nodes:
        v_vmart_node0007 (10.45.74.19)
    Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (UP)
    Creating database nodes
    Creating node v_vmart_node0008 (host 10.45.74.29)
    Creating node v_vmart_node0009 (host 10.45.74.39)
    Generating new configuration information
    Stopping single node db before adding additional nodes.
    Database shutdown complete
    Starting all nodes
Start hosts = ['10.45.74.19', '10.45.74.29', '10.45.74.39']
    Starting nodes:
        v_vmart_node0007 (10.45.74.19)
        v_vmart_node0008 (10.45.74.29)
        v_vmart_node0009 (10.45.74.39)
    Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (UP) v_vmart_node0008: (UP)
v_vmart_node0009: (UP)
Creating depot locations for 3 nodes
Communal storage detected: rebalancing shards

Waiting for rebalance shards. We will wait for at most 36000 seconds.
Installing AWS package
```

```
    Success: package AWS installed
Installing ComplexTypes package
    Success: package ComplexTypes installed
Installing MachineLearning package
    Success: package MachineLearning installed
Installing ParquetExport package
    Success: package ParquetExport installed
Installing VFunctions package
    Success: package VFunctions installed
Installing approximate package
    Success: package approximate installed
Installing flextable package
    Success: package flextable installed
Installing kafka package
    Success: package kafka installed
Installing logsearch package
    Success: package logsearch installed
Installing place package
    Success: package place installed
Installing txtindex package
    Success: package txtindex installed
Installing voltagesecure package
    Success: package voltagesecure installed
Syncing catalog on vmart with 2000 attempts.
Database creation SQL tasks completed successfully. Database vmart created
successfully.
```

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
61	s3://vertica/051/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07_0_0.dfs
145	s3://vertica/2c4/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d_0_0.dfs
146	s3://vertica/33c/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d_0_0.dfs

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
40	s3://vertica/382/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31_0_0.dfs
145	s3://vertica/42f/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a21/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a21_0_0.dfs
34	s3://vertica/472/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a25/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a25_0_0.dfs
41	s3://vertica/476/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a2d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a2d_0_0.dfs
61	s3://vertica/52a/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a5d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a5d_0_0.dfs
131	s3://vertica/5d2/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a19/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a19_0_0.dfs
91	s3://vertica/5f7/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a11/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a11_0_0.dfs
118	s3://vertica/82d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a15/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a15_0_0.dfs
115	s3://vertica/9a2/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a61/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a61_0_0.dfs

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
33	s3://vertica/acd/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29_0_0.dfs
133	s3://vertica/b98/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d_0_0.dfs
38	s3://vertica/db3/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49_0_0.dfs
38	s3://vertica/eba/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59_0_0.dfs
21521920	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2.tar
6865408	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021602/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602.tar
204217344	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021610/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610.tar
16109056	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0.tar
12853248	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021800/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800.tar

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
8937984	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a.tar
56260608	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000218b2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2.tar
53947904	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000219ba/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba.tar
44932608	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000219de/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de.tar
256306688	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar
8062464	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar
20024832	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.tar
10444	s3://vertica/metadata/VMart/cluster_config.json
823266	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/chkpt_1.cat.gz

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
254	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/completed
2958	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/chkpt_1.cat.gz
231	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/completed
822521	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/chkpt_1.cat.gz
231	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/completed
746513	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g14.cat
2596	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_3_g3.cat.gz
821065	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_4_g4.cat.gz
6440	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_5_g5.cat
8518	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_8_g8.cat

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat

Desactivación de las limitaciones de la transmisión

Este procedimiento se basa en la guía de Vertica para otro almacenamiento de objetos en las instalaciones y debe ser aplicable a StorageGRID.

1. Después de crear la base de datos, desactive la `AWSStreamingConnectionPercentage` configuración del parámetro configurándolo como `0`. Esta configuración es innecesaria para una instalación local en modo Eon con almacenamiento común. Este parámetro de configuración controla el número de conexiones al almacén de objetos que Vertica utiliza para las lecturas en streaming. En un entorno cloud, esta configuración ayuda a evitar que la transmisión de datos del almacén de objetos utilice todos los identificadores de archivos disponibles. Deja algunos identificadores de archivos disponibles para otras operaciones de almacén de objetos. Debido a la baja latencia de los almacenes de objetos en las instalaciones, esta opción es innecesaria.
2. Utilice un `vsql` instrucción para actualizar el valor del parámetro. La contraseña es la contraseña de la base de datos que se establece en “creación de la base de datos Eon en las instalaciones”. Por ejemplo, consulte el siguiente resultado de muestra:

```
[dbadmin@vertica-vm1 ~]$ vsql
Password:
Welcome to vsql, the Vertica Analytic Database interactive terminal.
Type:   \h or \? for help with vsql commands
        \g or terminate with semicolon to execute query
        \q to quit
dbadmin=> ALTER DATABASE DEFAULT SET PARAMETER
AWSStreamingConnectionPercentage = 0; ALTER DATABASE
dbadmin=> \q
```

Verificación de la configuración del almacén

La configuración predeterminada del almacén de la base de datos Vertica está habilitada (valor = 1) para las operaciones de lectura y escritura. NetApp recomienda mantener la configuración de estos almacenes habilitada para mejorar el rendimiento.

```
vsql -c 'show current all;' | grep -i UseDepot
DATABASE | UseDepotForReads | 1
DATABASE | UseDepotForWrites | 1
```

Carga de datos de muestra (opcional)

Si esta base de datos se utiliza para realizar pruebas y se eliminará, puede cargar datos de ejemplo en esta base de datos para realizar pruebas. Vertica incluye un conjunto de datos de muestra, VMart, que se encuentra en `/opt/vertica/examples/VMart_Schema/` En cada nodo Vertica. Puede encontrar más información acerca de este conjunto de datos de ejemplo "aquí".

Siga estos pasos para cargar los datos de ejemplo:

1. Inicie sesión como dbadmin en uno de los nodos Vertica: cd /opt/vertica/examples/VMart_Schema/
2. Cargue los datos de ejemplo en la base de datos e introduzca la contraseña de la base de datos cuando

se le solicite en los subpasos c y d:

a. `cd /opt/vertica/examples/VMart_Schema`

b. `./vmart_gen`

c. `vsql < vmart_define_schema.sql`

d. `vsql < vmart_load_data.sql`

3. Hay varias consultas SQL predefinidas, puede ejecutar algunas de ellas para confirmar que los datos de prueba se han cargado correctamente en la base de datos. Por ejemplo: `vsql < vmart_queries1.sql`

Dónde encontrar información adicional

Si quiere más información sobre el contenido de este documento, consulte los siguientes documentos o sitios web:

- "Documentación de producto de NetApp StorageGRID 11,7"
- "Especificaciones técnicas de StorageGRID"
- "Documentación de producto de Vertica 10.1"

Historial de versiones

Versión	Fecha	Historial de versiones del documento
Versión 1.0	Septiembre de 2021	Versión inicial.

Por Angela Cheng

Análisis de registros de StorageGRID mediante pila ELK

Con la función syslog Forward de StorageGRID 11.6, es posible configurar un servidor de syslog externo para recopilar y analizar mensajes de registro de StorageGRID. ELK (Elasticsearch, Logstash, Kibana) se ha convertido en una de las soluciones de análisis de registros más populares. Vea el "Análisis de registros de StorageGRID mediante vídeo ELK" Para ver una configuración DE EJEMPLO DE ELK y cómo se puede utilizar para identificar y solucionar las solicitudes fallidas de S3. Este artículo proporciona archivos de ejemplo de configuración de Logstash, consultas de Kibana, gráficos y panel para ofrecerle un inicio rápido para la gestión de registros y análisis de StorageGRID.

Requisitos

- StorageGRID 11.6.0.2 o superior
- ELK (Elasticsearch, Logstash y Kibana) 7.1x o superior instalado y en funcionamiento

Archivos de ejemplo

- "Descargue el paquete de archivos de ejemplo Logstash 7.x." + **md5 checksum** 148c23d0021d9a4bb4a6c0287464deab + **sha256 checksum**

f51ec9e2e3f842d5a7861566b167a561b4373b4e7bb3c8b3b3b3b3b3b3b3b3b3b3b3b3b3b3b3b3b3b3b5

- "Descargue el paquete de archivos de ejemplo Logstash 8.x." + **md5 checksum** e11bae3a662f87c310ef363d0fe06835 + **sha256 checksum** 5c670755742cfdfd5aa7a596ba087e0153a65bcaef3934afdb682f6cd1278d

Suposición

Los lectores están familiarizados con la terminología y las operaciones de StorageGRID y ELK.

Instrucción

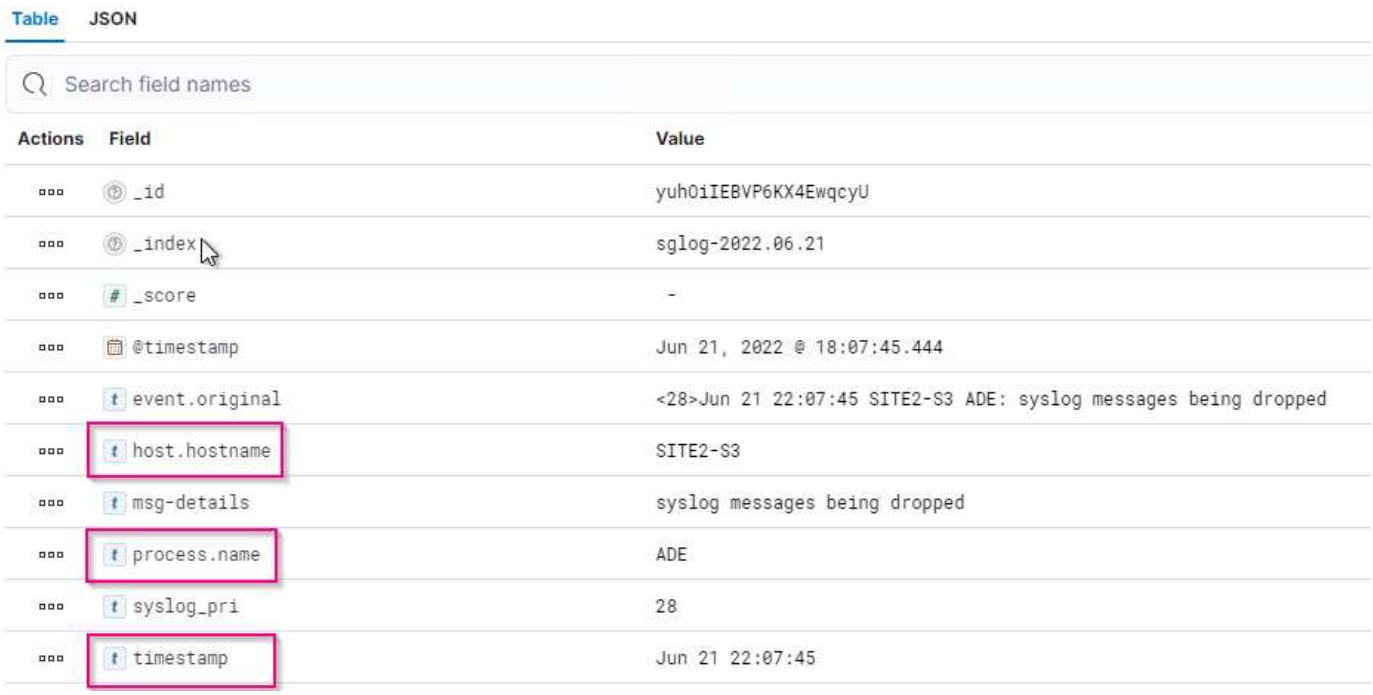
Se proporcionan dos versiones de ejemplo debido a diferencias en los nombres definidos por patrones de gruñido. + por ejemplo, el patrón de grok SYSLOGBASE en el archivo de configuración Logstash define los nombres de campo de forma diferente dependiendo de la versión de Logstash instalada.

```
match => {"message" => '<%{POSINT:syslog_pri}>%{SYSLOGBASE}
%{GREEDYDATA:msg-details}'}
```

Muestra de Logstash 7.17

Field	Value
t _id	7C1MaYEBRH8UbfKnIls8
t _index	sgrid2-2022.06.15
# _score	-
t _type	_doc
@timestamp	Jun 15, 2022 @ 17:36:46.038
t host	grid2-site2-s1
t logsource	SITE2-S1
t msg-details	Reloading syslog service
t pid	628
t program	update-sysl
t syslog_pri	37
t timestamp	Jun 15 21:36:46

Muestra de Logstash 8.23



- Pasos*

1. Descomprima el ejemplo proporcionado en función de la versión DE ELK instalada. + la carpeta de ejemplo incluye dos muestras de configuración de Logstash: + **sglog-2-file.conf:** este archivo de configuración envía mensajes de registro StorageGRID a un archivo en Logstash sin transformación de datos. Puede utilizar esta opción para confirmar que Logstash recibe mensajes de StorageGRID o para comprender los patrones de registro de StorageGRID. + **sglog-2-es.conf:** este archivo de configuración transforma los mensajes de registro StorageGRID utilizando varios patrones y filtros. Incluye ejemplos de sentencias DROP, que devuelven mensajes basados en patrones o filtros. La salida se envía a Elasticsearch para indizar. + Personalice el archivo de configuración seleccionado de acuerdo con la instrucción dentro del archivo.
2. Pruebe el archivo de configuración personalizado:

   ```
   /usr/share/logstash/bin/logstash --config.test_and_exit -f <config-file-path/file>
   ```

   Si la última línea devuelta es similar a la siguiente, el archivo de configuración no tiene errores de sintaxis:

   ```
   [LogStash::Runner] runner - Using config.test_and_exit mode. Config Validation Result: OK. Exiting Logstash
   ```

3. Copie el archivo conf personalizado en la configuración del servidor Logstash: /Etc/logstash/conf.d + Si no ha habilitado config.reload.automatic en /etc/logstash/logstash.yml, reinicie el servicio Logstash. De lo contrario, espere a que transcurra el intervalo de recarga de la configuración.

```
 grep reload /etc/logstash/logstash.yml
# Periodically check if the configuration has changed and reload the
pipeline
config.reload.automatic: true
config.reload.interval: 5s
```

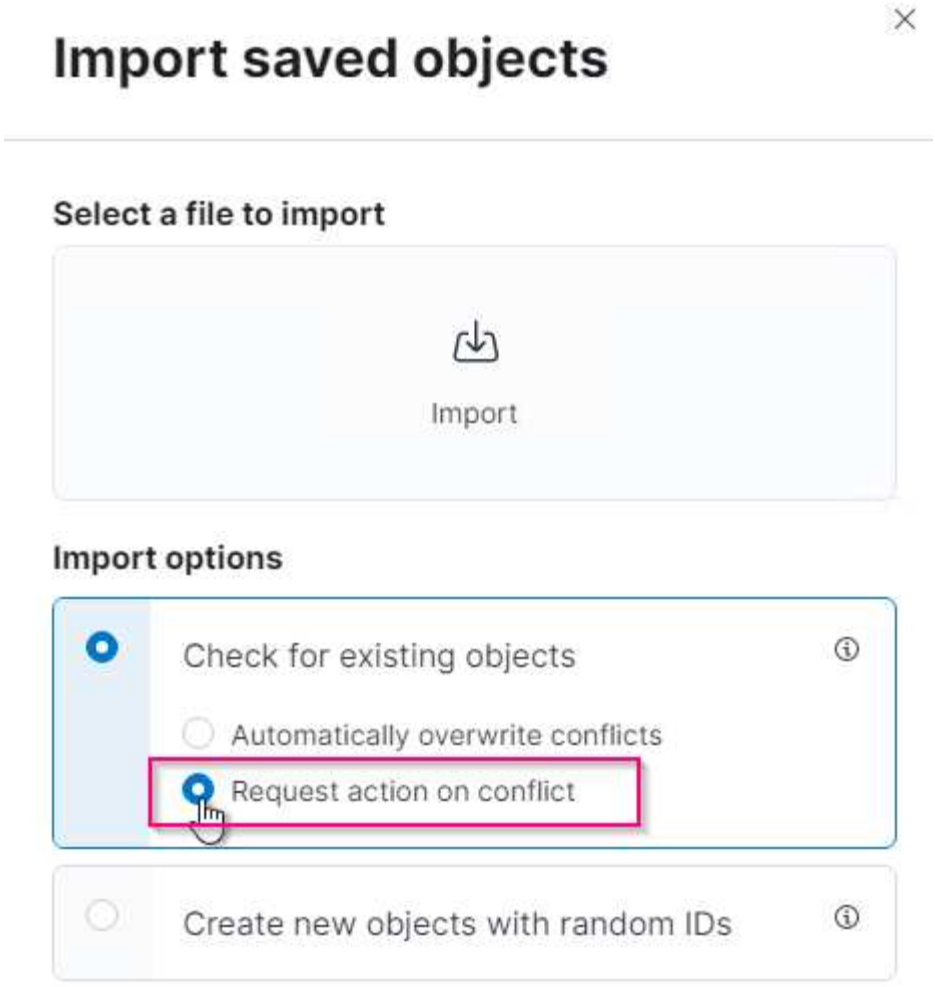
4. Compruebe /var/log/logstash/logstash-plain.log y confirme que no hay errores al iniciar Logstash con el nuevo archivo de configuración.
5. Confirme que el puerto TCP se ha iniciado y que está escuchando. + en este ejemplo, se utiliza el puerto TCP 5000.

```
netstat -ntpa | grep 5000
tcp6       0      0 :::5000                 :::*
LISTEN      25744/java
```

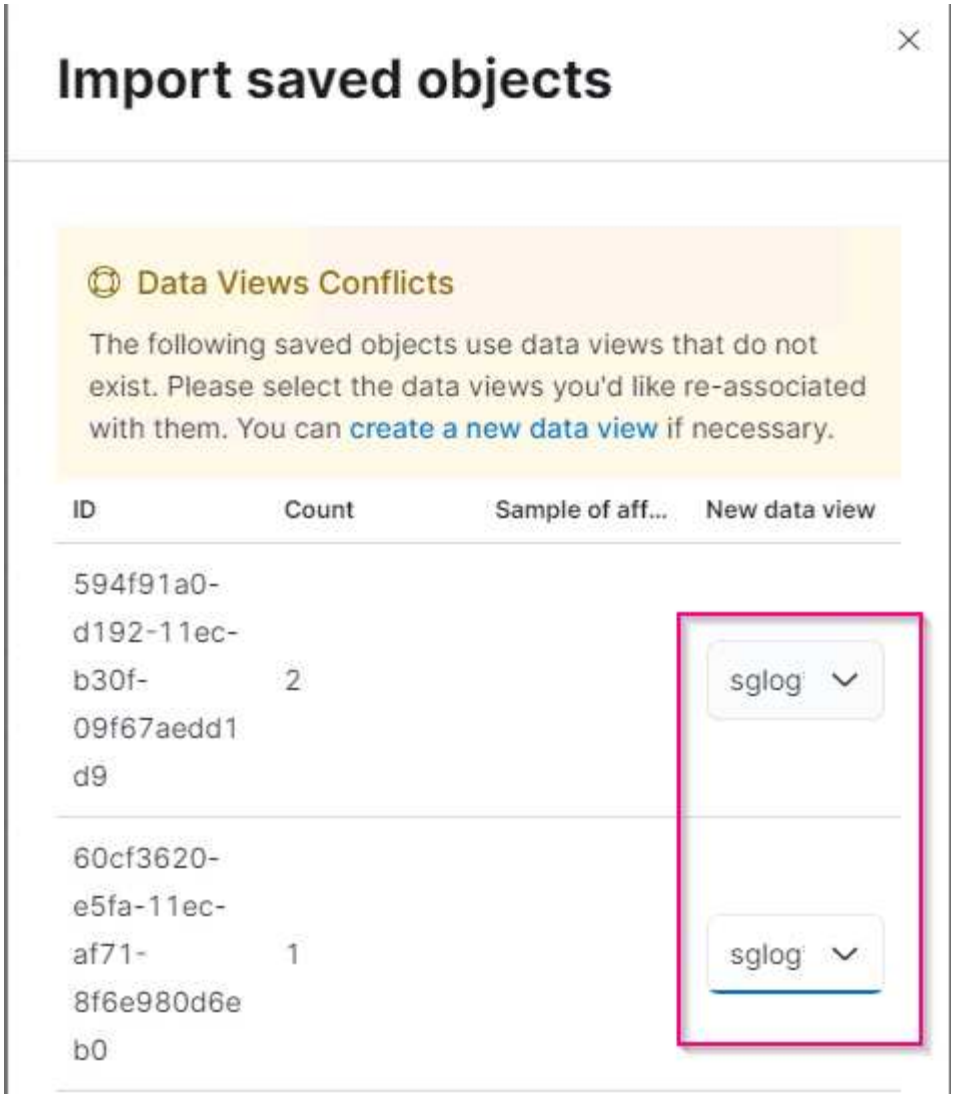
6. Desde la interfaz gráfica de usuario del administrador StorageGRID, configure el servidor de syslog externo para que envíe mensajes de registro a Logstash. Consulte la "vídeo de demostración" para obtener más detalles.
7. Debe configurar o deshabilitar el firewall en el servidor Logstash para permitir la conexión de nodos StorageGRID al puerto TCP definido.
8. En la interfaz gráfica de usuario de Kibana, seleccione Management → Dev Tools. En la página Console, ejecute este comando GET para confirmar que se han creado nuevos índices en Elasticsearch.

```
GET /_cat/indices/*?v=true&s=index
```

9. Desde la GUI de Kibana, cree un patrón de índice (ELK 7.x) o una vista de datos (ELK 8.x).
10. En la GUI de Kibana, escriba "objetos afeitados" en el cuadro de búsqueda que se encuentra en el centro superior. + en la página objetos guardados, seleccione Importar. En Opciones de importación, seleccione "solicitar acción en conflicto"



Importe eleLIMPORT <version>-QUERY-chart-sample.ndjson. + cuando se le solicite que resuelva el conflicto, seleccione el patrón de índice o la vista de datos que creó en el paso 8.



Se importan los siguientes objetos de Kibana: + **Consulta** + * audit-msg-s3rq-orlm + * mensajes relacionados con bycast log s3 + * advertencia de nivel de registro o superior + * evento de seguridad fallido + **Gráfico** + * las solicitudes s3 cuentan en función de bycast.log + * código de estado HTTP + * desglose de msg de auditoría por tipo + * respuesta media de s3 Tiempo + **Panel** + * panel de solicitudes S3 utilizando los gráficos anteriores.

Ahora está listo para realizar el análisis de registros de StorageGRID con Kibana.

Recursos adicionales

- "syslog101"
- "¿Qué es LA pila ELK"
- "Lista de patrones GROK"
- "Guía para principiantes de Logstash: GROK"
- "Una guía práctica para Logstash: Syslog profundizar"
- "Guía de Kibana: Explorar el documento"
- "Referencia de mensajes de registro de auditoría de StorageGRID"

Por Angela Cheng

Utilice Prometheus y Grafana para ampliar la retención de métricas

En este informe técnico, se proporcionan instrucciones detalladas para la configuración de NetApp StorageGRID 11.6 con servicios externos Prometheus y Grafana.

Introducción

StorageGRID almacena métricas mediante Prometheus y proporciona visualizaciones de estas métricas a través de los paneles de Grafana integrados. Se puede acceder a las métricas de Prometheus de manera segura desde StorageGRID mediante la configuración de certificados de acceso de clientes y la habilitación del acceso a prometheus para el cliente especificado. Hoy en día, la retención de estos datos métricos está limitada por la capacidad de almacenamiento del nodo de administración. Para obtener una mayor duración y una capacidad para crear visualizaciones personalizadas de estas métricas, implementaremos un nuevo servidor Prometheus y Grafana, configuraremos nuestro nuevo servidor para reunir las métricas de la instancia de StorageGRIDs y creamos un panel con las métricas que somos importantes para nosotros. Puede obtener más información sobre las métricas de Prometheus recopiladas en el "Documentación de StorageGRID".

Federar Prometheus

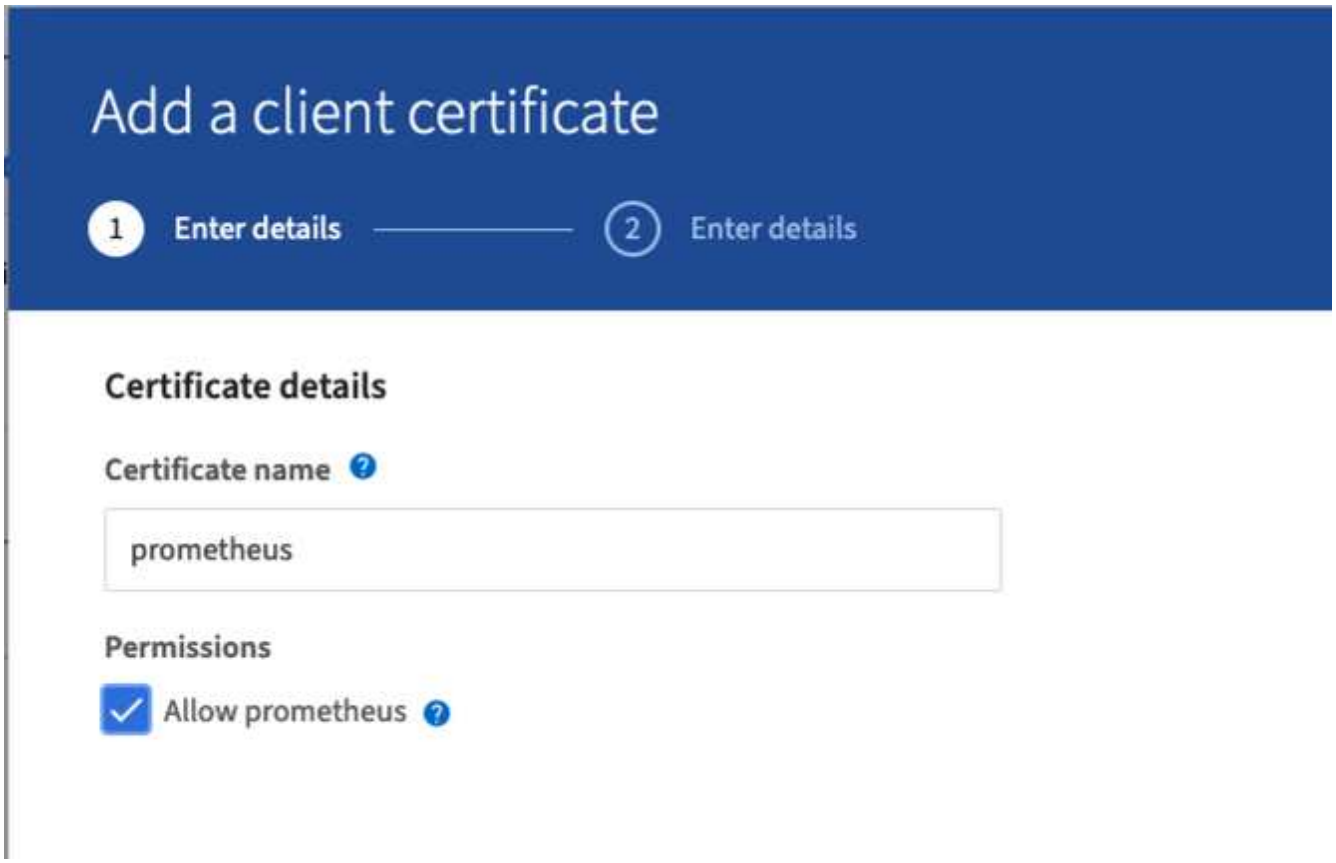
Detalles del laboratorio

A efectos de este ejemplo, usaré todas las máquinas virtuales para nodos StorageGRID 11.6 y un servidor Debian 11. La interfaz de gestión StorageGRID se configura con un certificado de CA de confianza pública. Este ejemplo no pasará por la instalación y configuración del sistema StorageGRID o la instalación de Debian linux. Puede utilizar cualquier sabor de Linux que desee, con el apoyo de Prometheus y Grafana. Tanto Prometheus como Grafana pueden instalar como contenedores Docker, crear a partir de código fuente o binarios precompilados. En este ejemplo, instalaré los binarios Prometheus y Grafana directamente en el mismo servidor Debian. Descargue y siga las instrucciones de instalación básica de https://prometheus.io y.. https://grafana.com/grafana/ respectivamente.

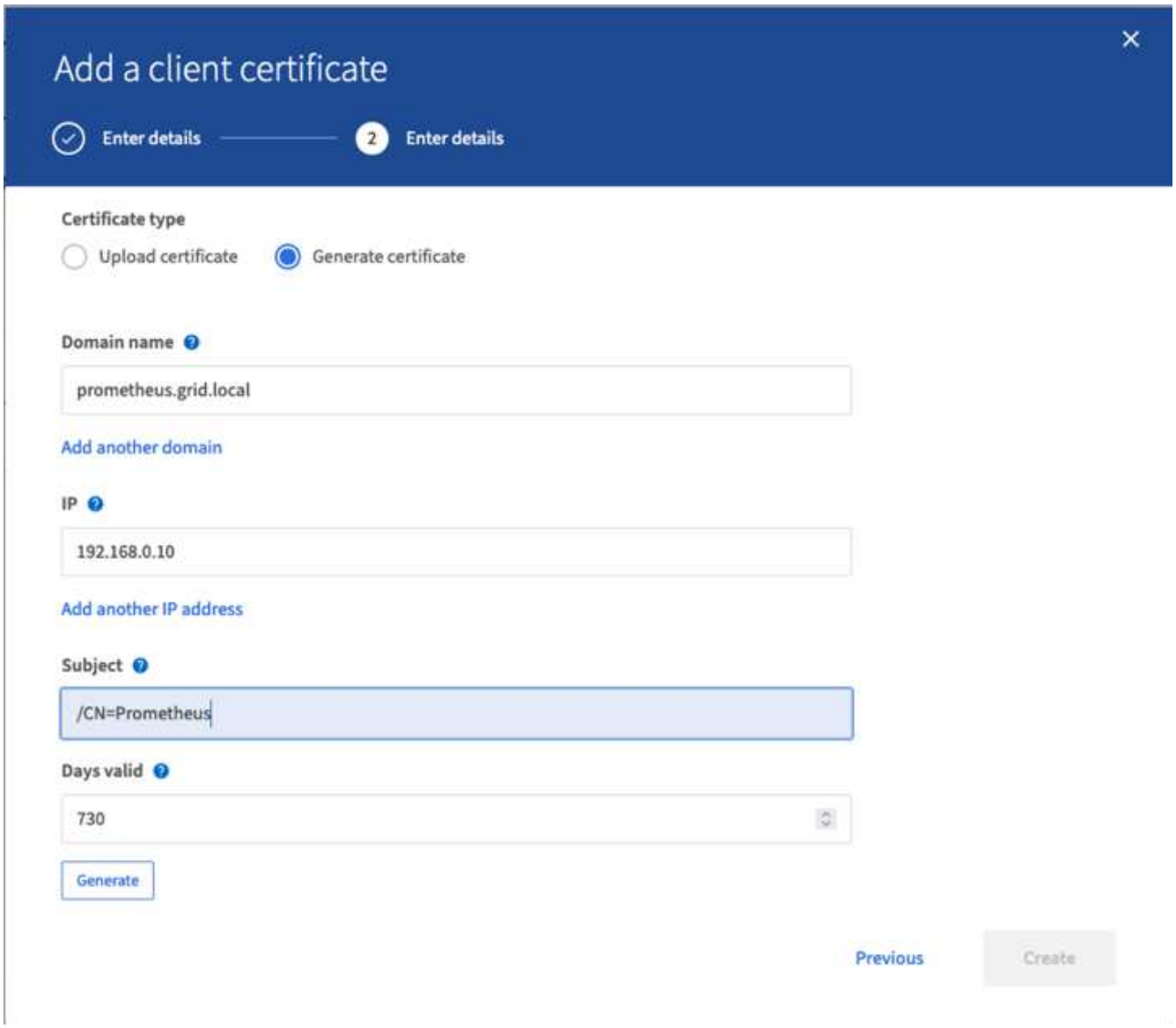
Configure StorageGRID para el acceso de Prometheus Client

Para poder acceder a las métricas de prometheus almacenadas de StorageGRID, debe generar o cargar un certificado de cliente con clave privada y habilitar el permiso para el cliente. La interfaz de gestión de StorageGRID debe tener un certificado SSL. El servidor prometheus debe confiar en este certificado mediante una CA de confianza o de forma manual si se firma automáticamente. Para leer más, visite la "Documentación de StorageGRID".

1. En la interfaz de gestión StorageGRID, seleccione "CONFIGURACIÓN" en la parte inferior izquierda y, en la segunda columna de "Seguridad", haga clic en certificados.
2. En la página certificados, seleccione la ficha "Cliente" y haga clic en el botón "Agregar".
3. Proporcione un nombre para el cliente al que se le otorgará acceso y que utilice este certificado. Haga clic en la casilla de "permisos", delante de "permitir Prometheus" y haga clic en el botón continuar.

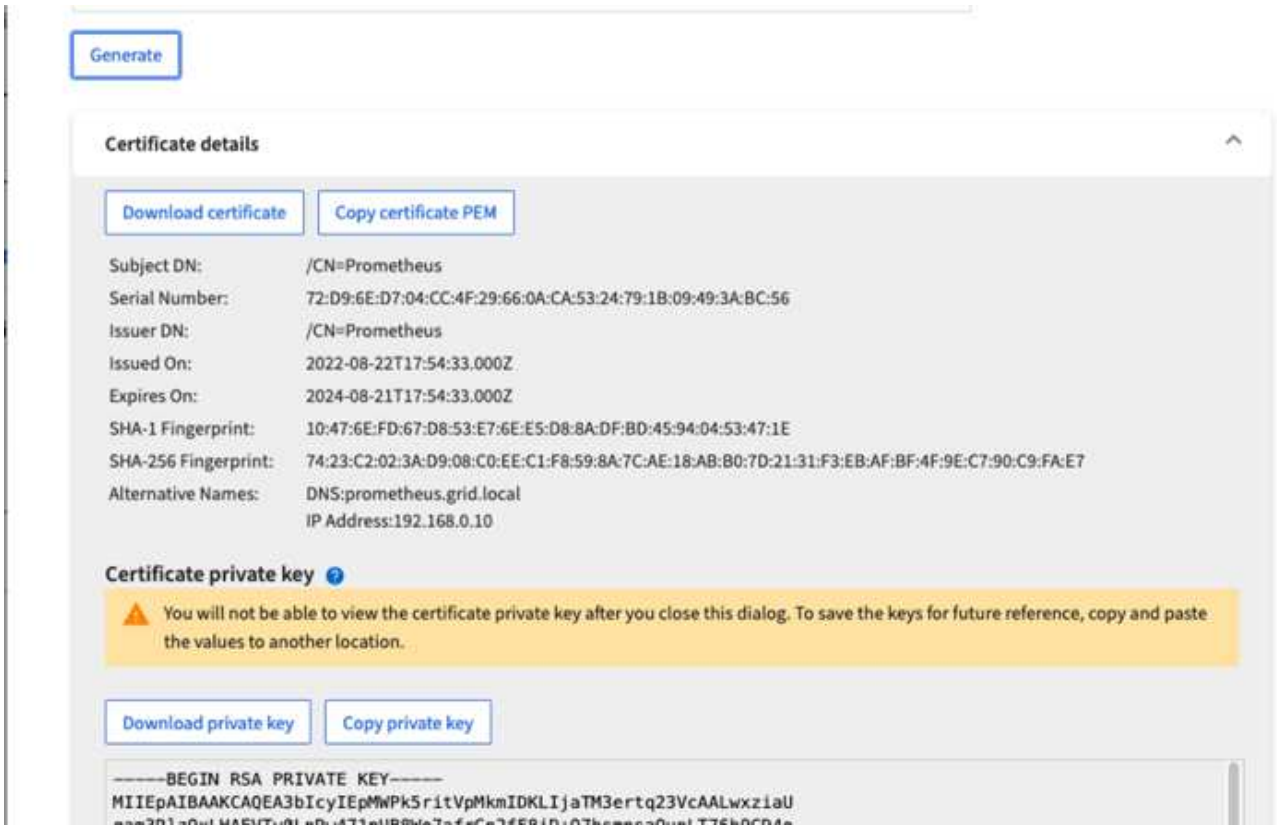


4. Si tiene un certificado firmado por CA, puede seleccionar el botón de opción "cargar certificado", pero en nuestro caso vamos a dejar que StorageGRID genere el certificado de cliente seleccionando el botón de opción "generar certificado". Los campos obligatorios se mostrarán para rellenar. Introduzca el FQDN del servidor cliente, la IP del servidor, el asunto y los días válidos. A continuación, haga clic en el botón "generar".



> Be mindful of the certificate days valid entry as you will need to renew this certificate in both StorageGRID and the Prometheus server before it expires to maintain uninterrupted collection.

1. Descargue el archivo pem del certificado y el archivo pem de la clave privada.



> This is the only time you can download the private key, so make sure you do not skip this step.

Preparar el servidor Linux para la instalación de Prometheus

Antes de instalar Prometheus, quiero preparar mi entorno con un usuario Prometheus, la estructura de directorio y configurar la capacidad para la ubicación del almacenamiento de métricas.

1. Cree el usuario Prometheus.

```
sudo useradd -M -r -s /bin/false Prometheus
```

2. Crear los directorios de Prometheus, certificado de cliente y datos de métricas.

```
sudo mkdir /etc/Prometheus /etc/Prometheus/cert /var/lib/Prometheus
```

3. He formateado el disco que estoy usando para la retención de métricas con un sistema de archivos ext4.

```
mkfs -t ext4 /dev/sdb
```

4. A continuación, he montado el sistema de archivos en el directorio de métricas Prometheus.

```
sudo mount -t auto /dev/sdb /var/lib/prometheus/
```

5. Obtenga el UUID del disco que utiliza para los datos de métricas.

```
sudo ls -al /dev/disk/by-uuid/
   lrwxrwxrwx 1 root root   9 Aug 18 17:02 9af2c5a3-bfc2-4ec1-85d9-
ebab850bb4a1 -> ../../sdb
```

6. Agregar una entrada en /etc/fstab/ haciendo que el montaje persista en reinicios utilizando el UUID de /dev/sdb.

```
/etc/fstab
UUID=9af2c5a3-bfc2-4ec1-85d9-ebab850bb4a1 /var/lib/prometheus   ext4
defaults    0   0
```

Instalar y configurar Prometheus

Ahora que el servidor está listo, puedo iniciar la instalación de Prometheus y configurar el servicio.

1. Extraiga el paquete de instalación Prometheus

```
tar xzf prometheus-2.38.0.linux-amd64.tar.gz
```

2. Copie los archivos binarios en /usr/local/bin y cambie la propiedad al usuario prometheus creado anteriormente

```
sudo cp prometheus-2.38.0.linux-amd64/{prometheus,promtool}
/usr/local/bin
sudo chown prometheus:prometheus /usr/local/bin/{prometheus,promtool}
```

3. Copie las consolas y bibliotecas en /etc/prometheus

```
sudo cp -r prometheus-2.38.0.linux-amd64/{consoles,console_libraries}
/etc/prometheus/
```

4. Copie el certificado de cliente y los archivos de pem de claves privadas descargados anteriormente de StorageGRID a /etc/prometheus/certs
5. Cree el archivo yaml de configuración de prometheus

```
sudo nano /etc/prometheus/prometheus.yml
```

6. Inserte la siguiente configuración. El nombre del trabajo puede ser cualquier cosa que desee. Cambie el "-Targets: ["]" al FQDN del nodo admin y, si modificó los nombres del certificado y los nombres de los archivos de claves privadas, actualice la sección tls_config para que coincida. a continuación, guarde el archivo. Si la interfaz de gestión de grid utiliza un certificado autofirmado, descargue el certificado y colóquelo con el certificado de cliente con un nombre único, y en la sección tls_config añada CA_file: /Etc/prometheus/cert/UIcert.pem

   a. En este ejemplo estoy recopilando todas las métricas que empiezan con alertManager, cassandra, nodo y StorageGRID. Puede ver más información sobre la métrica Prometheus en la "Documentación de StorageGRID".

```
# my global config
global:
  scrape_interval: 60s # Set the scrape interval to every 15 seconds.
Default is every 1 minute.

scrape_configs:
  - job_name: 'StorageGRID'
    honor_labels: true
    scheme: https
    metrics_path: /federate
    scrape_interval: 60s
    scrape_timeout: 30s
    tls_config:
      cert_file: /etc/prometheus/cert/certificate.pem
      key_file: /etc/prometheus/cert/private_key.pem
    params:
      match[]:
        -
'{__name__=~"alertmanager_.*|cassandra_.*|node_.*|storagegrid_.*"}'
    static_configs:
    - targets: ['sgdemo-rtp.netapp.com:9091']
```

Si la interfaz de gestión de grid utiliza un certificado autofirmado, descargue el certificado y colóquelo con el certificado de cliente con un nombre único. En la sección tls_config, agregue el certificado encima del certificado de cliente y las líneas de clave privada

```
        ca_file: /etc/prometheus/cert/UIcert.pem
```

1. Cambie la propiedad de todos los archivos y directorios en /etc/prometheus y /var/lib/prometheus al usuario prometheus

```
sudo chown -R prometheus:prometheus /etc/prometheus/
sudo chown -R prometheus:prometheus /var/lib/prometheus/
```

2. Cree un archivo de servicio prometheus en /etc/systemd/system

```
sudo nano /etc/systemd/system/prometheus.service
```

3. Inserte las siguientes líneas, observe el --Storage.tsdb.retention.time=1 año que establece la retención de los datos de la métrica en 1 año. También puede usar --Storage.tsdb.retention.size=300GIB para basar la retención en los límites de almacenamiento. Esta es la única ubicación donde se establece la retención de las métricas.

```
[Unit]
Description=Prometheus Time Series Collection and Processing Server
Wants=network-online.target
After=network-online.target

[Service]
User=prometheus
Group=prometheus
Type=simple
ExecStart=/usr/local/bin/prometheus \
        --config.file /etc/prometheus/prometheus.yml \
        --storage.tsdb.path /var/lib/prometheus/ \
        --storage.tsdb.retention.time=1y \
        --web.console.templates=/etc/prometheus/consoles \
        --web.console.libraries=/etc/prometheus/console_libraries

[Install]
WantedBy=multi-user.target
```

4. Vuelva a cargar el servicio systemd para registrar el nuevo servicio prometheus. a continuación, inicie y habilite el servicio prometheus.

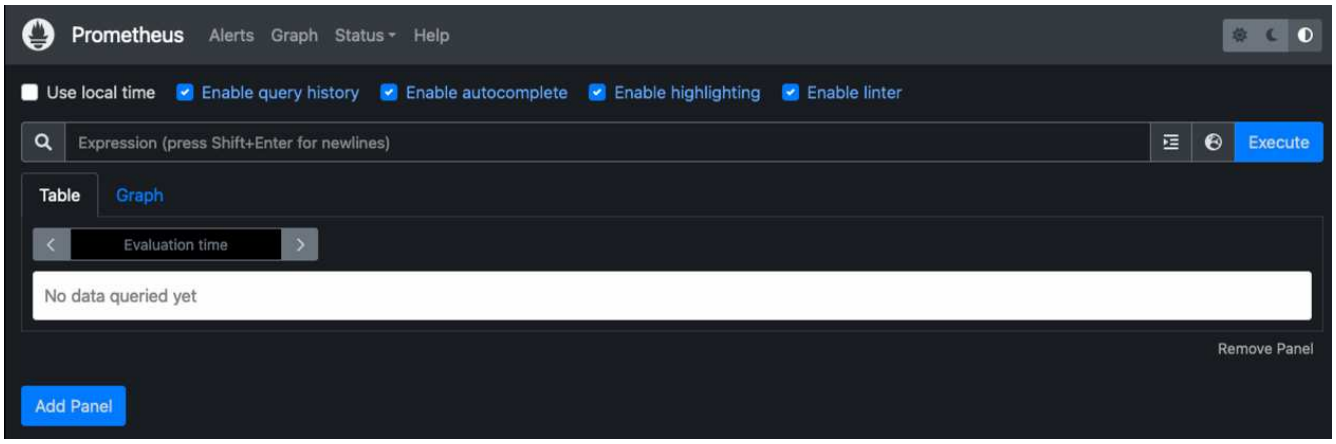
```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Compruebe que el servicio está funcionando correctamente

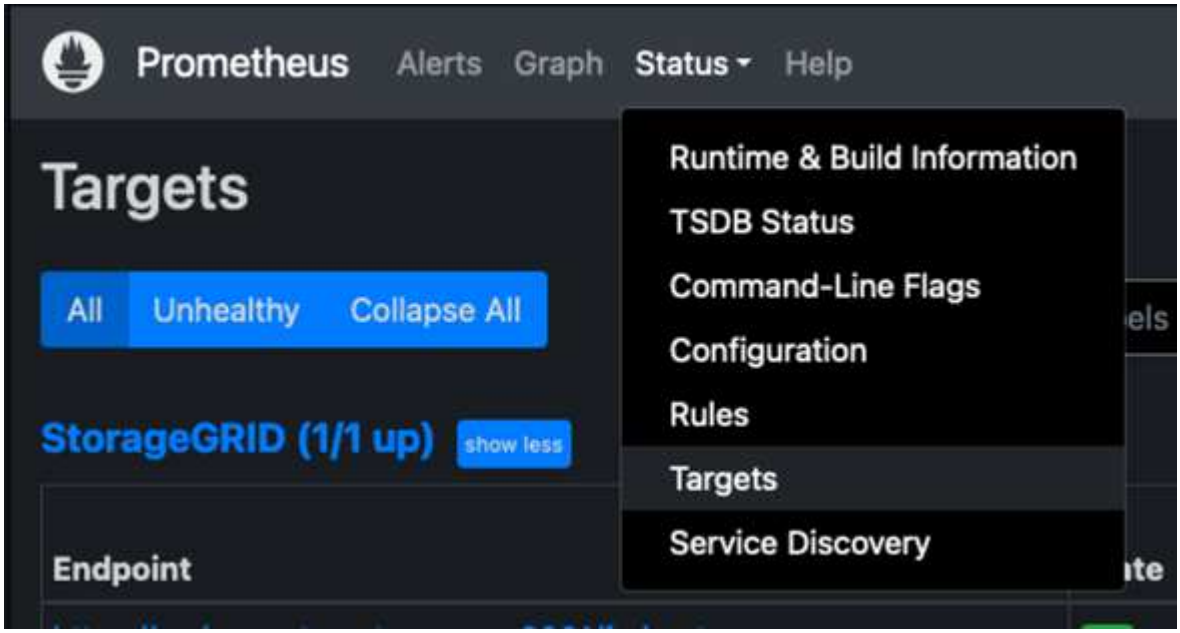
```
sudo systemctl status prometheus
```

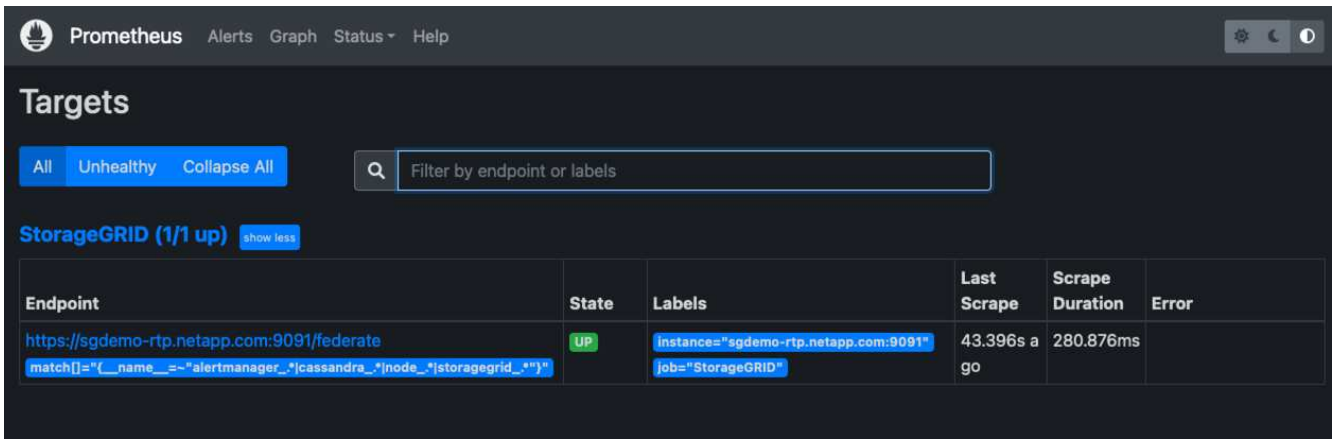
```
● prometheus.service - Prometheus Time Series Collection and Processing
Server
     Loaded: loaded (/etc/systemd/system/prometheus.service; enabled;
vendor preset: enabled)
     Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago
   Main PID: 6498 (prometheus)
      Tasks: 13 (limit: 28818)
     Memory: 107.7M
        CPU: 1.143s
     CGroup: /system.slice/prometheus.service
             └─6498 /usr/local/bin/prometheus --config.file
/etc/prometheus/prometheus.yml --storage.tsdb.path /var/lib/prometheus/
--web.console.templates=/etc/prometheus/consoles --web.con>

Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.510Z caller=head.go:544 level=info component=tsdb
msg="Replaying WAL, this may take a while"
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.816Z caller=head.go:615 level=info component=tsdb msg="WAL
segment loaded" segment=0 maxSegment=1
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.816Z caller=head.go:615 level=info component=tsdb msg="WAL
segment loaded" segment=1 maxSegment=1
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.816Z caller=head.go:621 level=info component=tsdb msg="WAL
replay completed" checkpoint_replay_duration=55.57µs wal_rep>
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.831Z caller=main.go:997 level=info fs_type=EXT4_SUPER_MAGIC
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.831Z caller=main.go:1000 level=info msg="TSDB started"
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.831Z caller=main.go:1181 level=info msg="Loading
configuration file" filename=/etc/prometheus/prometheus.yml
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.832Z caller=main.go:1218 level=info msg="Completed loading
of configuration file" filename=/etc/prometheus/prometheus.y>
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.832Z caller=main.go:961 level=info msg="Server is ready to
receive web requests."
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.832Z caller=manager.go:941 level=info component="rule
manager" msg="Starting rule manager..."
```

6. Ahora debe poder acceder a la interfaz de usuario de su servidor prometheus http://Prometheus-server:9090 Y ver la interfaz de usuario

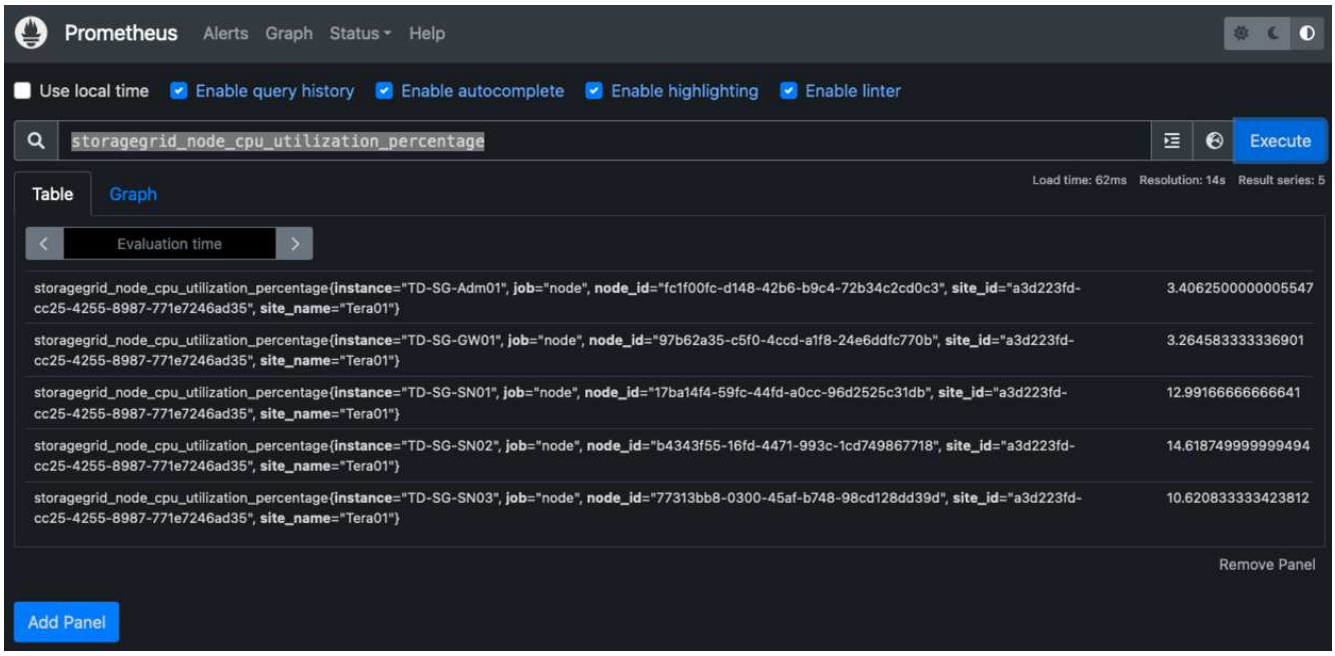


7. En "Estado", puede ver el estado del extremo StorageGRID que hemos configurado en prometheus.yml





8. En la página Graph, puede ejecutar una consulta de prueba y comprobar que los datos se están raspando correctamente. Por ejemplo, introduzca "storagegrid_node_cpu_Utilization_porcentual" en la barra de consultas y haga clic en el botón Execute.



Instalar y configurar Grafana

Ahora que prometheus está instalado y en funcionamiento, podemos pasar a la instalación de Grafana y configurar una consola

Grafana Instalation

1. Instale la última edición empresarial de Grafana

```
sudo apt-get install -y apt-transport-https
sudo apt-get install -y software-properties-common wget
sudo wget -q -O /usr/share/keyrings/grafana.key
https://packages.grafana.com/gpg.key
```

2. Añada este repositorio para versiones estables:

```
echo "deb [signed-by=/usr/share/keyrings/grafana.key]
https://packages.grafana.com/enterprise/deb stable main" | sudo tee -a
/etc/apt/sources.list.d/grafana.list
```

3. Después de agregar el repositorio.

```
sudo apt-get update
sudo apt-get install grafana-enterprise
```

4. Vuelva a cargar el servicio systemd para registrar el nuevo servicio grafana. A continuación, inicie y habilite el servicio Grafana.

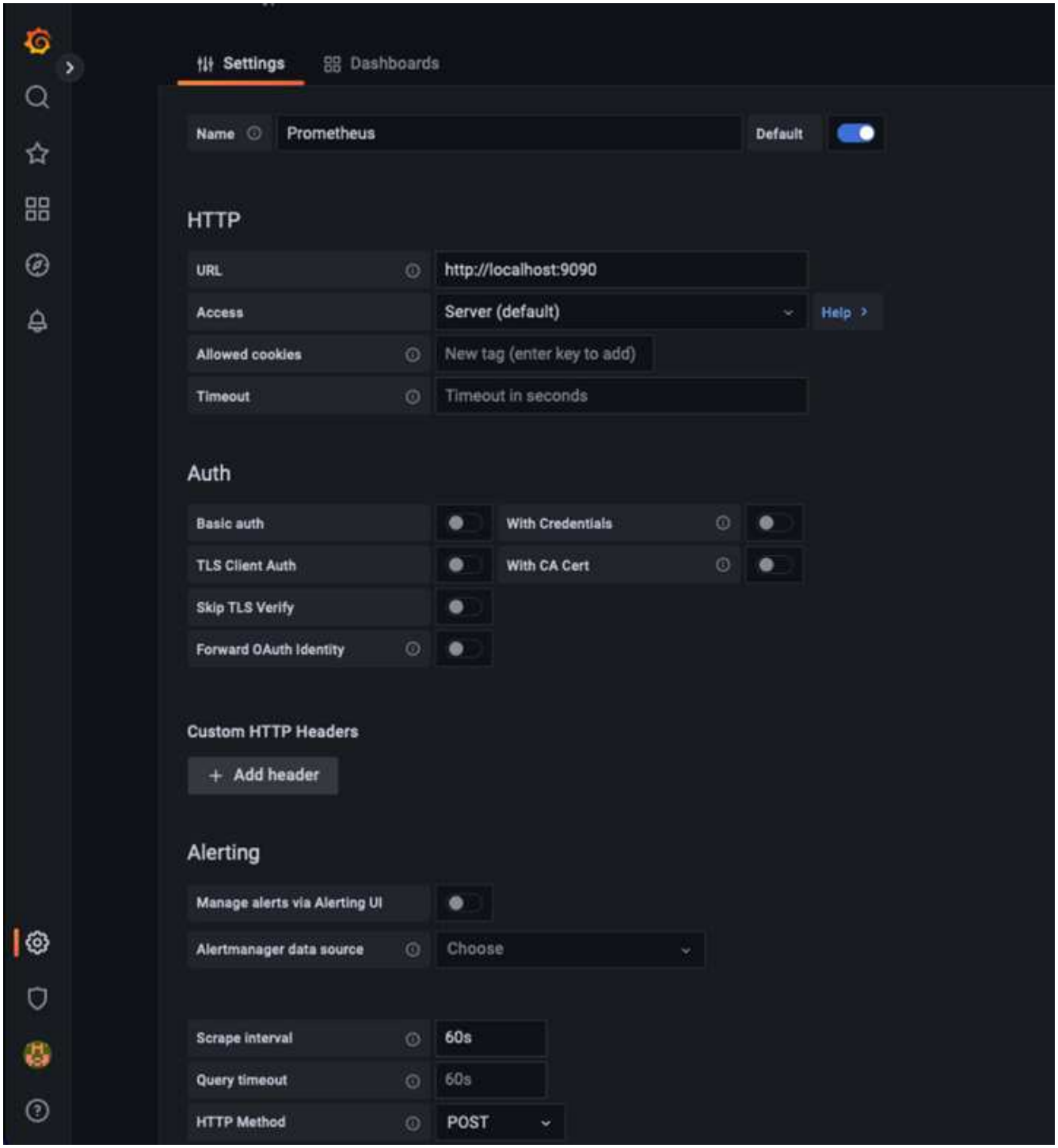
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. Grafana ya está instalado y en funcionamiento. Cuando abra un navegador a HTTP://Prometheus-Server:3000 recibirá la página de inicio de sesión de Grafana.
6. Las credenciales de inicio de sesión predeterminadas son admin/admin, y debe configurar una contraseña nueva cuando le solicite.

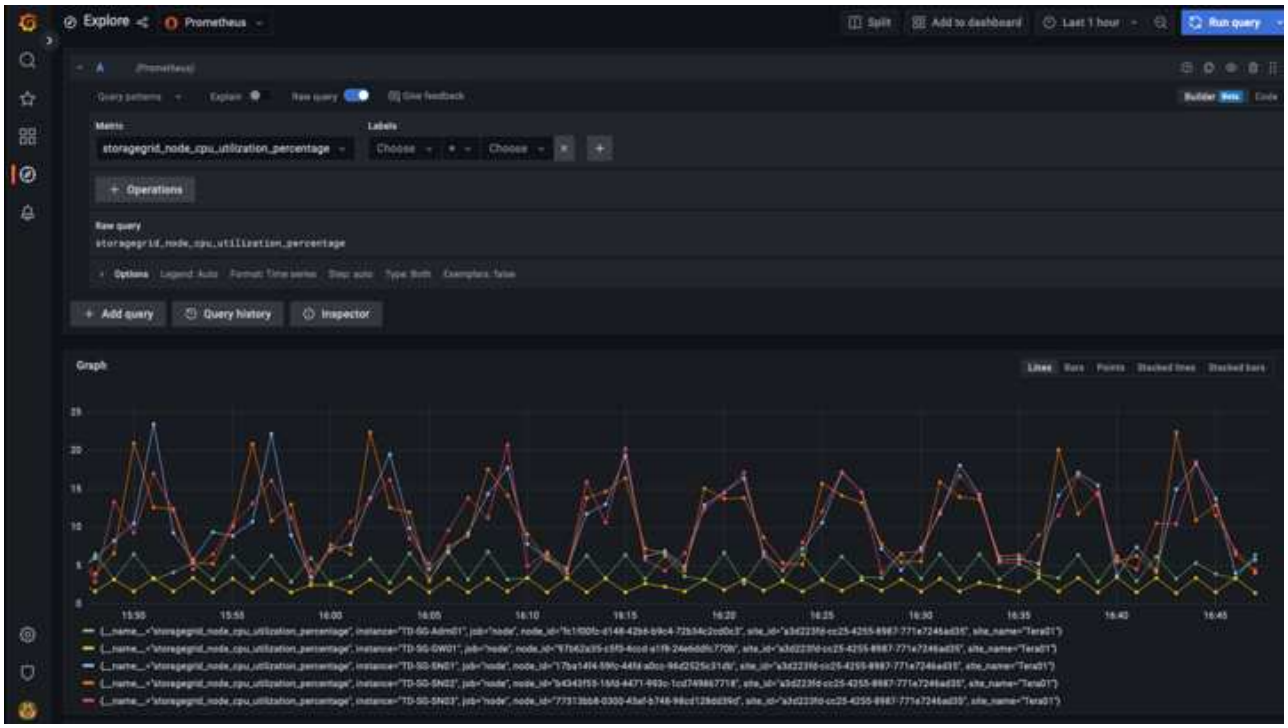
Cree un panel de Grafana para StorageGRID

Con Grafana y Prometheus instalados y en ejecución, ahora es hora de conectar los dos mediante la creación de un origen de datos y la creación de un panel

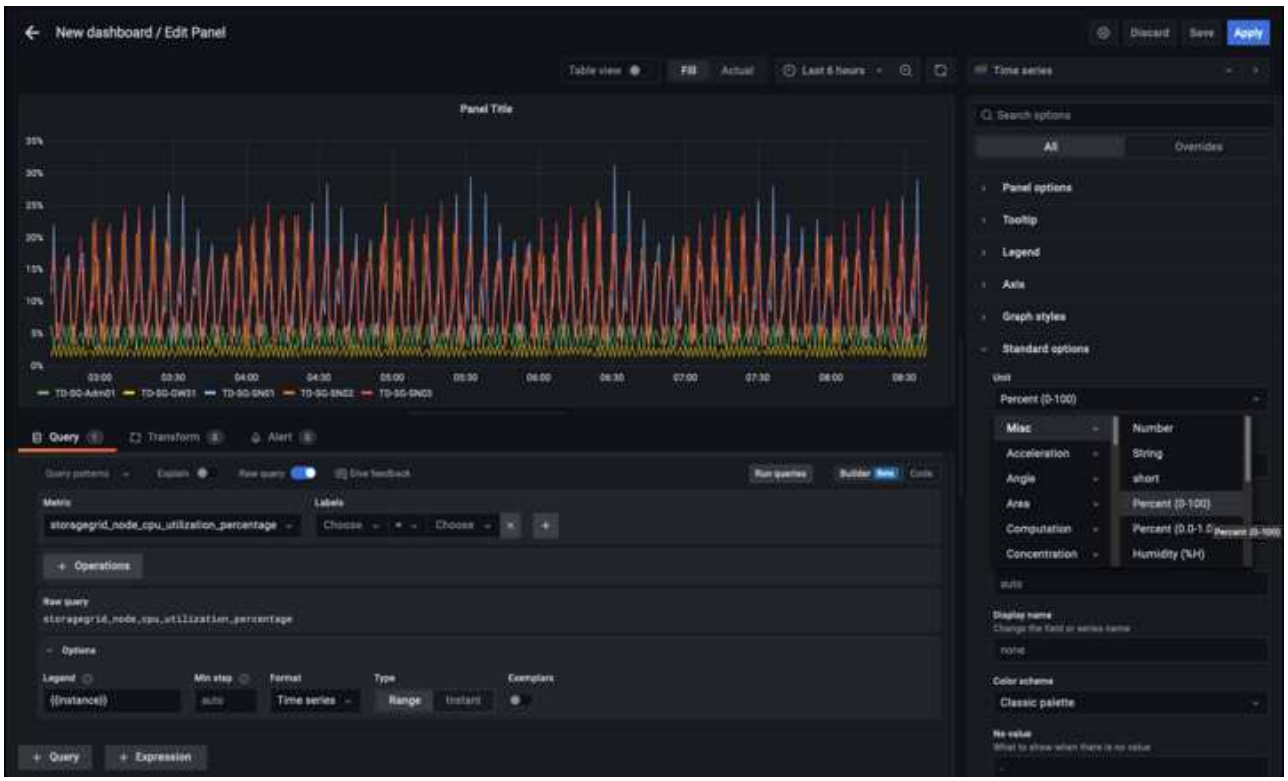
1. En el panel izquierdo, expanda "Configuración" y seleccione "orígenes de datos" y, a continuación, haga clic en el botón "Agregar origen de datos"
2. Prometheus será una de las principales fuentes de datos entre las que elegir. Si no lo es, utilice la barra de búsqueda para localizar "Prometheus"
3. Para configurar el origen Prometheus, introduzca la URL de la instancia prometheus y el intervalo de raspado para que coincidan con el intervalo Prometheus. También he deshabilitado la sección de alertas, ya que no configuré el administrador de alertas en prometheus.



4. Con la configuración deseada introducida, desplácese hacia abajo hasta la parte inferior y haga clic en "Guardar y probar"
5. Una vez que la prueba de configuración se haya realizado correctamente, haga clic en el botón explorar.
    a. En la ventana explorar, puede utilizar la misma métrica que probamos Prometheus con "storagegrid_node_cpu_Utilization_porcentual" y hacer clic en el botón "Ejecutar consulta"



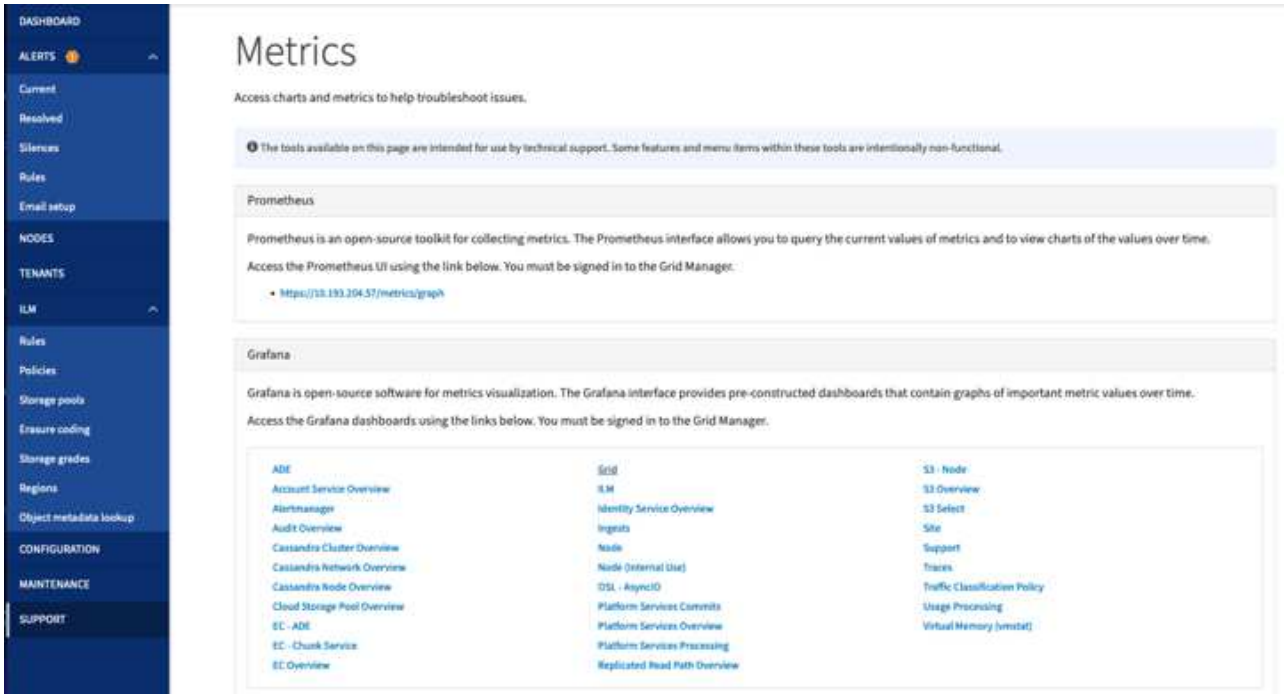
6. Ahora que tenemos configurado el origen de datos, podemos crear un panel de control.
   a. En el panel izquierdo, expanda "Paneles" y seleccione "+ New Dashboard".
   b. Seleccione "Añadir un nuevo panel"
   c. Configure el nuevo panel seleccionando una métrica, de nuevo utilizaré "storagegrid_node_cpu_Utilization_Percent", Introduzca un título para el panel, expanda "Opciones" en la parte inferior y para que la leyenda cambie a personalizado e introduzca "{{Instance}" para definir los nombres de los nodos y, en el panel derecho, en "Opciones estándar", defina "Unidad" en "Misc-100). A continuación, haga clic en "aplicar" para guardar el panel en el tablero de a bordo.



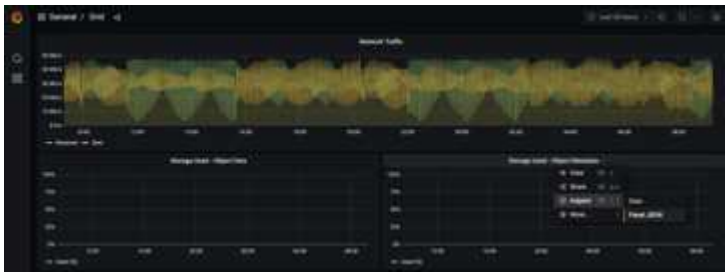
7. Podríamos seguir desarrollando nuestro panel de control como este para cada métrica que deseamos, pero por suerte StorageGRID ya dispone de paneles con paneles que podemos copiar en nuestros paneles personalizados.

   a. En el panel izquierdo de la interfaz de gestión de StorageGRID, seleccione «Soporte» y, en la parte inferior de la columna «Herramientas», haga clic en «Métricas».

   b. Dentro de las métricas, voy a seleccionar el enlace "Grid" en la parte superior de la columna central.

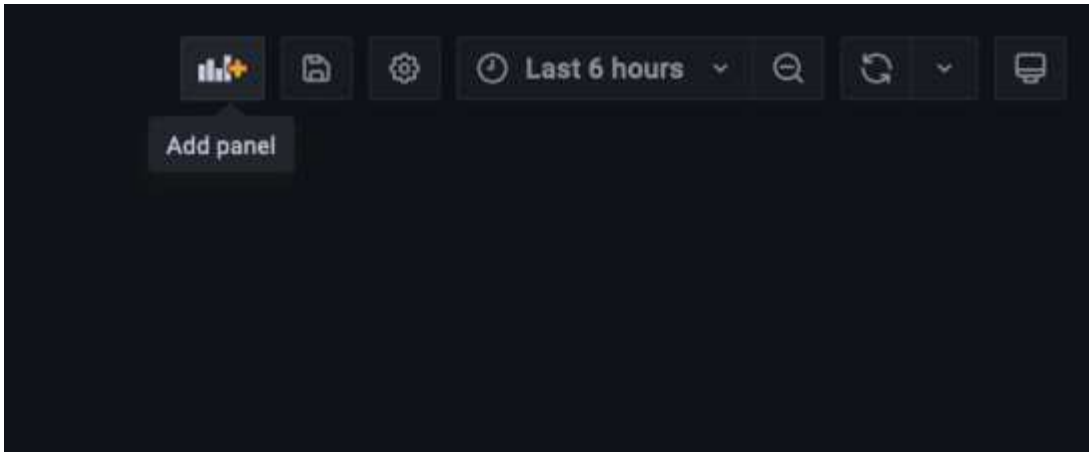


   c. En el panel de cuadrícula, permite seleccionar el panel "almacenamiento usado - metadatos de objeto". Haga clic en la flecha abajo y en el final del título del panel para ver un menú desplegable. En este menú, seleccione "inspeccionar" y "Panel JSON".



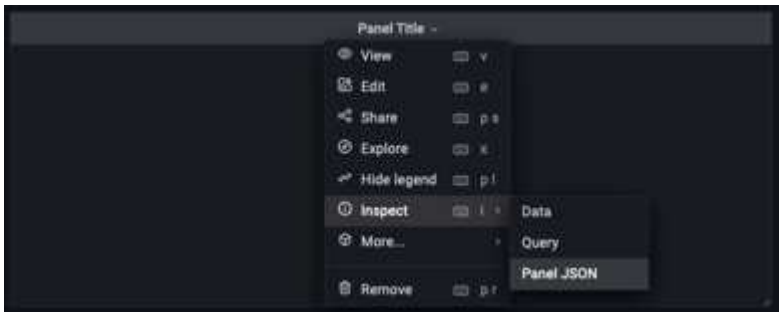
   d. Copie el código JSON y cierre la ventana.

e. En nuestro nuevo panel, haga clic en el icono para añadir un nuevo panel.

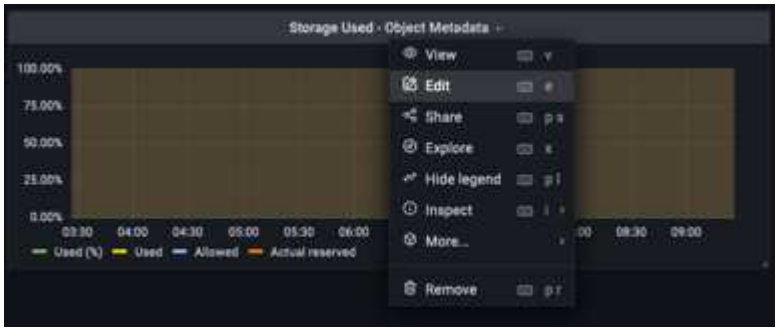


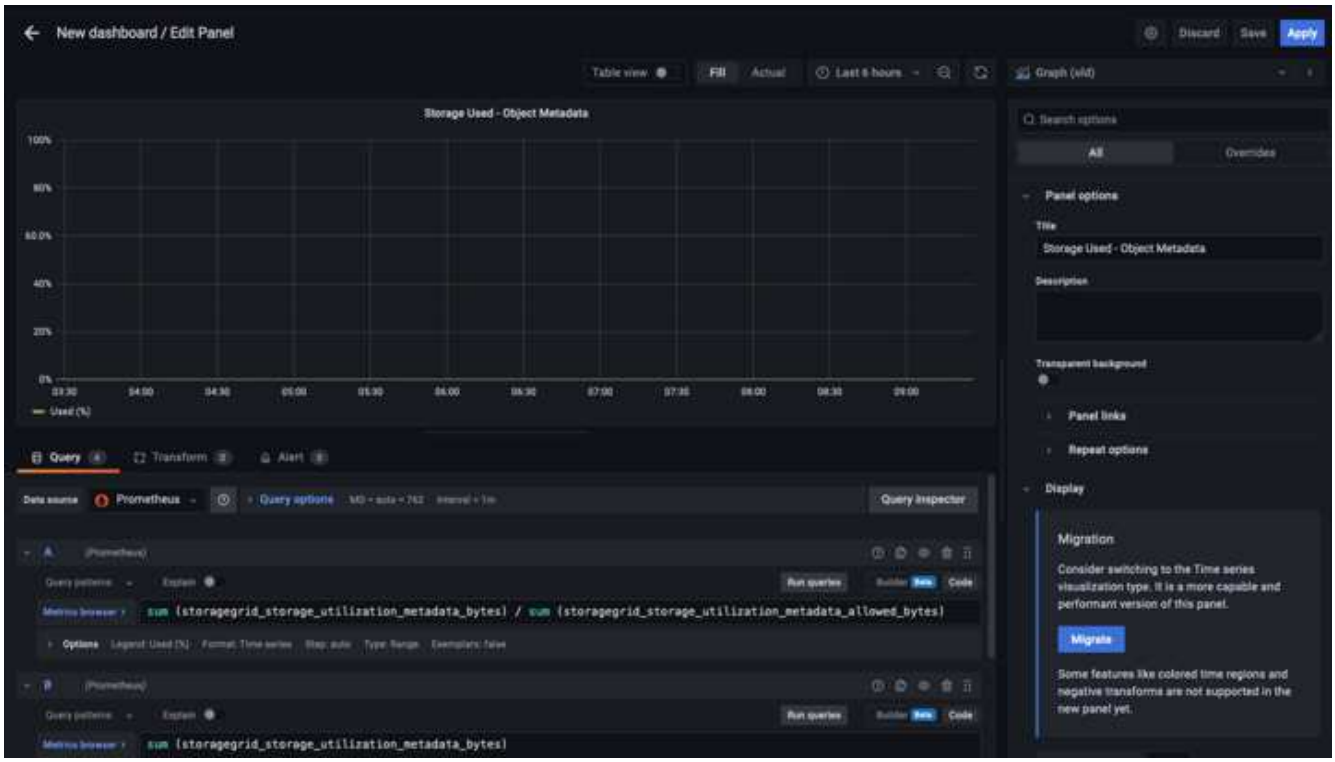
f. Aplique el nuevo panel sin realizar cambios

g. Al igual que con el panel StorageGRID, inspeccione el JSON. Quite todo el código JSON y sustitúyalo por el código copiado del panel StorageGRID.



h. Edite el nuevo panel y, a la derecha, verá un mensaje de migración con el botón "migrar". Haga clic en el botón y, a continuación, en el botón "aplicar".





8. Una vez que tenga todos los paneles en su lugar y configurados como desee. Guarde el panel haciendo clic en el icono de disco de la parte superior derecha y asigne un nombre a su panel.

Conclusión

Ahora disponemos de un servidor Prometheus con capacidad personalizable de almacenamiento y retención de datos. De este modo, podemos desarrollar nuestros propios paneles con las métricas más relevantes para nuestras operaciones. Puede obtener más información sobre las métricas de Prometheus recopiladas en el "Documentación de StorageGRID".

Por Aron Klein

Configuración de SNMP de Datadog

Configurar Datadog para recopilar métricas y capturas snmp de StorageGRID.

Configurar Datadog

Datadog es una solución de supervisión que ofrece métricas, visualizaciones y alertas. La siguiente configuración se implementó con el agente de linux versión 7.43.1 en un host de Ubuntu 22.04.1 implementado localmente en el sistema StorageGRID.

Perfil de datos y archivos de captura generados a partir del archivo MIB de StorageGRID

Datadog proporciona un método para convertir archivos MIB de producto en archivos de referencia de datadog necesarios para asignar los mensajes SNMP.

Este archivo yaml de StorageGRID para la asignación de resolución de solapamiento de Datadog se generó siguiendo la instrucción encontrada "aquí". + Coloque este archivo en /etc/datadog-agent/conf.d/snmp.d/traps_dB/ +

- "Descargue el archivo yaml de captura" +
  - **md5 checksum** 42e274210719945a46172b98c379517 +
  - * sha256 checksum* d0fe5c8e6ca3c902d054f854b70a85f928cba8b7c76391d356f05d2cf73b6887 +

Este archivo yaml de perfil de StorageGRID para la asignación de métricas de Datadog generada siguiendo la instrucción encontrada "aquí". + Coloque este archivo en /etc/datatog-agent/conf.d/snmp.d/profiles/ +

- "Descargue el archivo yaml de perfil" +
  - * md5 checksum* 72bb7784f4801adda4e0c3ea77df19aa +
  - * sha256 checksum* b6b7fadd33063422a8bb8e39b3ead8ab38349ee0229926eadc85f0087b8cee +

Configuración de Datadog de SNMP para Metrics

La configuración de SNMP para métricas se puede administrar de dos maneras. Puede configurar para la detección automática proporcionando un rango de direcciones de red que contenga los sistemas StorageGRID o definiendo las IP de los dispositivos individuales. La ubicación de la configuración es diferente en función de la decisión tomada. La detección automática se define en el archivo yaml del agente datatog. Las definiciones de dispositivos explícitas se configuran en el archivo yaml de configuración de snmp. A continuación, se muestran ejemplos de cada uno para el mismo sistema StorageGRID.

Detección automática

configuración ubicada en /etc/datadir-agent/datadir.yaml

```
listeners:
  - name: snmp
snmp_listener:
  workers: 100  # number of workers used to discover devices concurrently
  discovery_interval: 3600  # interval between each autodiscovery in
seconds
  loader: core  # use core check implementation of SNMP integration.
recommended
  use_device_id_as_hostname: true  # recommended
  configs:
    - network_address: 10.0.0.0/24  # CIDR subnet
      snmp_version: 2
      port: 161
      community_string: 'st0r@gegrid'  # enclose with single quote
      profile: netapp-storagegrid
```

Dispositivos individuales

/etc/datatog-agent/conf.d/snmp.d/conf.yaml

```
init_config:
  loader: core  # use core check implementation of SNMP integration.
recommended
  use_device_id_as_hostname: true  # recommended
instances:
- ip_address: '10.0.0.1'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'  # enclose with single quote
- ip_address: '10.0.0.2'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'
- ip_address: '10.0.0.3'
  profile: netapp-storagegrid
 community_string: 'st0r@gegrid'
- ip_address: '10.0.0.4'
  profile: netapp-storagegrid
  community_string: 'st0r@gegrid'
```

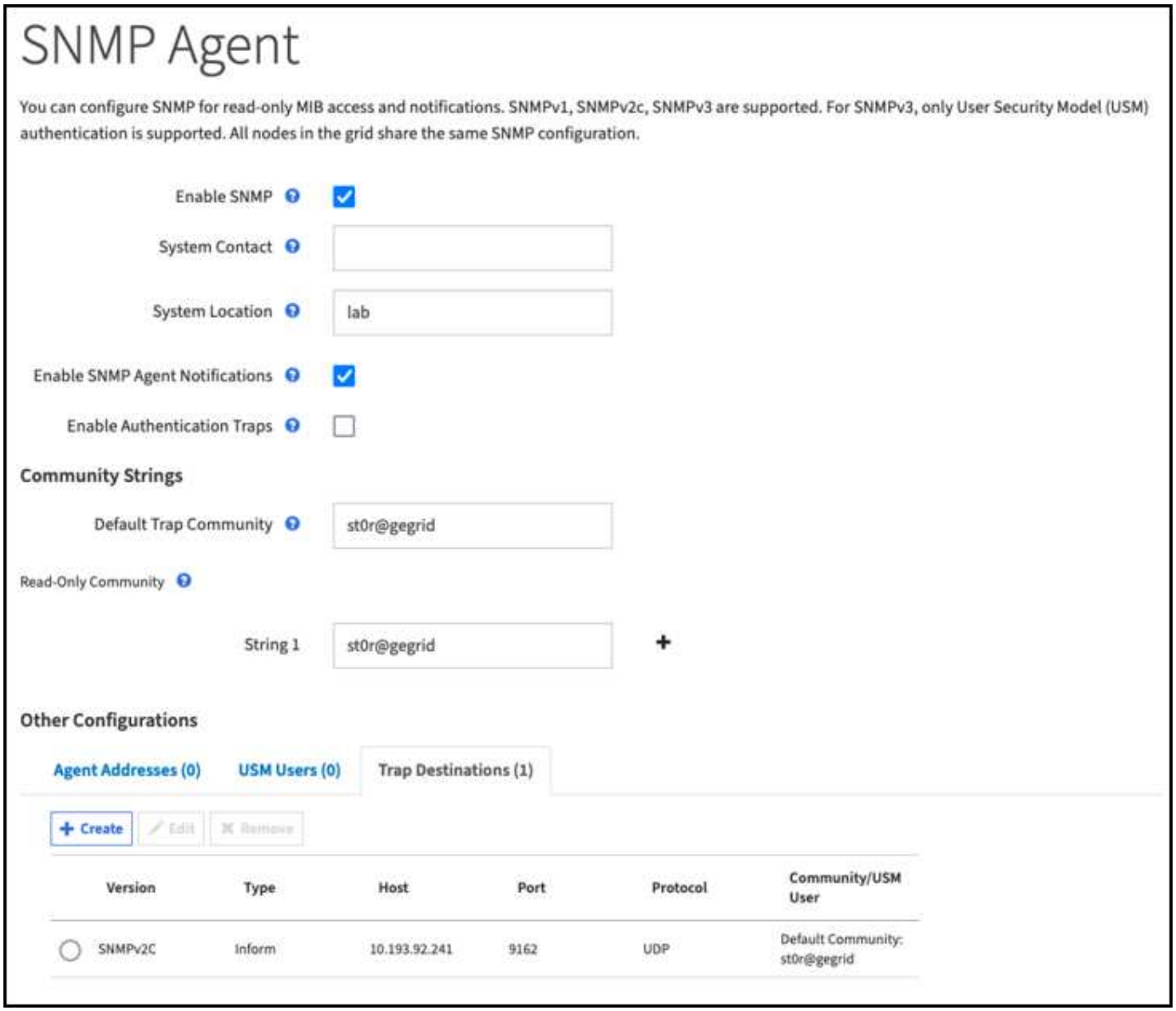
Configuración de SNMP para las capturas

La configuración de los solapamientos SNMP se define en el archivo de configuración de datadir ylma /etc/datadir-agent/datadog.ylma

```
network_devices:
  namespace:  # optional, defaults to "default".
  snmp_traps:
    enabled: true
    port: 9162 # on which ports to listen for traps
    community_strings: # which community strings to allow for v2 traps
      - st0r@gegrid
```

Ejemplo de configuración de SNMP de StorageGRID

El agente SNMP del sistema StorageGRID se encuentra en la pestaña Configuración, columna Supervisión. Habilite SNMP e introduzca la información que desee. Si desea configurar capturas, seleccione "Destinos de capturas" y cree un destino para el host del agente de datos que contenga la configuración de capturas.



Por Aron Klein

Utilice rclone para migrar, PONER y ELIMINAR objetos en StorageGRID

Rclone es una herramienta de línea de comandos gratuita y un cliente para operaciones S3. Es posible usar rclone para migrar, copiar y eliminar datos de objetos en StorageGRID. rclone incluye la capacidad de eliminar cubos incluso cuando no están vacíos con una función de “purga” como se muestra en el siguiente ejemplo.

Instalar y configurar rclone

Para instalar rclone en una estación de trabajo o servidor, descárguelo de "rclone.org".

Pasos de configuración inicial

1. Cree el archivo de configuración rclone ejecutando el script de configuración o creando manualmente el archivo.
2. Para este ejemplo usaré sgdemo para el nombre del punto final remoto de StorageGRID S3 en la configuración rclone.
   a. Cree el archivo de configuración ~/.config/rclone/rclone.conf

```
[sgdemo]
type = s3
provider = Other
access_key_id = ABCDEFGH123456789JKL
secret_access_key = 123456789ABCDEFGHIJKLMN0123456789PQRST+V
endpoint = sgdemo.netapp.com
```

   b. Ejecute rclone config

rclone config

```
2023/04/13 14:22:45 NOTICE: Config file
"/root/.config/rclone/rclone.conf" not found - using defaults
No remotes found - make a new one
n) New remote
s) Set configuration password
q) Quit config
n/s/q> n
name> sgdemo
```

```
Option Storage.
Type of storage to configure.
Enter a string value. Press Enter for the default ("").
Choose a number from below, or type in your own value.
 1 / 1Fichier
   \ "fichier"
 2 / Alias for an existing remote
   \ "alias"
 3 / Amazon Drive
   \ "amazon cloud drive"
 4 / Amazon S3 Compliant Storage Providers including AWS,
Alibaba, Ceph, Digital Ocean, Dreamhost, IBM COS, Minio,
SeaweedFS, and Tencent COS
   \ "s3"
 5 / Backblaze B2
   \ "b2"
 6 / Better checksums for other remotes
   \ "hasher"
 7 / Box
   \ "box"
 8 / Cache a remote
   \ "cache"
 9 / Citrix Sharefile
   \ "sharefile"
10 / Compress a remote
   \ "compress"
11 / Dropbox
   \ "dropbox"
12 / Encrypt/Decrypt a remote
   \ "crypt"
13 / Enterprise File Fabric
   \ "filefabric"
14 / FTP Connection
```

```
   \ "ftp"
15 / Google Cloud Storage (this is not Google Drive)
   \ "google cloud storage"
16 / Google Drive
   \ "drive"
17 / Google Photos
   \ "google photos"
18 / Hadoop distributed file system
   \ "hdfs"
19 / Hubic
   \ "hubic"
20 / In memory object storage system.
   \ "memory"
21 / Jottacloud
   \ "jottacloud"
22 / Koofr
   \ "koofr"
23 / Local Disk
   \ "local"
24 / Mail.ru Cloud
   \ "mailru"
25 / Mega
   \ "mega"
26 / Microsoft Azure Blob Storage
   \ "azureblob"
27 / Microsoft OneDrive
   \ "onedrive"
28 / OpenDrive
   \ "opendrive"
29 / OpenStack Swift (Rackspace Cloud Files, Memset Memstore,
OVH)
   \ "swift"
30 / Pcloud
   \ "pcloud"
31 / Put.io
   \ "putio"
32 / QingCloud Object Storage
   \ "qingstor"
33 / SSH/SFTP Connection
   \ "sftp"
34 / Sia Decentralized Cloud
   \ "sia"
35 / Sugarsync
   \ "sugarsync"
36 / Tardigrade Decentralized Cloud Storage
   \ "tardigrade"
```

```
37 / Transparently chunk/split large files
   \ "chunker"
38 / Union merges the contents of several upstream fs
   \ "union"
39 / Uptobox
   \ "uptobox"
40 / Webdav
   \ "webdav"
41 / Yandex Disk
   \ "yandex"
42 / Zoho
   \ "zoho"
43 / http Connection
   \ "http"
44 / premiumize.me
   \ "premiumizeme"
45 / seafile
   \ "seafile"
```

```
Storage> 4
```

```
Option provider.
Choose your S3 provider.
Enter a string value. Press Enter for the default ("").
Choose a number from below, or type in your own value.
 1 / Amazon Web Services (AWS) S3
   \ "AWS"
 2 / Alibaba Cloud Object Storage System (OSS) formerly Aliyun
   \ "Alibaba"
 3 / Ceph Object Storage
   \ "Ceph"
 4 / Digital Ocean Spaces
   \ "DigitalOcean"
 5 / Dreamhost DreamObjects
   \ "Dreamhost"
 6 / IBM COS S3
   \ "IBMCOS"
 7 / Minio Object Storage
   \ "Minio"
 8 / Netease Object Storage (NOS)
   \ "Netease"
 9 / Scaleway Object Storage
   \ "Scaleway"
10 / SeaweedFS S3
   \ "SeaweedFS"
11 / StackPath Object Storage
   \ "StackPath"
12 / Tencent Cloud Object Storage (COS)
   \ "TencentCOS"
13 / Wasabi Object Storage
   \ "Wasabi"
14 / Any other S3 compatible provider
   \ "Other"
provider> 14
```

```
Option env_auth.
Get AWS credentials from runtime (environment variables or
EC2/ECS meta data if no env vars).
Only applies if access_key_id and secret_access_key is blank.
Enter a boolean value (true or false). Press Enter for the
default ("false").
Choose a number from below, or type in your own value.
 1 / Enter AWS credentials in the next step.
   \ "false"
 2 / Get AWS credentials from the environment (env vars or IAM).
   \ "true"
env_auth> 1
```

```
Option access_key_id.
AWS Access Key ID.
Leave blank for anonymous access or runtime credentials.
Enter a string value. Press Enter for the default ("").
access_key_id> ABCDEFGH123456789JKL
```

```
Option secret_access_key.
AWS Secret Access Key (password).
Leave blank for anonymous access or runtime credentials.
Enter a string value. Press Enter for the default ("").
secret_access_key> 123456789ABCDEFGHIJKLMN0123456789PQRST+V
```

```
Option region.
Region to connect to.
Leave blank if you are using an S3 clone and you don't have a
region.
Enter a string value. Press Enter for the default ("").
Choose a number from below, or type in your own value.
   / Use this if unsure.
 1 | Will use v4 signatures and an empty region.
   \ ""
   / Use this only if v4 signatures don't work.
 2 | E.g. pre Jewel/v10 CEPH.
   \ "other-v2-signature"
region> 1
```

```
Option endpoint.
Endpoint for S3 API.
Required when using an S3 clone.
Enter a string value. Press Enter for the default ("").
endpoint> sgdemo.netapp.com
```

```
Option location_constraint.
Location constraint - must be set to match the Region.
Leave blank if not sure. Used when creating buckets only.
Enter a string value. Press Enter for the default ("").
location_constraint>
```

```
Option acl.
Canned ACL used when creating buckets and storing or copying
objects.
This ACL is used for creating objects and if bucket_acl isn't
set, for creating buckets too.
For more info visit
https://docs.aws.amazon.com/AmazonS3/latest/dev/acl-
overview.html#canned-acl
Note that this ACL is applied when server-side copying objects as
S3
doesn't copy the ACL from the source but rather writes a fresh
one.
Enter a string value. Press Enter for the default ("").
Choose a number from below, or type in your own value.
   / Owner gets FULL_CONTROL.
 1 | No one else has access rights (default).
   \ "private"
   / Owner gets FULL_CONTROL.
 2 | The AllUsers group gets READ access.
   \ "public-read"
   / Owner gets FULL_CONTROL.
 3 | The AllUsers group gets READ and WRITE access.
   | Granting this on a bucket is generally not recommended.
   \ "public-read-write"
   / Owner gets FULL_CONTROL.
 4 | The AuthenticatedUsers group gets READ access.
   \ "authenticated-read"
   / Object owner gets FULL_CONTROL.
 5 | Bucket owner gets READ access.
   | If you specify this canned ACL when creating a bucket,
Amazon S3 ignores it.
   \ "bucket-owner-read"
   / Both the object owner and the bucket owner get FULL_CONTROL
over the object.
 6 | If you specify this canned ACL when creating a bucket,
Amazon S3 ignores it.
   \ "bucket-owner-full-control"
acl>
```

```
Edit advanced config?
y) Yes
n) No (default)
y/n> n
```

```
--------------------
[sgdemo]
type = s3
provider = Other
access_key_id = ABCDEFGH123456789JKL
secret_access_key = 123456789ABCDEFGHIJKLMN0123456789PQRST+V
endpoint = sgdemo.netapp.com:443
--------------------
y) Yes this is OK (default)
e) Edit this remote
d) Delete this remote
y/e/d>
```

```
Current remotes:
```

```
Name                 Type
====                 ====
sgdemo               s3
```

```
e) Edit existing remote
n) New remote
d) Delete remote
r) Rename remote
c) Copy remote
s) Set configuration password
q) Quit config
e/n/d/r/c/s/q> q
```

Ejemplos de comandos básicos

- * Crear un cubo:*

  `rclone mkdir remote:bucket`

  ```
  # rclone mkdir sgdemo:test01
  ```

  Utilice --no-check-certificate si necesita ignorar los certificados SSL.

- **Lista todos los cucharones:**

```
rclone lsd remote:
```

```
# rclone lsd sgdemo:
```

- **Lista objetos en un cubo específico:**

```
rclone ls remote:bucket
```

```
# rclone ls sgdemo:test01
```

```
    65536 TestObject.0
    65536 TestObject.1
    65536 TestObject.10
    65536 TestObject.12
    65536 TestObject.13
    65536 TestObject.14
    65536 TestObject.15
    65536 TestObject.16
    65536 TestObject.17
    65536 TestObject.18
    65536 TestObject.2
    65536 TestObject.3
    65536 TestObject.5
    65536 TestObject.6
    65536 TestObject.7
    65536 TestObject.8
    65536 TestObject.9
  33554432 bigobj
      102 key.json
       47 locked01.txt
4294967296 sequential-read.0.0
       15 test.txt
      116 version.txt
```

- **Eliminar un cucharón:**

```
rclone rmdir remote:bucket
```

```
# rclone rmdir sgdemo:test02
```

- * Poner un objeto:*

```
rclone copy filename remote:bucket
```

```
# rclone copy ~/test/testfile.txt sgdemo:test01
```

- **Obtener un objeto:**

```
rclone copy remote:bucket/objectname filename
```

```
# Rclone copy sgdemo:test01/Testfile.txt ~/test/testfileS3.txt
```

- **Eliminar un objeto:**

```
rclone delete remote:bucket/objectname
```

```
# rclone delete sgdemo:test01/testfile.txt
```

- **Migrar objetos en un cubo**

```
rclone sync source:bucket destination:bucket --progress

rclone sync source_directory destination:bucket --progress
```

```
# rclone sync sgdemo:test01 sgdemo:clone01 --progress

Transferred:        4.032 GiB / 4.032 GiB, 100%, 95.484 KiB/s, ETA
0s
Transferred:           22 / 22, 100%
Elapsed time:       1m4.2s
```

Utilice --Progress o -P para mostrar el progreso de la tarea. De lo contrario, no habrá ninguna salida.

- * Eliminar un cubo y todo el contenido del objeto*

```
rclone purge remote:bucket --progress
```

rclone purge sgdemo:test01 --progress

```
Transferred:              0 B / 0 B, -, 0 B/s, ETA -
Checks:                46 / 46, 100%
Deleted:               23 (files), 1 (dirs)
Elapsed time:         10.2s
```

rclone ls sgdemo:test01

```
2023/04/14 09:40:51 Failed to ls: directory not found
```

Por Siegfried Hepp y Aron Klein

Prácticas recomendadas de StorageGRID para la puesta en marcha con Veeam Backup and Replication

Esta guía se centra en la configuración de NetApp StorageGRID y, en parte, de backups y replicación de Veeam. Este documento está dirigido a administradores de almacenamiento y redes que estén familiarizados con los sistemas Linux y que tengan la tarea de mantener o implementar un sistema NetApp StorageGRID en combinación con Veeam Backup and Replication.

Descripción general

Los administradores de almacenamiento buscan gestionar el crecimiento de sus datos con soluciones que cumplen con la disponibilidad, los objetivos de recuperación rápida, los escalan para satisfacer sus necesidades y automatizan su política para la retención a largo plazo de datos. Estas soluciones también deben proporcionar protección frente a pérdidas o ataques maliciosos. Juntos, Veeam y NetApp se han asociado para crear una solución de protección de datos que combina backup y recuperación de Veeam con NetApp StorageGRID para el almacenamiento de objetos on-premises.

Veeam y NetApp StorageGRID ofrecen una solución fácil de usar que trabajan conjuntamente para ayudar a satisfacer las demandas del rápido crecimiento de los datos y las crecientes regulaciones en todo el mundo. El almacenamiento de objetos basado en cloud es conocido por su resiliencia, capacidad de escalado, eficiencia operativa y de costes que lo convierten en la opción lógica como destino para sus backups. En este documento se proporcionan instrucciones y recomendaciones para la configuración de su solución Veeam Backup y del sistema StorageGRID.

La carga de trabajo de objetos de Veeam crea una gran cantidad de OPERACIONES simultáneas DE PUT, DELETE y LIST DE objetos pequeños. Habilitar la inmutabilidad se añadirá al número de solicitudes al almacén de objetos para establecer la retención y mostrar versiones. El proceso de un trabajo de copia de seguridad incluye la escritura de objetos para el cambio diario. Después de que las nuevas escrituras se hayan completado, el trabajo eliminará cualquier objeto basado en la política de retención de la copia de seguridad. La programación de las tareas de backup casi siempre se superpondrá. Este solapamiento dará como resultado una gran parte de la ventana de backup que consiste en 50/50 carga de trabajo PUT/DELETE en el almacén de objetos. Realizar ajustes en Veeam al número de operaciones simultáneas con la

configuración de la ranura de tareas, aumentar el tamaño del objeto aumentando el tamaño del bloque de trabajos de copia de seguridad, reduciendo el número de objetos en las solicitudes de eliminación de objetos múltiples, y elegir la ventana de tiempo máximo para completar los trabajos optimizará el rendimiento y el costo de la solución.

Asegúrese de leer la documentación del producto para "Veeam Backup and Replication" y.. "StorageGRID" antes de empezar. Veeam ofrece calculadoras para entender el tamaño de la infraestructura de Veeam y los requisitos de capacidad que se deben usar antes de dimensionar su solución de StorageGRID. Consulte siempre las configuraciones validadas de Veeam-NetApp en la página web del programa Veeam Ready para "Veeam Ready Object, inmutabilidad de objetos y repositorio".

Configuración de Veeam

Versión recomendada

Siempre se recomienda mantenerse al día y aplicar las revisiones más recientes para su sistema Veeam Backup & Replication 12. Actualmente recomendamos instalar como mínimo el parche de Veeam P20230718.

S3 Configuración del repositorio

Un repositorio de backup de escalado horizontal (SOBR) es el nivel de capacidad del almacenamiento de objetos S3. El nivel de capacidad es una extensión del repositorio primario que proporciona períodos de retención de datos más largos y una solución de almacenamiento de menor coste. Veeam ofrece la capacidad de proporcionar inmutabilidad a través de la API de bloqueo de objetos S3. Veeam 12 puede utilizar múltiples buckets en un repositorio de escalado horizontal. StorageGRID no tiene un límite en cuanto al número de objetos o de capacidad de un único bloque. El uso de varios bloques puede mejorar el rendimiento cuando se realizan backups de conjuntos de datos de gran tamaño donde los datos de backup pueden llegar a escalarse a petabytes en objetos.

En función del dimensionamiento de su solución y sus requisitos específicos, puede que sea necesario limitar las tareas simultáneas. La configuración predeterminada especifica una ranura de tareas de repositorio para cada núcleo de CPU y para cada ranura de tareas un límite de ranura de tareas simultáneas de 64. Por ejemplo, si el servidor tiene 2 núcleos de CPU, se utilizará un total de 128 subprocesos simultáneos para el almacén de objetos. Esto incluye PUT, GET y SUPR por lotes. Se recomienda seleccionar un límite conservador para las ranuras de tareas con las que empezar y ajustar este valor una vez que los backups de Veeam hayan alcanzado un estado constante de nuevos backups y datos de backup caducados. Trabaje con el equipo de su cuenta de NetApp para dimensionar el sistema de StorageGRID correctamente para satisfacer el rendimiento y los periodos de tiempo deseados. Es posible que sea necesario ajustar el número de ranuras de tareas y el límite de tareas por ranura para proporcionar la solución óptima.

Configuración de trabajos de copia de seguridad

Los trabajos de copia de seguridad de Veeam se pueden configurar con diferentes opciones de tamaño de bloque que se deben considerar cuidadosamente. El tamaño de bloque predeterminado es de 1MB KB y, con las eficiencias de almacenamiento que Veeam proporciona compresión y la deduplicación crea tamaños de objeto de aproximadamente 500KB KB para el backup completo inicial y objetos de 100-200kB MB para los trabajos incrementales. Podemos aumentar considerablemente el rendimiento y reducir los requisitos del almacén de objetos eligiendo un tamaño de bloque de backup mayor. Aunque el tamaño de bloque mayor realiza grandes mejoras en el rendimiento del almacén de objetos, se logra a costa de unos requisitos de capacidad de almacenamiento primario potencialmente mayores gracias a la reducción del rendimiento de la eficiencia del almacenamiento. Se recomienda que los trabajos de backup se configuren con un tamaño de bloque de 4MB KB que cree aproximadamente 2MB objetos para los backups completos y tamaños de objeto de 700KB a 1MB KB para los incrementales. Los clientes pueden considerar incluso configurar tareas de backup con un tamaño de bloque de 8 MB, que se puede habilitar con la ayuda del soporte de Veeam.

La implementación de backups inmutables utiliza S3 Object Lock en el almacén de objetos. La opción Inmutabilidad genera un mayor número de solicitudes al almacén de objetos para mostrar y actualizar la retención en los objetos.

A medida que las retenciones de copia de seguridad vencen, los trabajos de copia de seguridad procesarán la eliminación de objetos. Veeam envía las solicitudes de eliminación al almacén de objetos en solicitudes de eliminación de objetos múltiples de 1000 objetos por solicitud. Para soluciones pequeñas, esto puede necesitar ser ajustado para reducir el número de objetos por solicitud. Al reducir este valor, se añadirá la ventaja de distribuir de forma más uniforme las solicitudes de eliminación entre los nodos del sistema StorageGRID. Se recomienda utilizar los valores de la siguiente tabla como punto de partida en la configuración del límite de eliminación de objetos múltiples. Multiplique el valor de la tabla por el número de nodos para el tipo de dispositivo elegido para obtener el valor de la configuración en Veeam. Si este valor es igual o mayor que 1000, no es necesario ajustar el valor predeterminado. Si este valor necesita ser ajustado, por favor trabaje con el soporte de Veeam para hacer el cambio.

Modelo de dispositivo	S3MultiObjectDeleteLimit por nodo
SG5712	34
SG5760	75
SG6060	200

Trabaje en colaboración con su equipo de cuenta de NetApp para la configuración recomendada de acuerdo con sus necesidades específicas. Las recomendaciones de ajustes de configuración de Veeam incluirán:

- Tamaño del bloque de tareas de backup = 4MB
- Límite de ranura de tarea SOBR = 2-16
- Límite de eliminación de objetos múltiples = 34-1000

Configuración de StorageGRID

Versión recomendada

NetApp StorageGRID 11,6 o 11,7 con la revisión más reciente son las versiones recomendadas para implementaciones de Veeam. Muchas funciones de optimización se introdujeron en StorageGRID 11.6.0.11 y 11.7.0.4, que beneficiarán a las cargas de trabajo de Veeam. Siempre se recomienda estar al día y aplicar las correcciones urgentes más recientes para su sistema StorageGRID.

Configuración del balanceador de carga y del extremo S3

Veeam requiere que el punto final se conecte solo a través de HTTPS. Veeam no admite una conexión no cifrada. El certificado SSL puede ser un certificado autofirmado, una entidad de certificación privada de confianza o una entidad de certificación pública de confianza. Para garantizar el acceso continuo al repositorio de S3, se recomienda utilizar al menos dos equilibradores de carga en una configuración de alta disponibilidad. Los balanceadores de carga pueden ser un servicio de balanceador de carga integrado proporcionado por StorageGRID ubicado en cada nodo de administración y nodo de puerta de enlace o solución de terceros como F5, Kemp, HAProxy, Loadbalanacer.org, etc. El uso de un equilibrador de carga StorageGRID proporcionará la capacidad de establecer clasificadores de tráfico (reglas de QoS) que puedan priorizar la carga de trabajo de Veeam o limitar Veeam para no afectar a las cargas de trabajo de mayor prioridad en el sistema StorageGRID.

Bloque de S3

StorageGRID es un sistema de almacenamiento multi-tenancy seguro. Se recomienda crear un inquilino dedicado para la carga de trabajo de Veeam. Se puede asignar opcionalmente una cuota de almacenamiento. Como práctica recomendada habilitar “Usar fuente de identidad propia”. Proteger al usuario de gestión raíz de inquilinos con la contraseña adecuada. Veeam Backup 12 requiere una fuerte consistencia para bloques de S3. StorageGRID ofrece varias opciones de coherencia configuradas a nivel del bloque. Para implementaciones multi-sitio con Veeam accediendo a los datos desde múltiples ubicaciones, seleccione “strong-global”. Si los backups y restauraciones de Veeam se producen en un solo sitio, el nivel de consistencia debe establecerse en «sitio fuerte». Para obtener más información sobre los niveles de coherencia de bloques, revise la "documentación". Para usar StorageGRID para los backups de inmutabilidad de Veeam, el bloqueo de objetos de S3 debe habilitarse globalmente y configurarse en el bloque durante la creación del bloque.

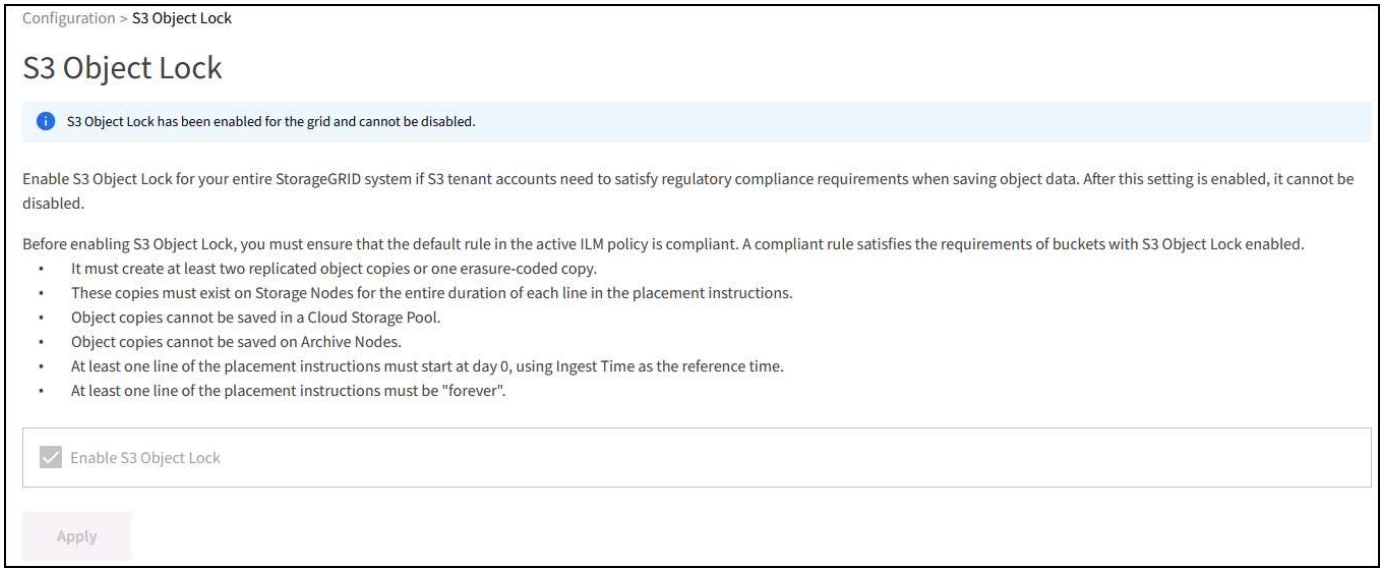
Gestión del ciclo de vida

StorageGRID admite la replicación y el código de borrado para la protección a nivel de objeto en nodos y sitios de StorageGRID. El código de borrado requiere un tamaño de objeto de 200kB KB como mínimo. El tamaño de bloque predeterminado para Veeam de 1MB produce tamaños de objeto que a menudo pueden estar por debajo de este tamaño mínimo recomendado de 200kB después de las eficiencias de almacenamiento de Veeam. Para mejorar el rendimiento de la solución, no se recomienda utilizar un perfil de código de borrado que abarque varios sitios a menos que la conectividad entre los sitios sea suficiente para no agregar latencia ni restringir el ancho de banda del sistema StorageGRID. En un sistema StorageGRID multisitio, la regla de gestión del ciclo de vida de la información se puede configurar para almacenar una sola copia en cada sitio. Para la máxima durabilidad, se puede configurar una regla para almacenar una copia codificada de borrado en cada sitio. El uso de dos copias locales en los servidores de Veeam Backup es la implementación más recomendada para esta carga de trabajo.

Puntos clave de implementación

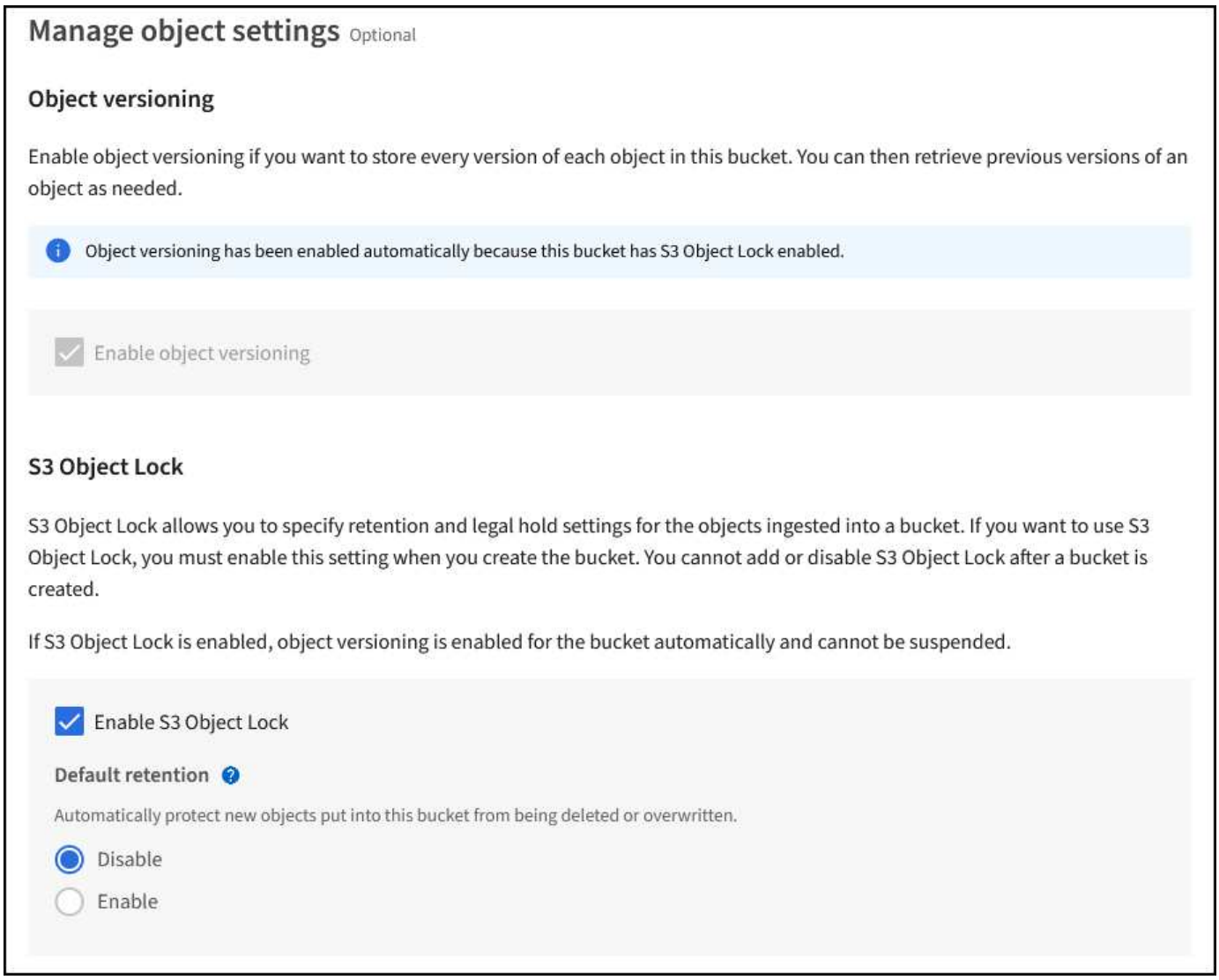
StorageGRID

Asegúrese de que el bloqueo de objetos está activado en el sistema StorageGRID si es necesario inmutabilidad. Busque la opción en la interfaz de usuario de administración en Configuration/S3 Object Lock.

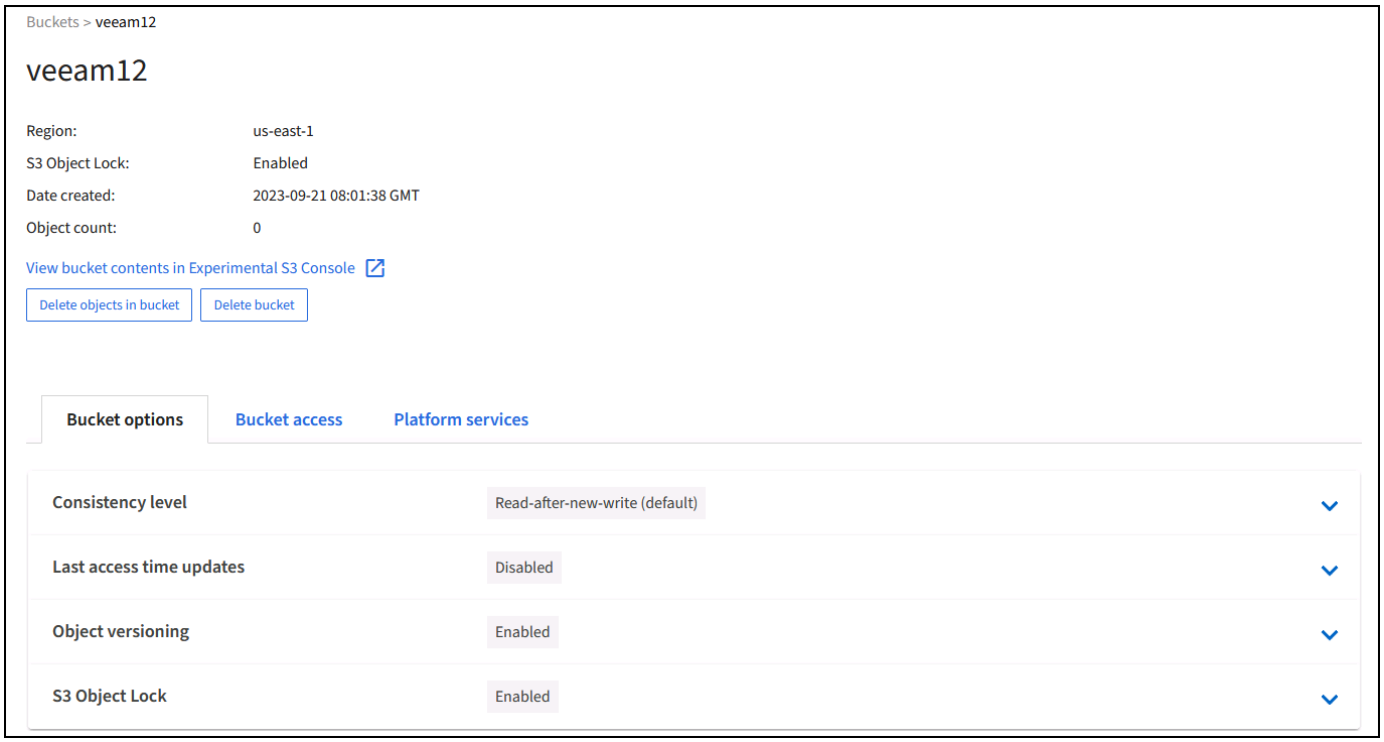


Al crear el bucket, seleccione «Enable Object Lock» (Habilitar bloqueo de objetos S3) si este bucket se va a utilizar para backups de inmutabilidad. Esto habilitará automáticamente el control de versiones de bloques.

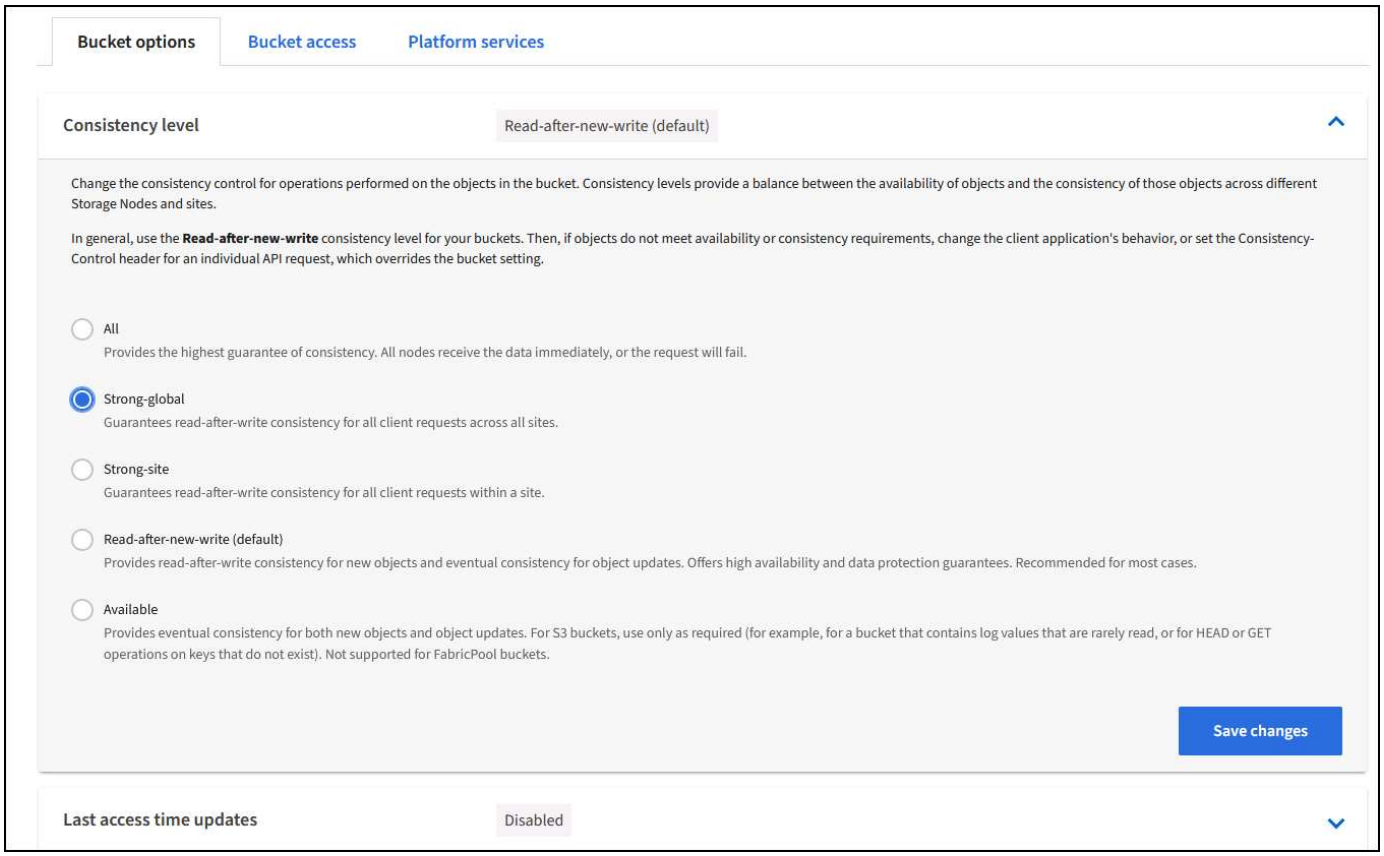
Deje desactivada la retención predeterminada, ya que Veeam establecerá la retención de objetos de forma explícita. El control de versiones y el bloqueo de objetos S3 no se deben seleccionar si Veeam no está creando copias de seguridad inmutables.



Una vez creado el bloque, vaya a la página de detalles del bloque creado. Seleccione el nivel de coherencia.



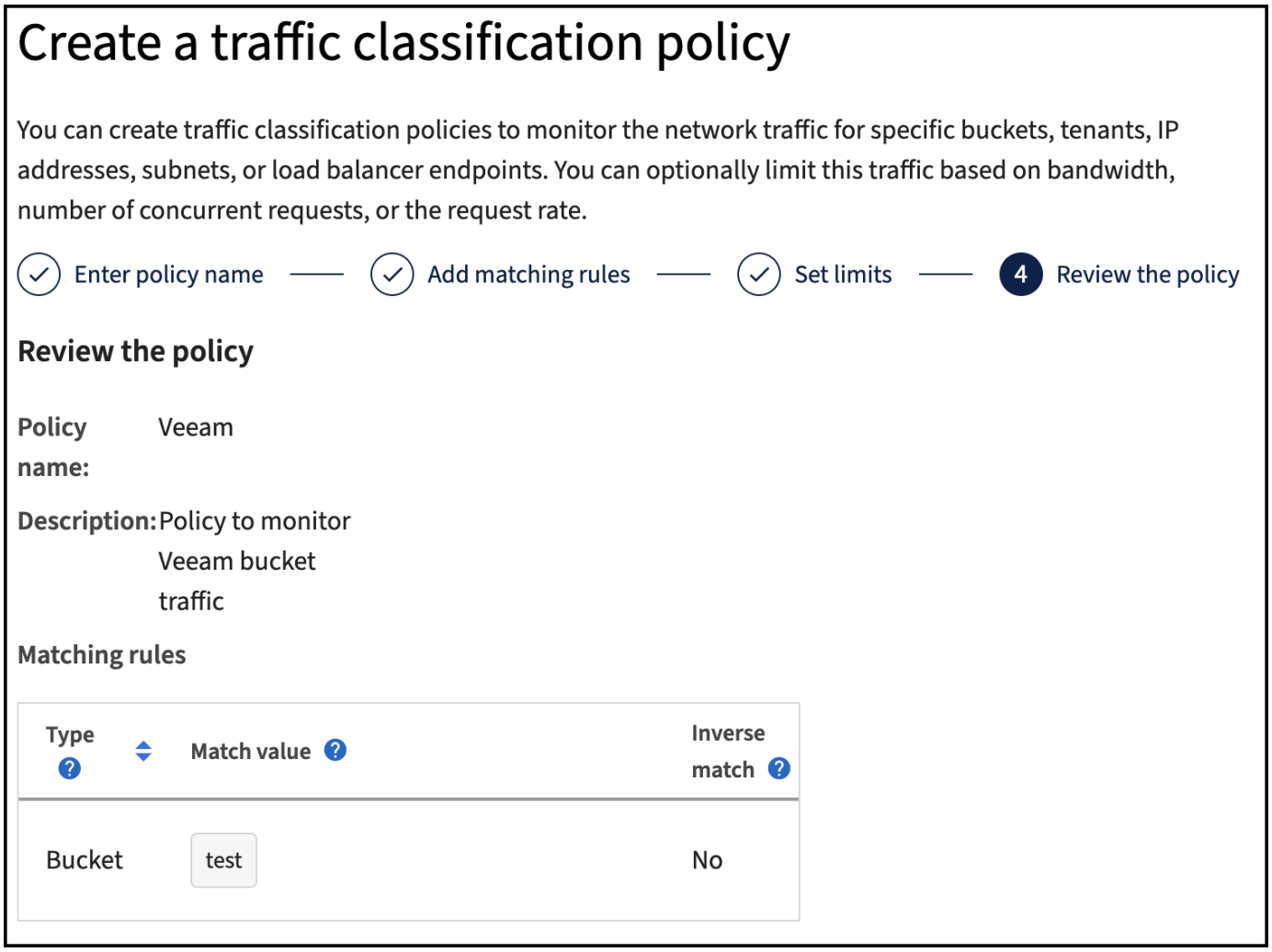
Veeam requiere una gran coherencia para bloques de S3. Por lo tanto, para implementaciones de múltiples sitios con Veeam accediendo a los datos desde múltiples ubicaciones, seleccione “strong-global”. Si los backups y restauraciones de Veeam se producen en un solo sitio, el nivel de consistencia debe establecerse en «sitio fuerte». Guarde los cambios.



StorageGRID ofrece un servicio de balanceo de carga integrado en cada nodo de administración y nodos de

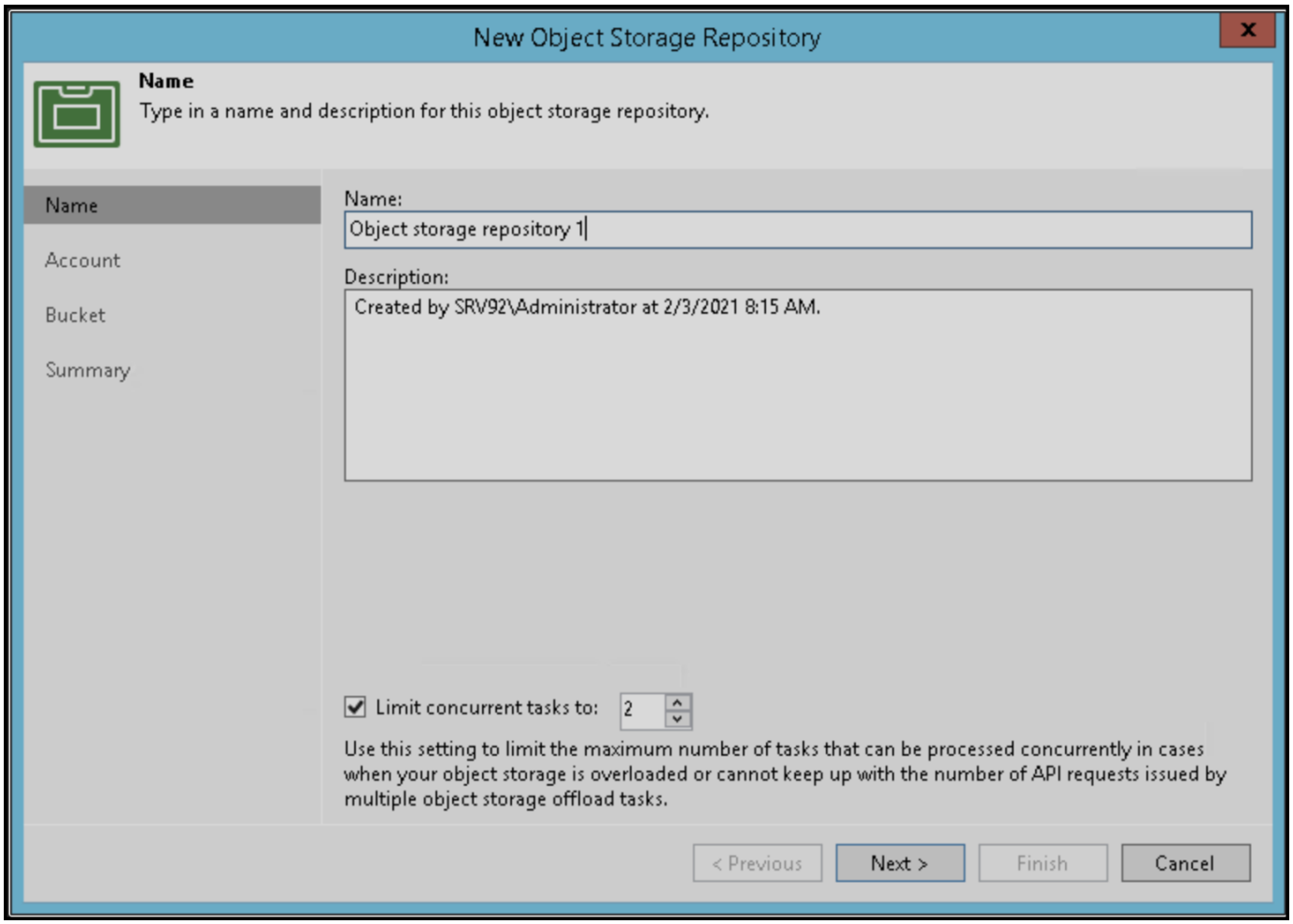
pasarela dedicados. Una de las muchas ventajas de usar este equilibrador de carga es la capacidad de configurar políticas de clasificación de tráfico (QoS). Aunque se utilizan principalmente para limitar el impacto de las aplicaciones en otras cargas de trabajo de clientes o priorizar una carga de trabajo sobre otras, también proporcionan una bonificación de la recopilación de métricas adicionales para ayudar a la supervisión.

En la pestaña de configuración, seleccione “Clasificación de tráfico” y cree una nueva política. Asigne un nombre a la regla y seleccione el tipo de cubo o arrendatario. Introduzca los nombres de los cubos o arrendatarios. Si la QoS es necesaria, establece un límite, pero para la mayoría de las implementaciones, solo queremos agregar los beneficios de monitoreo que proporciona, así que no establezcas un límite.

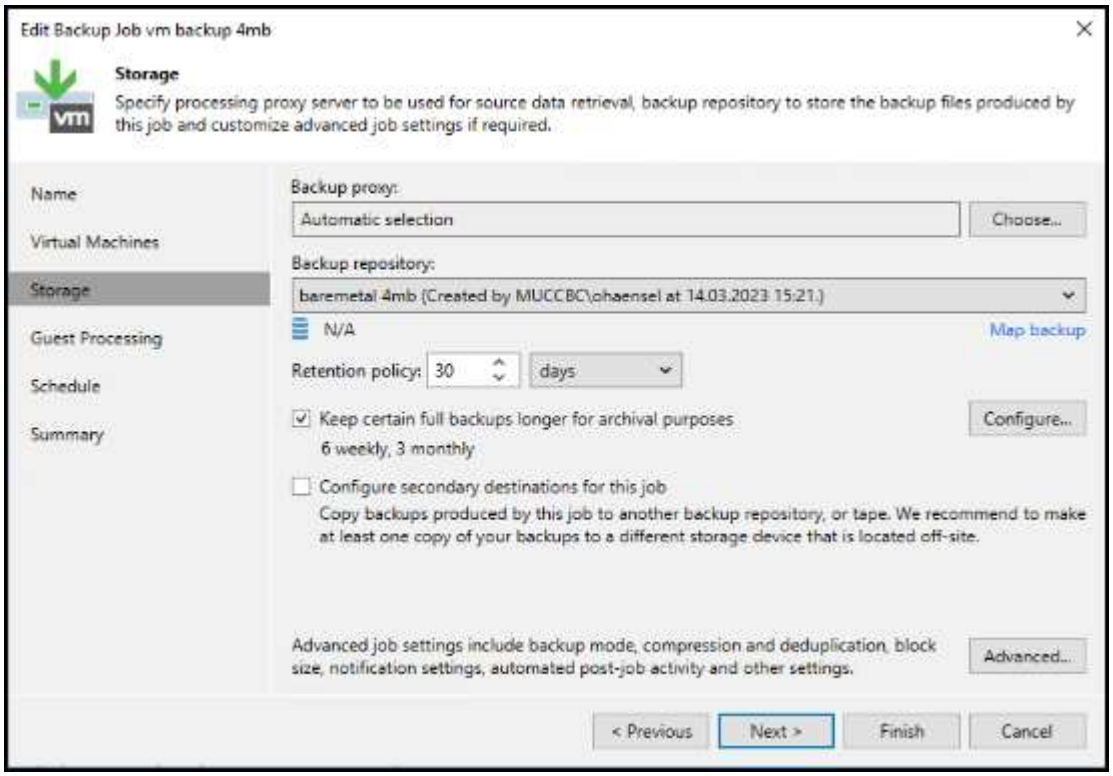


Veeam

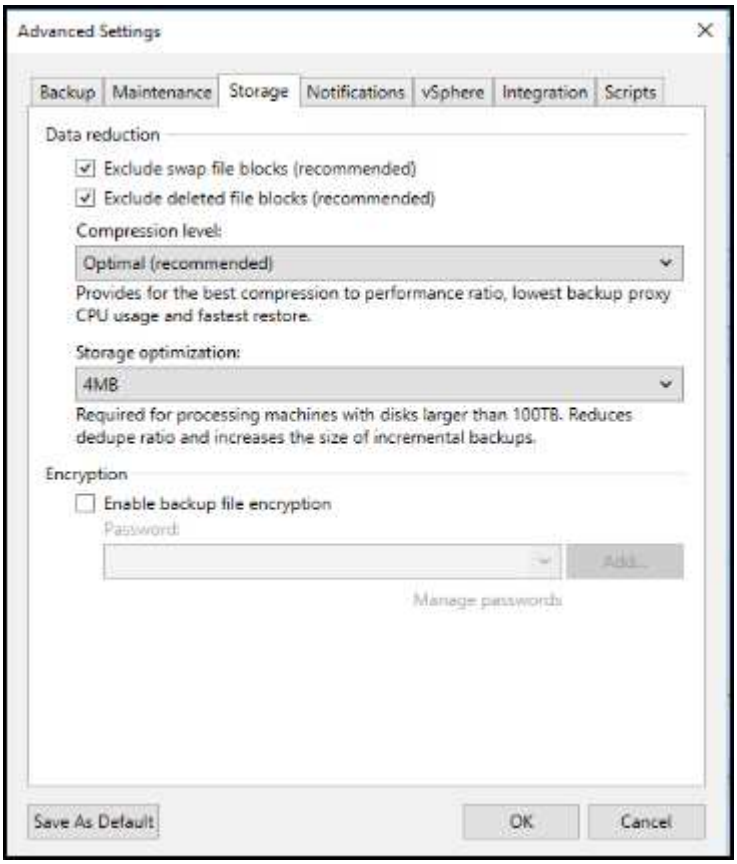
En función del modelo y la cantidad de dispositivos StorageGRID, puede que sea necesario seleccionar y configurar un límite para el número de operaciones simultáneas en el bloque.



Siga la documentación de Veeam sobre la configuración del trabajo de copia de seguridad en la consola de Veeam para iniciar el asistente. Después de agregar VM, seleccione el repositorio SOBR.



Haga clic en Configuración avanzada y cambie la configuración de optimización de almacenamiento a 4 MB o más. Se activará la compresión y la deduplicación. Cambie la configuración de invitado según sus requisitos y configure la programación de trabajos de copia de seguridad.



Supervisión de StorageGRID

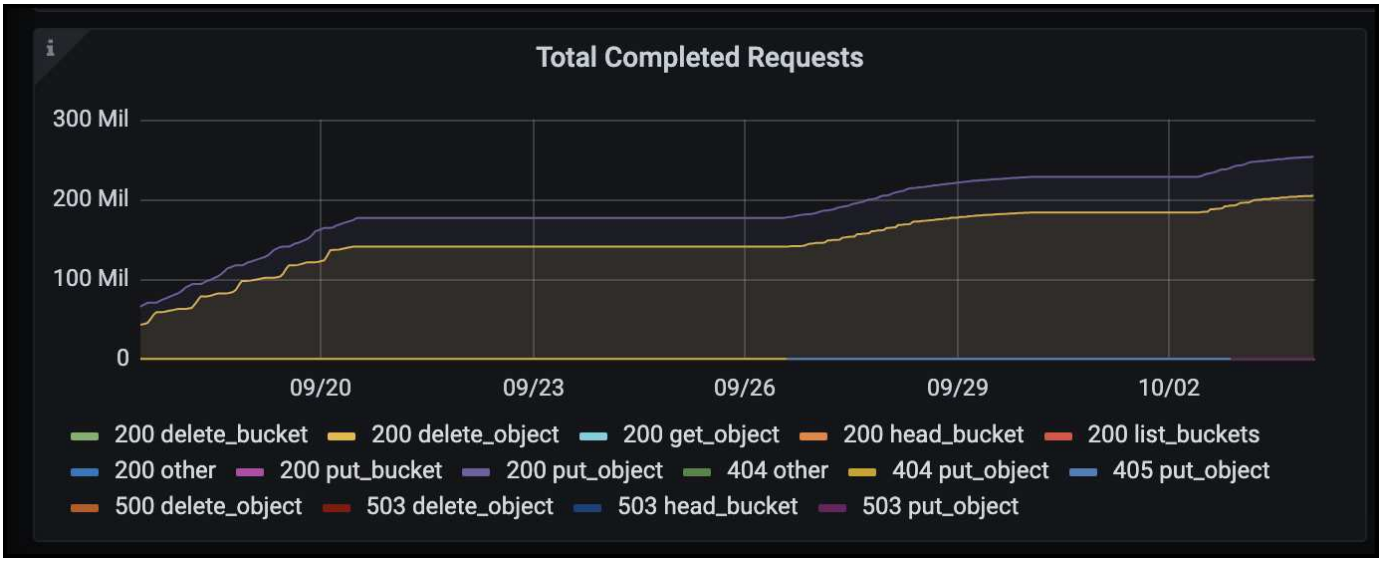
Para obtener una visión completa del rendimiento que están teniendo Veeam y StorageGRID juntos, deberá esperar hasta que haya caducado el tiempo de retención de los primeros backups. Hasta este punto, la carga de trabajo de Veeam consta principalmente de OPERACIONES PUT y no se han producido eliminaciones. Una vez que los datos de backup caducan y se producen limpiezas, ahora puede ver el uso consistente completo en el almacén de objetos y ajustar la configuración en Veeam si es necesario.

StorageGRID proporciona gráficos prácticos para supervisar el funcionamiento del sistema ubicado en la página Métricas de la pestaña Soporte. Los paneles principales a ver serán la información general de S3, ILM y la normativa de clasificación de tráfico si se crea una normativa. En el panel de Descripción general de S3 encontrará información sobre las tasas de funcionamiento de S3, las latencias y las respuestas de las solicitudes.

Si observa las tasas de S3 y las solicitudes activas, puede ver cuánta carga está manejando cada nodo y el número general de solicitudes por tipo.

El gráfico Duración Media muestra el tiempo medio que cada nodo está tomando para cada tipo de solicitud. Esta es la latencia media de la solicitud y puede ser un buen indicador de que se puede requerir un ajuste adicional, o hay espacio para que el sistema StorageGRID asuma más carga.

En el gráfico Total de Solicitudes Completadas, puede ver las solicitudes por tipo y códigos de respuesta. Si ve respuestas distintas de 200 (OK) para las respuestas, esto puede indicar un problema como que el sistema StorageGRID está recibiendo una carga elevada enviando respuestas 503 (Lento) y puede que sea necesario realizar algún ajuste adicional, o que haya llegado el momento de expandir el sistema para aumentar la carga.



En la consola de gestión de la vida útil de la información, puede supervisar el rendimiento de eliminación del sistema StorageGRID. StorageGRID usa una combinación de eliminaciones síncronas y asíncronas en cada nodo a fin de intentar optimizar el rendimiento general de todas las solicitudes.

Con una Política de clasificación de tráfico, podemos ver las métricas del rendimiento de la solicitud del equilibrador de carga, las tasas, la duración, así como los tamaños de los objetos que Veeam envía y recibe.

Dónde encontrar información adicional

Si quiere más información sobre el contenido de este documento, consulte los siguientes documentos o sitios web:

- "Documentación de producto de NetApp StorageGRID 11,7"
- "Veeam Backup and Replication"

Por Oliver Haensel y Aron Klein

Configurar el origen de datos de Dremio con StorageGRID

Dremio admite una variedad de fuentes de datos, incluido el almacenamiento de objetos en las instalaciones o basado en cloud. Puede configurar Dremio para que utilice StorageGRID como origen de datos de almacenamiento de objetos.

Configurar el origen de datos de Dremio

Requisitos previos

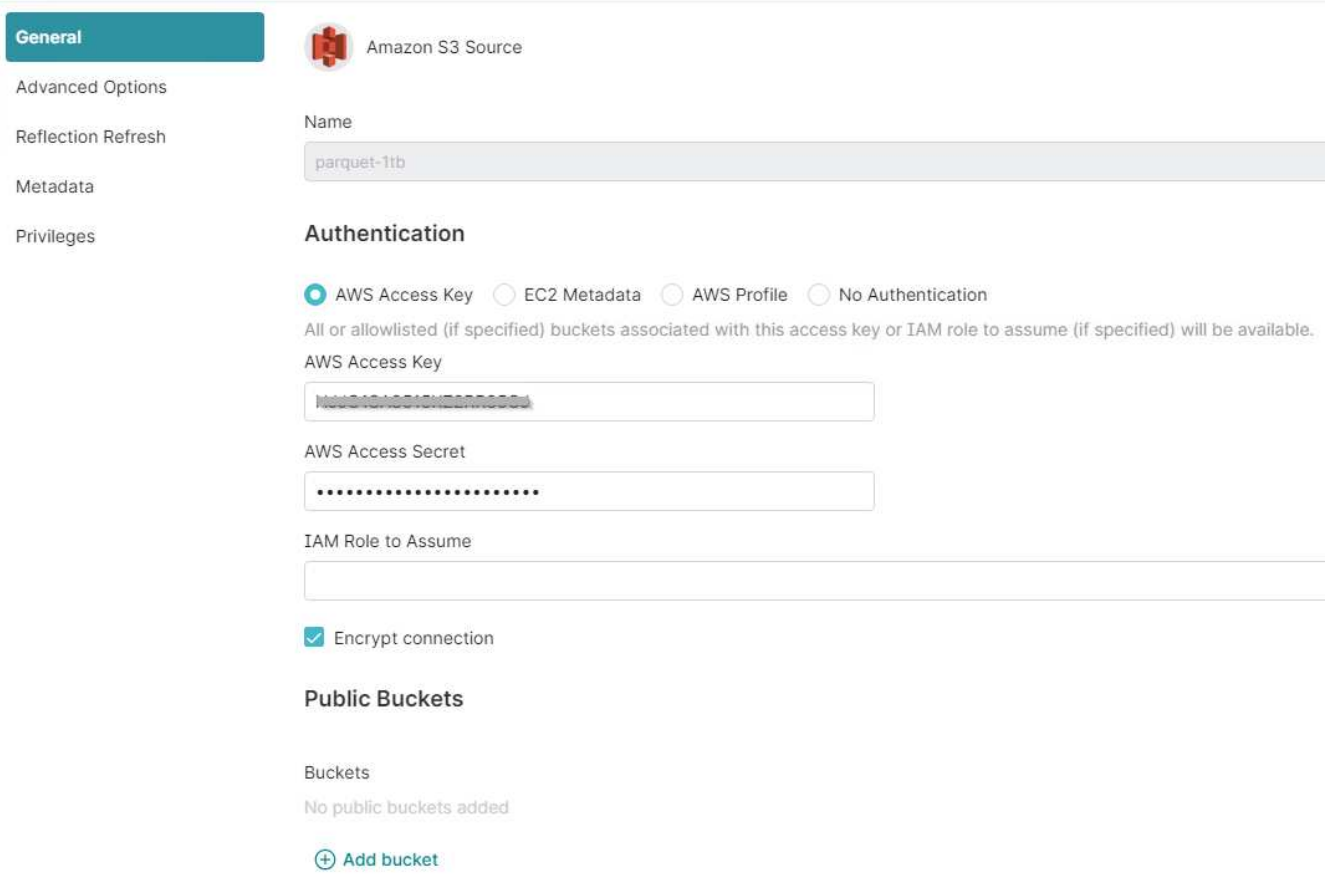
- Una URL de extremo de StorageGRID S3, un ID de clave de acceso de inquilino S3 y una clave de acceso secreta.
- Recomendación de configuración de StorageGRID: Deshabilitar la compresión (deshabilitada de forma predeterminada).
  Dremio utiliza el rango de bytes GET para recuperar diferentes rangos de bytes dentro del mismo objeto simultáneamente durante la consulta. El tamaño típico de las solicitudes de rango de bytes es 1MB. El objeto comprimido degrada el RENDIMIENTO DE LA OBTENCIÓN por rango de bytes.

Guía de Dremio

"Conexión a Amazon S3: Configuración de almacenamiento compatible con S3".

Instrucción

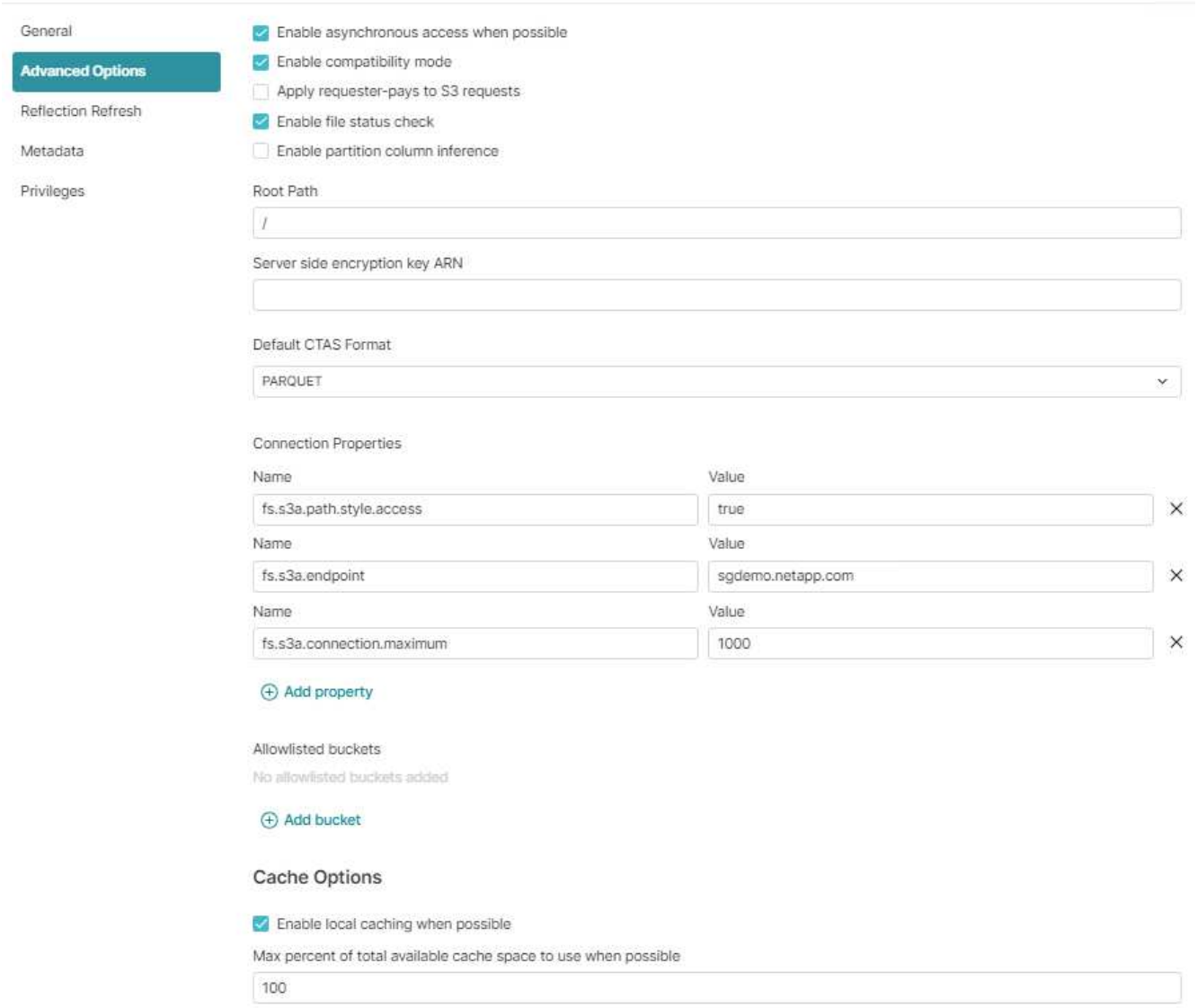
1. En la página Dremio Datasets, haga clic en el signo + para agregar una fuente, seleccione 'Amazon S3'.
2. Introduzca un nombre para este nuevo origen de datos, ID de clave de acceso de inquilino de StorageGRID S3 y clave de acceso secreta.
3. Active la casilla 'Cifrar conexión' si utiliza https para la conexión al punto final StorageGRID S3.
   Si utiliza el certificado de CA autofirmado para este punto final S3, siga las instrucciones de la guía de Dremio para agregar este certificado de CA al servidor <JAVA_HOME>/jre/lib/security + de Dremio
   Captura de pantalla de ejemplo



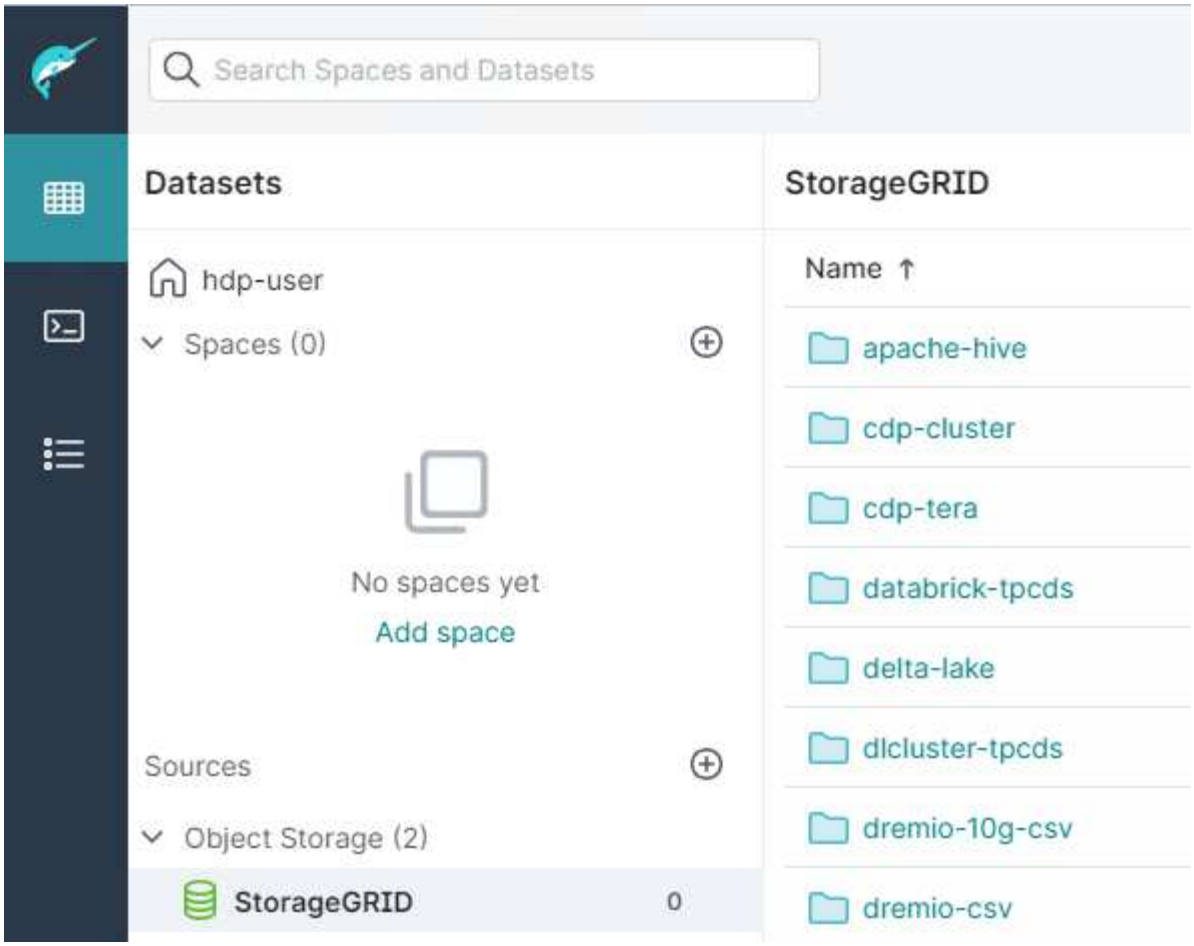
4. Haga clic en 'Opciones avanzadas', seleccione 'Activar modo de compatibilidad'
5. En Propiedades de conexión, haga clic en + Agregar propiedades y agregue estas S3A propiedades.
6. fs.s3a.connection.el valor por defecto máximo es 100. Si los conjuntos de datos S3 incluyen archivos de parquet grandes con 100 o más columnas, debe introducir un valor mayor que 100. Consulte la guía de Dremio para conocer este ajuste.

Nombre	Valor
fs.s3a.endpoint	*<StorageGRID S3 endpoint:Port>*
fs.s3a.path.style.access	verdadero
fs.s3a.conexión.máximo	*<un valor mayor que 100>*

Captura de pantalla de ejemplo



7. Configure otras opciones de Dremio según los requisitos de su organización o aplicación.
8. Haga clic en el botón Guardar para crear este nuevo origen de datos.
9. Una vez que el origen de datos StorageGRID se haya agregado correctamente, se mostrará una lista de cubos en el panel izquierdo.
   Captura de pantalla de ejemplo



Por Angela Cheng

NetApp StorageGRID con GitLab

NetApp ha probado StorageGRID con GitLab. Consulte la configuración de GitLab de ejemplo a continuación. Consulte "Guía de configuración de almacenamiento de objetos de GitLab" para obtener más detalles.

Ejemplo de conexión de almacenamiento de objetos

Para las instalaciones de Linux Package, este es un ejemplo de `connection` configuración en el formulario consolidado. Editar `/etc/gitlab/gitlab.rb` y agregue las siguientes líneas, sustituyendo los valores que desee:

```
# Consolidated object storage configuration
gitlab_rails['object_store']['enabled'] = true
gitlab_rails['object_store']['proxy_download'] = true
gitlab_rails['object_store']['connection'] = {
  'provider' => 'AWS',
  'region' => 'us-east-1',
  'endpoint' => 'https://<storagegrid-s3-endpoint:port>',
  'path_stype' => 'true',
  'aws_access_key_id' => '<AWS_ACCESS_KEY_ID>',
  'aws_secret_access_key' => '<AWS_SECRET_ACCESS_KEY>'
}
# OPTIONAL: The following lines are only needed if server side encryption
is required
gitlab_rails['object_store']['storage_options'] = {
  'server_side_encryption' => 'AES256'
}
gitlab_rails['object_store']['objects']['artifacts']['bucket'] = 'gitlab-
artifacts'
gitlab_rails['object_store']['objects']['external_diffs']['bucket'] =
'gitlab-mr-diffs'
gitlab_rails['object_store']['objects']['lfs']['bucket'] = 'gitlab-lfs'
gitlab_rails['object_store']['objects']['uploads']['bucket'] = 'gitlab-
uploads'
gitlab_rails['object_store']['objects']['packages']['bucket'] = 'gitlab-
packages'
gitlab_rails['object_store']['objects']['dependency_proxy']['bucket'] =
'gitlab-dependency-proxy'
gitlab_rails['object_store']['objects']['terraform_state']['bucket'] =
'gitlab-terraform-state'
gitlab_rails['object_store']['objects']['pages']['bucket'] = 'gitlab-
pages'
```