



Cómo habilitar StorageGRID en su entorno

How to enable StorageGRID in your environment

NetApp
April 26, 2024

Tabla de contenidos

- Cómo habilitar StorageGRID en su entorno 1
- Soluciones de terceros validadas 2
 - Soluciones de terceros validadas: Información general 2
 - Soluciones de terceros validadas con StorageGRID 11,8 2
 - Soluciones de terceros validadas con StorageGRID 11,7 4
 - Soluciones de terceros validadas de StorageGRID 11.6 7
 - Soluciones de terceros validadas de StorageGRID 11.5 9
 - Soluciones de terceros validadas de StorageGRID 11.4 11
 - Soluciones de terceros validadas de StorageGRID 11.3 13
 - Soluciones de terceros validadas de StorageGRID 11.2 15
- Guías de características de productos 17
 - Cree Cloud Storage Pool para AWS o Google Cloud 17
 - Cree Cloud Storage Pool para el almacenamiento BLOB de Azure 18
 - Utilice un pool de almacenamiento en el cloud para el backup 18
 - Configure el servicio de integración de búsqueda StorageGRID 19
 - Clon de nodo 35
 - Cómo utilizar el remap de puertos 38
 - Reubicación del sitio de grid y procedimiento de cambio de red en todo el sitio 49
- Guías de herramientas y aplicaciones 55
 - Utilice el conector Hadoop S3A de Cloudera con StorageGRID 55
 - Use S3cmd para probar y demostrar el acceso S3 en StorageGRID 62
 - Vertica Eon mode Database utilizando NetApp StorageGRID como almacenamiento comunitario 63
 - Análisis de registros de StorageGRID mediante pila ELK 77
 - Utilice Prometheus y Grafana para ampliar la retención de métricas 83
 - Configuración de SNMP de Datadog 99
 - Utilice rclone para migrar, PONER y ELIMINAR objetos en StorageGRID 102
 - Prácticas recomendadas de StorageGRID para la puesta en marcha con Veeam Backup and Replication 114
 - Configurar el origen de datos de Dremio con StorageGRID 125
 - NetApp StorageGRID con GitLab 128
- Procedimientos y ejemplos de API 130
 - Pruebe y muestre opciones de cifrado de S3 en StorageGRID 130
 - Pruebe y muestre el bloqueo de objetos de S3 en StorageGRID 133
 - Ejemplo de políticas de bloque y grupo(IAM) 138
- Informes técnicos 145
 - NetApp StorageGRID y análisis de Big Data 145
 - Ajuste Hadoop S3A 149
- Blogs de StorageGRID de NetApp 156
- Documentación de StorageGRID de NetApp 158
- Avisos legales 159
 - Derechos de autor 159
 - Marcas comerciales 159
 - Estadounidenses 159

Política de privacidad	159
Código abierto	159

Cómo habilitar StorageGRID en su entorno

Soluciones de terceros validadas

Soluciones de terceros validadas: Información general

NetApp, en colaboración con nuestros partners, ha validado estas soluciones para su uso con StorageGRID. Revise la información de esta sección para saber qué soluciones se han validado y para obtener instrucciones adicionales, si procede.

Una sus fuerzas a NetApp para acelerar la innovación en la cartera de productos, ampliar el reconocimiento de la Marca y aumentar las ventas al crear las mejores soluciones probadas de NetApp. ["Conviértase en un partner de la alianza"](#).

Soluciones de terceros validadas con StorageGRID 11,8

Las siguientes soluciones de terceros se han validado para el uso con StorageGRID 11,8.

Si la solución que busca no aparece en la lista, póngase en contacto con su representante de cuentas de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Alluxio
- Apache Kafka
- Punto de montaje de AWS
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- Collibra (Calidad de datos mínima de Collibra versión 2024,02)
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Datos de Diskover
- Dremio
- Imán
- Archivo de objetos de Fujifilm
- Servidor empresarial de GitHub

- IBM Filenet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Clústeres de Big Data de Microsoft SQL Server
- Modelo 9
- Modzy
- Moonwalk Universal
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- Archivo informativo de OpenText 16 EP7
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- PixitMedia ngenera
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Reveille v10, versión 220706 o superior
- Rubrik CDM
- s3a
- Signiant
- Copo de nieve
- Spectra Logic en el glaciar local
- Smartstore de Splunk
- Almacenamiento simplificado
- Trino
- Empresa de barniz 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Tejido de almacenamiento de Virtualica
- Weka v3.10 o posterior

Soluciones de terceros validadas en StorageGRID con bloqueo de objetos

Estas soluciones se han probado en colaboración con los partners correspondientes.

- CommVault 11 Feature Release 26
- IBM Filenet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 y posterior

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas con StorageGRID 11,7

Las siguientes soluciones de terceros se han validado para el uso con StorageGRID 11,7. + Si la solución que busca no está en esta lista, póngase en contacto con su representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Alluxio
- Apache Kafka
- Punto de montaje de AWS
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- Collibra (Calidad de datos mínima de Collibra versión 2024,02)
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Datos de Diskover
- Dremio
- Imán
- Archivo de objetos de Fujifilm
- Servidor empresarial de GitHub
- IBM Filenet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Clústeres de Big Data de Microsoft SQL Server
- Modelo 9
- Modzy
- Moonwalk Universal
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- Archivo informativo de OpenText 16 EP7
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura

- PixitMedia ngenera
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Reveille v10, versión 220706 o superior
- Rubrik CDM
- s3a
- Signiant
- Copo de nieve
- Spectra Logic en el glaciar local
- Smartstore de Splunk
- Almacenamiento simplificado
- Trino
- Empresa de barniz 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Tejido de almacenamiento de Virtualica
- Weka v3.10 o posterior

Soluciones de terceros validadas en StorageGRID con bloqueo de objetos

Estas soluciones se han probado en colaboración con los partners correspondientes.

- CommVault 11 Feature Release 26
- IBM Filenet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 y posterior

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360

- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas de StorageGRID 11.6

Las siguientes soluciones de terceros han sido validadas para su uso con StorageGRID 11.6. + Si la solución que busca no está en esta lista, póngase en contacto con su representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Alluxio
- Apache Kafka
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Datos de Diskover
- Dremio

- Imán
- Archivo de objetos de Fujifilm
- Servidor empresarial de GitHub
- IBM Filenet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Clústeres de Big Data de Microsoft SQL Server
- Modelo 9
- Modzy
- Moonwalk Universal
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- Archivo informativo de OpenText 16 EP7
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- PixitMedia ngenera
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Reveille v10, versión 220706 o superior
- Rubrik CDM
- s3a
- Signiant
- Copo de nieve
- Spectra Logic en el glaciar local
- Smartstore de Splunk
- Almacenamiento simplificado
- Trino
- Empresa de barniz 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x

- Vidispine
- Tejido de almacenamiento de Virtualica
- Weka v3.10 o posterior

Soluciones de terceros validadas en StorageGRID con bloqueo de objetos

Estas soluciones se han probado en colaboración con los partners correspondientes.

- CommVault 11 Feature Release 26
- IBM Filenet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 y posterior

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas de StorageGRID 11.5

Las siguientes soluciones de terceros han sido validadas para su uso con StorageGRID 11.5. + Si la solución que busca no está en esta lista, póngase en contacto con su

representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Alluxio
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Interica
- Komprise
- Moonwalk Universal
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- Archivo informativo de OpenText 16 EP7
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Rubrik CDM
- s3a
- Signiant
- Smartstore de Splunk
- Trino
- Empresa de barniz 6.0.4
- Veeam 11
- Veritas Enterprise Vault 11

- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Tejido de almacenamiento de Virtualica

Soluciones de terceros validadas en StorageGRID con bloqueo de objetos

Estas soluciones se han probado en colaboración con los partners correspondientes.

- OpenText Documentum 21.4
- Veeam 11

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas de StorageGRID 11.4

Las siguientes soluciones de terceros han sido validadas para su uso con StorageGRID 11.4. + Si la solución que busca no está en esta lista, póngase en contacto con su representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Interica
- Komprise
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- Archivo informativo de OpenText 16 EP7
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Rubrik CDM
- Signiant
- Smartstore de Splunk
- Empresa de barniz 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas de StorageGRID 11.3

Las siguientes soluciones de terceros han sido validadas para su uso con StorageGRID 11.3. + Si la solución que busca no está en esta lista, póngase en contacto con su representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX

- DefendX
- Interica
- Komprise
- MUJ BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- Smartstore de Splunk
- Empresa de barniz 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH

- SilverTrak
- SoftNAS
- QStar
- Velasea

Soluciones de terceros validadas de StorageGRID 11.2

Las siguientes soluciones de terceros han sido validadas para su uso con StorageGRID 11.2. + Si la solución que busca no está en esta lista, póngase en contacto con su representante de cuenta de NetApp.

Soluciones de terceros validadas en StorageGRID

Estas soluciones se han probado en colaboración con los partners correspondientes.

- Con razón
- Bridgestor
- Cantemo
- Colaboración de contenidos Citrix
- CommVault 11
- CTERA Portal 6
- Dalet
- Datadobi
- Dinámicas de datos StorageX
- DefendX
- Interica
- Komprise
- MUY BIEN
- Nasuní
- OpenText Documentum 16.4
- OpenText Media Management 16.5 con CyanGate Cloud
- Panzura
- Puerta de enlace de archivado de punto 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext cuántico 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- Smartstore de Splunk
- Empresa de barniz 6.0.4

- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Soluciones de terceros compatibles con StorageGRID

Estas soluciones han sido probadas.

- Software de archivado
- Comunicaciones de Axis
- Congruity360
- Marcos de datos
- Plataforma EcoDigital DIVA
- Encoding.com
- Archivo de objetos de Fujifilm
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Sistemas Milestone
- ONSSI
- Motor REACH
- SilverTrak
- SoftNAS
- QStar
- Velasea

Guías de características de productos

Cree Cloud Storage Pool para AWS o Google Cloud

Puede usar un pool de almacenamiento en cloud si desea mover objetos de StorageGRID a un bloque de S3 externo. El bloque externo puede pertenecer a Amazon S3 (AWS) o Google Cloud.

Lo que necesitará

- Se configuró StorageGRID 11.6.
- Ya ha configurado un bloque de S3 externo en AWS o Google Cloud.

Pasos

1. En Grid Manager, navegue hasta **ILM > agrupaciones de almacenamiento**.
2. En la sección Cloud Storage Pools de la página, seleccione **Crear**.

Se mostrará la ventana emergente Create Cloud Storage Pool.

3. Introduzca un nombre para mostrar.
4. Seleccione **Amazon S3** en la lista desplegable Tipo de proveedor.

Este tipo de proveedor funciona con AWS S3 o Google Cloud.

5. Introduzca el URI del bloque de S3 que se va a utilizar para el Cloud Storage Pool.

Se permiten dos formatos:

`https://host:port`

`http://host:port`

6. Introduzca el nombre de bloque de S3.

El nombre que especifique debe coincidir exactamente con el nombre del bloque de S3; de lo contrario, se producirá un error al crear el pool de almacenamiento en cloud. No se puede cambiar este valor después de guardar el pool de almacenamiento en cloud.

7. De manera opcional, introduzca el identificador de clave de acceso y la clave de acceso secreta.
8. Seleccione **no verificar certificado** en la lista desplegable.
9. Haga clic en **Guardar**.

Resultado esperado

Confirme que se ha creado un Cloud Storage Pool para Amazon S3 o Google Cloud.

Por Jonathan Wong

Cree Cloud Storage Pool para el almacenamiento BLOB de Azure

Puede usar un pool de almacenamiento en cloud si desea mover objetos de StorageGRID a un contenedor de Azure externo.

Lo que necesitará

- Se configuró StorageGRID 11.6.
- Ya ha configurado un contenedor de Azure externo.

Pasos

1. En Grid Manager, navegue hasta **ILM > agrupaciones de almacenamiento**.
2. En la sección Cloud Storage Pools de la página, seleccione **Crear**.

Se mostrará la ventana emergente Create Cloud Storage Pool.

3. Introduzca un nombre para mostrar.
4. Seleccione **Azure Blob Storage** en la lista desplegable Provider Type.
5. Introduzca el URI del bloque de S3 que se va a utilizar para el Cloud Storage Pool.

Se permiten dos formatos:

`https://host:port`

`http://host:port`

6. Introduzca el nombre del contenedor de Azure.

El nombre que especifique debe coincidir exactamente con el nombre del contenedor de Azure; de lo contrario, se producirá un error al crear el pool de almacenamiento en cloud. No se puede cambiar este valor después de guardar el pool de almacenamiento en cloud.

7. De forma opcional, introduzca el nombre de cuenta y la clave de cuenta asociados del contenedor de Azure para la autenticación.
8. Seleccione **no verificar certificado** en la lista desplegable.
9. Haga clic en **Guardar**.

Resultado esperado

Confirme que se ha creado un Cloud Storage Pool para el almacenamiento BLOB de Azure.

Por Jonathan Wong

Utilice un pool de almacenamiento en el cloud para el backup

Puede crear una regla de ILM para mover objetos a un pool de almacenamiento en el cloud para backup.

Lo que necesitará

- Se configuró StorageGRID 11.6.
- Ya ha configurado un contenedor de Azure externo.

Pasos

1. En Grid Manager, navegue hasta **ILM > Reglas > Crear**.
2. Introduzca una descripción.
3. Introduzca un criterio para activar la regla.
4. Haga clic en **Siguiente**.
5. Replique el objeto en nodos de almacenamiento.
6. Agregue una regla de colocación.
7. Replique el objeto en el pool de almacenamiento en cloud
8. Haga clic en **Siguiente**.
9. Haga clic en **Guardar**.

Resultado esperado

Confirmar que el diagrama de retención muestra los objetos almacenados localmente en StorageGRID y en un pool de almacenamiento en cloud para backup.

Confirme que, cuando se activa la regla de ILM, existe una copia en el Cloud Storage Pool y puede recuperar el objeto localmente sin realizar una restauración de objetos.

Por Jonathan Wong

Configure el servicio de integración de búsqueda StorageGRID

En esta guía, se ofrecen instrucciones detalladas para configurar el servicio de integración de búsquedas de NetApp StorageGRID 11.6 con Amazon OpenSearch Service o Elasticsearch en las instalaciones.

Introducción

StorageGRID admite tres tipos de servicios de plataforma.

- **Replicación CloudMirror de StorageGRID.** Reflejar objetos específicos desde un bloque de StorageGRID en un destino externo especificado.
- **Notificaciones.** Notificaciones de eventos por bloque para enviar notificaciones sobre acciones específicas realizadas en los objetos a un Amazon simple Notification Service (Amazon SNS) externo especificado.
- **Servicio de integración de búsqueda.** Envíe metadatos de objetos de simple Storage Service (S3) a un índice de Elasticsearch especificado, donde se pueden buscar o analizar los metadatos usando el servicio externo.

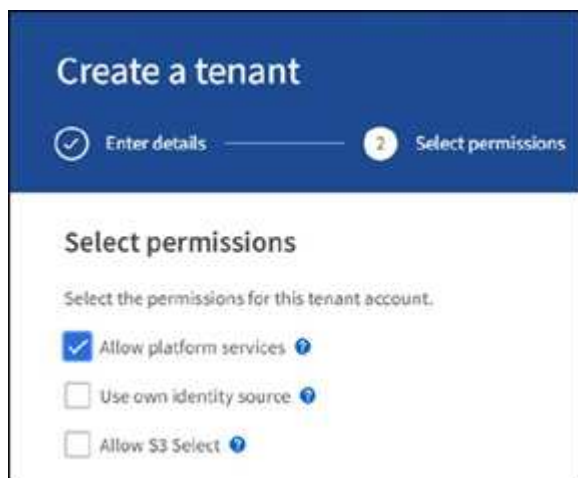
Los servicios de plataforma se configuran por el inquilino de S3 a través de la IU del administrador de inquilinos. Para obtener más información, consulte ["Consideraciones sobre el uso de servicios de plataforma"](#).

Este documento sirve como suplemento a la "[Guía de inquilinos de StorageGRID 11.6](#)" y proporciona instrucciones paso a paso y ejemplos para la configuración de endpoints y bloques para los servicios de integración de búsqueda. Las instrucciones de configuración de Amazon Web Services (AWS) o Elasticsearch en las instalaciones que se incluyen aquí son solo para fines de prueba o demostración básicos.

Las audiencias deben estar familiarizadas con Grid Manager y Tenant Manager y tener acceso al navegador S3 para realizar operaciones básicas de carga (PUT) y descarga (GET) para las pruebas de integración de búsqueda de StorageGRID.

Cree un inquilino y habilite los servicios de plataforma

1. Cree un inquilino S3 mediante Grid Manager, introduzca un nombre para mostrar y seleccione el protocolo S3.
2. En la página permisos, seleccione la opción permitir servicios de plataforma. Opcionalmente, seleccione otros permisos, si es necesario.



3. Configure la contraseña inicial del usuario raíz del inquilino o, si Identify federation está habilitada en la cuadrícula, seleccione el grupo federado con permiso de acceso raíz para configurar la cuenta de inquilino.
4. Haga clic en Iniciar sesión como raíz y seleccione cucharón: Crear y administrar cucharones.

Esto le lleva a la página Administrador de inquilinos.

5. En Tenant Manager, seleccione My Access Keys para crear y descargar la clave de acceso S3 para realizar pruebas posteriores.

Servicios de integración de búsqueda con Amazon OpenSearch

Configuración del servicio Amazon OpenSearch (antes Elasticsearch)

Utilice este procedimiento para una configuración rápida y sencilla del servicio OpenSearch sólo con fines de prueba/demostración. Si utiliza Elasticsearch en las instalaciones para los servicios de integración de búsqueda, consulte la sección [Servicios de integración de búsqueda con Elasticsearch en las instalaciones](#).



Debe tener un inicio de sesión de la consola de AWS válido, una clave de acceso, una clave de acceso secreta y permisos para suscribirse al servicio OpenSearch.

1. Cree un nuevo dominio siguiendo las instrucciones de "[Introducción al servicio AWS OpenSearch](#)",

excepto lo siguiente:

- Paso 4. Nombre de dominio: Sgdemo
- Paso 10. Control de acceso detallado: Anule la selección de la opción Habilitar control de acceso detallado.
- Paso 12. Política de acceso: Seleccione Configure Level Access Policy, seleccione la pestaña JSON para modificar la política de acceso mediante el ejemplo siguiente:
 - Reemplace el texto resaltado por su propio ID y nombre de usuario de gestión de acceso e identidades (IAM) de AWS.
 - Reemplace el texto resaltado (la dirección IP) por la dirección IP pública del equipo local que utilizó para acceder a la consola de AWS.
 - Abra una pestaña del navegador a "<https://checkip.amazonaws.com>" Para encontrar su IP pública.

```
{  
  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Effect": "Allow",  
      "Principal":  
        {"AWS": "arn:aws:iam:: nnnnnn:user/xyzabc"},  
      "Action": "es:*",  
      "Resource": "arn:aws:es:us-east-1:nnnnn:domain/sgdemo/*"  
    },  
    {  
      "Effect": "Allow",  
      "Principal": {"AWS": "*"},  
      "Action": [  
        "es:ESHttp*"  
      ],  
      "Condition": {  
        "IpAddress": {  
          "aws:SourceIp": [ "nnn.nnn.nn.n/nn"  
        ]  
      }  
    },  
    "Resource": "arn:aws:es:us-east-1:nnnnn:domain/sgdemo/*"  
  ]  
}
```


Fine-grained access control

Fine-grained access control provides numerous features to help you keep your data secure. Features include document-level security, field-level security, read-only users, and OpenSearch Dashboards/Kibana tenants. Fine-grained access control requires a master user. [Learn more](#)

Enable fine-grained access control

SAML authentication for OpenSearch Dashboards/Kibana

SAML authentication lets you use your existing identity provider for single sign-on for OpenSearch Dashboards/Kibana. [Learn more](#)

■ Prepare SAML authentication

To use SAML authentication, you must first enable fine-grained access control.

Amazon Cognito authentication

Enable to use Amazon Cognito authentication for OpenSearch Dashboards/Kibana. Amazon Cognito supports a variety of identity providers for username-password authentication. [Learn more](#)

Enable Amazon Cognito authentication

Access policy

Access policies control whether a request is accepted or rejected when it reaches the Amazon OpenSearch Service domain. If you specify an account, user, or role in this policy, you must sign your requests. [Learn more](#)

Domain access policy

- Only use fine-grained access control
Allow open access to the domain.
- Do not set domain level access policy
All requests to the domain will be denied.
- Configure domain level access policy**

Visual editor

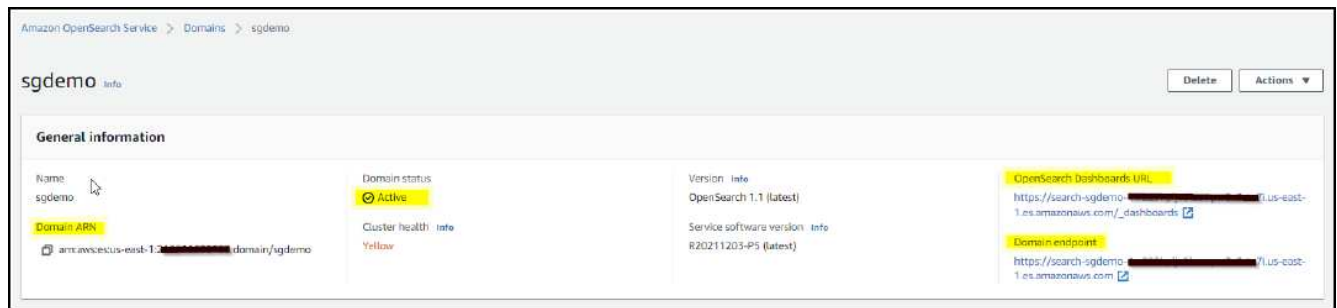
JSON

Import policy

Access policy

```
3-   "Statement": [  
4-     {  
5-       "Effect": "Allow",  
6-       "Principal": {  
7-         "AWS": "arn:aws:iam::222222222222:user/ashwin"  
8-       },  
9-       "Action": "es:*",  
10-      "Resource": "arn:aws:es:us-east-1:222222222222:domain/sgdemo/**"  
11-    },  
12-    {  
13-      "Effect": "Allow",  
14-      "Principal": {  
15-        "AWS": "*"   
16-      },  
17-      "Action": [  
18-        "es:ESHttpPost"  
19-      ],  
20-      "Condition": {  
21-        "IpAddress": {  
22-          "aws:SourceIp": [  
23-            "216.24.24.0/24"  
24-          ]  
25-        }  
26-      },  
27-      "Resource": "arn:aws:es:us-east-1:222222222222:domain/sgdemo/**"  
28-    }  
  ]  
}
```

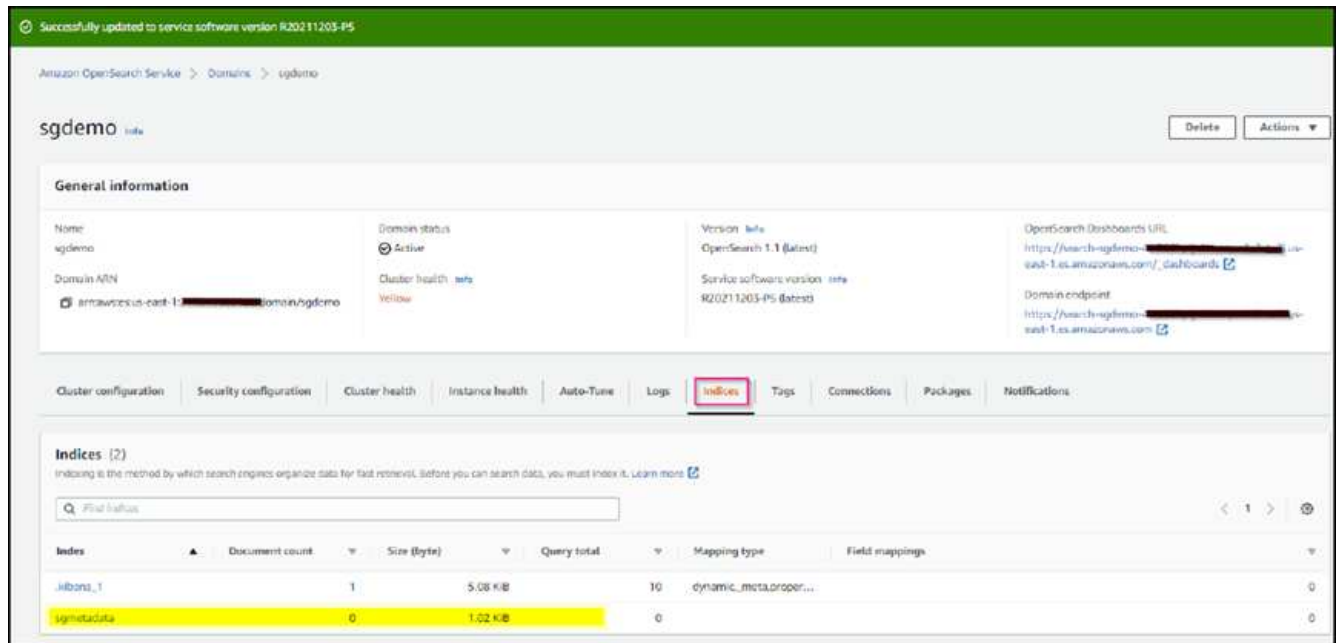
2. Espere de 15 a 20 minutos para que el dominio se active.



3. Haga clic en OpenSearch Dashboards URL para abrir el dominio en una nueva pestaña para tener acceso al panel. Si obtiene un error de acceso denegado, compruebe que la dirección IP de origen de la directiva de acceso esté correctamente configurada en la IP pública del equipo para permitir el acceso al panel de control de dominio.
4. En la página de bienvenida del panel, seleccione explorar por su cuenta. En el menú, vaya a Management → Dev Tools
5. En Herramientas de desarrollo → Consola , escriba `PUT <index>` Donde se usa el índice para almacenar metadatos de objetos StorageGRID. Utilizamos el nombre de índice 'gmetadata' en el siguiente ejemplo. Haga clic en el símbolo de triángulo pequeño para ejecutar el comando PUT. El resultado esperado se muestra en el panel derecho como se muestra en la siguiente captura de pantalla de ejemplo.



6. Verifique que el índice sea visible desde la IU de Amazon OpenSearch en `sgdomain > Indices`.



Configuración de extremos de servicios de plataforma

Para configurar los extremos de servicios de la plataforma, siga estos pasos:

1. En el administrador de inquilinos, vaya a ALMACENAMIENTO (S3) > extremos de servicios de la plataforma.
2. Haga clic en Create Endpoint, introduzca lo siguiente y haga clic en Continue:
 - Ejemplo de nombre para mostrar `aws-opensearch`
 - El extremo de dominio en la captura de pantalla de ejemplo bajo el paso 2 del procedimiento anterior en el campo URI.
 - El dominio ARN utilizado en el paso 2 del procedimiento anterior en el campo URN y agregue `<index>/_doc` Al final de ARN.

En este ejemplo, URN se convierte en `arn:aws:es:us-east-1:211234567890:domain/sgdemo/sgmetadata/_doc`.

Create endpoint

Enter details
 2 Select authentication type Optional
 Verify server Optional

Authentication type ?

Select the method used to authenticate connections to the endpoint.

Access Key ▼

Access key ID ?

AKIA[REDACTED]UWO

Secret access key ?

[REDACTED] 👁

[Previous](#) [Continue](#)

- Para verificar el punto final, seleccione usar certificado CA del sistema operativo y probar y crear punto final. Si la verificación se realiza correctamente, aparece una pantalla de extremo similar a la siguiente figura. Si se produce un error de verificación, compruebe que URN incluya `/<index>/_doc` Al final de la ruta, la clave de acceso y la clave secreta de AWS son correctas.

Platform services endpoints

A platform services endpoint stores the information StorageGRID needs to use an external resource as a target for a platform service (CloudMirror replication, notifications, or search integration). You must configure an endpoint for each platform service you plan to use.

1 endpoint [Create endpoint](#)

[Delete endpoint](#)

<input type="checkbox"/>	Display name ?	Last error ?	Type ?	URI ?	URN ?
<input type="checkbox"/>	aws-opensearch		Search	https://search-sgdemo-1.es.amazonaws.com/	arn:aws:es:us-east-1:[REDACTED]:domain/sgdemo/sgmetadata/_doc

Servicios de integración de búsqueda con Elasticsearch en las instalaciones

Configuración de Elasticsearch en las instalaciones

Este procedimiento es para una configuración rápida de Elasticsearch en las instalaciones y Kibana usando docker solo para fines de pruebas. Si ya existe el servidor Elasticsearch y Kibana, vaya al paso 5.

1. Siga este ["Procedimiento de instalación de Docker"](#) para instalar el docker. Utilizamos la ["Procedimiento de instalación de CentOS Docker"](#) en esta configuración.

```
sudo yum install -y yum-utils
sudo yum-config-manager --add-repo
https://download.docker.com/linux/centos/docker-ce.repo
sudo yum install docker-ce docker-ce-cli containerd.io
sudo systemctl start docker
```

- Para iniciar docker después del reinicio, introduzca lo siguiente:

```
sudo systemctl enable docker
```

- Ajuste la `vm.max_map_count` valor hasta 262144:

```
sysctl -w vm.max_map_count=262144
```

- Para mantener el ajuste después del reinicio, introduzca lo siguiente:

```
echo 'vm.max_map_count=262144' >> /etc/sysctl.conf
```

2. Siga la ["Guía de inicio rápido de Elasticsearch"](#) Sección autogestionada para instalar y ejecutar Elasticsearch y Kibana docker. En este ejemplo, instalamos la versión 8.1.



Tenga en cuenta el nombre de usuario/contraseña y el token creados por Elasticsearch, necesita esos elementos para iniciar la autenticación del extremo de la plataforma StorageGRID y la interfaz de usuario de Kibana.

Install and run Elasticsearch

1. Install and start [Docker Desktop](#).
2. Run:

```
docker network create elastic
docker pull docker.elastic.co/elasticsearch/elasticsearch:8.1.0
docker run --name es-node01 --net elastic -p 9200:9200 -p 9300:9300 -it
```

When you start Elasticsearch for the first time, the following security configuration occurs automatically:

- [Certificates and keys](#) are generated for the transport and HTTP layers.
- The Transport Layer Security (TLS) configuration settings are written to `elasticsearch.yml`.
- A password is generated for the `elastic` user.
- An enrollment token is generated for Kibana.



You might need to scroll back a bit in the terminal to view the password and enrollment token.

3. Copy the generated password and enrollment token and save them in a secure location. These values are shown only when you start Elasticsearch for the first time. You'll use these to enroll Kibana with your Elasticsearch cluster and log in.



If you need to reset the password for the `elastic` user or other built-in users, run the [elasticsearch-reset-password](#) tool. To generate new enrollment tokens for Kibana or Elasticsearch nodes, run the [elasticsearch-create-enrollment-token](#) tool. These tools are available in the Elasticsearch `bin` directory.

Install and run Kibana

To analyze, visualize, and manage Elasticsearch data using an intuitive UI, install Kibana.

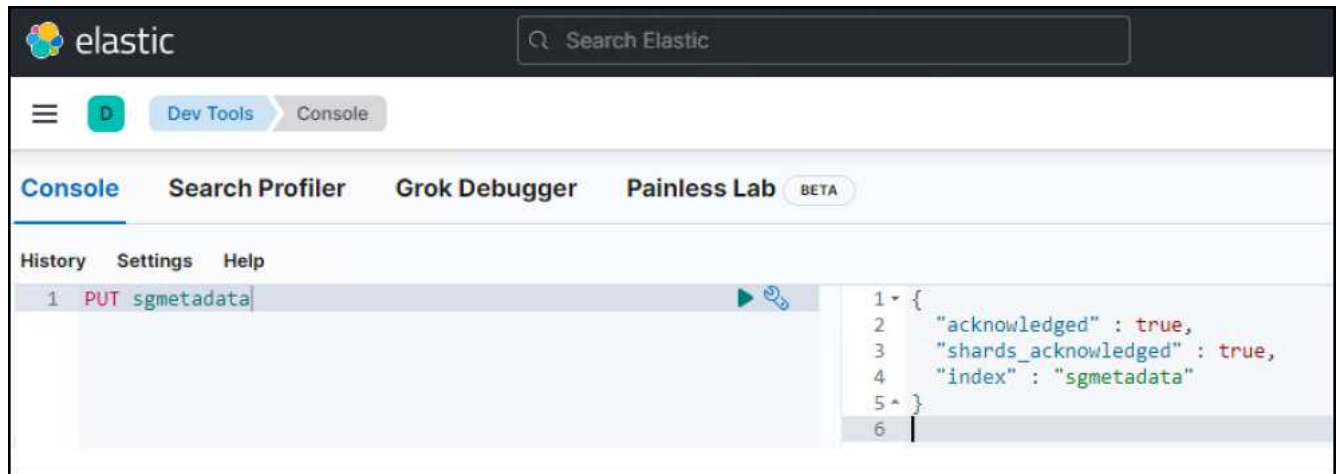
1. In a new terminal session, run:

```
docker pull docker.elastic.co/kibana/kibana:8.1.0
docker run --name kib-01 --net elastic -p 5601:5601 docker.elastic.co/k
```

When you start Kibana, a unique link is output to your terminal.

2. To access Kibana, click the generated link in your terminal.
 - a. In your browser, paste the enrollment token that you copied and click the button to connect your Kibana instance with Elasticsearch.
 - b. Log in to Kibana as the `elastic` user with the password that was generated when you started Elasticsearch.

- Después de que se haya iniciado el contenedor de Docker de Kibana, el enlace de URL `https://0.0.0.0:5601` aparecen en la consola. Reemplace 0.0.0.0 por la dirección IP del servidor en la dirección URL.
- Inicie sesión en la interfaz de usuario de Kibana con el nombre de usuario `elastic` Y la contraseña generada por Elastic en el paso anterior.
- Para iniciar sesión por primera vez, en la página de bienvenida del panel, seleccione explorar por su cuenta. En el menú, seleccione Management > Dev Tools.
- En la pantalla Dev Tools Console, introduzca `PUT <index>` Dónde se usa este índice para almacenar metadatos de objetos StorageGRID. Usamos el nombre del índice `sgmetadata` en este ejemplo. Haga clic en el símbolo de triángulo pequeño para ejecutar el comando PUT. El resultado esperado se muestra en el panel derecho como se muestra en la siguiente captura de pantalla de ejemplo.



Configuración de extremos de servicios de plataforma

Para configurar extremos para servicios de plataforma, siga estos pasos:

- En el Administrador de inquilinos, vaya a ALMACENAMIENTO (S3) > extremos de servicios de la plataforma
- Haga clic en Create Endpoint, introduzca lo siguiente y haga clic en Continue:
 - Ejemplo de nombre para mostrar: `elasticsearch`
 - URI: `https://<elasticsearch-server-ip or hostname>:9200`
 - URN: `urn:<something>:es:::<some-unique-text>/<index-name>/_doc` Donde el nombre de índice es el nombre que utilizó en la consola de Kibana. Ejemplo:
`urn:local:es:::sgmd/sgmetadata/_doc`

Create endpoint

1 Enter details — 2 Select authentication type Optional — 3 Verify server Optional

Enter endpoint details

Enter the endpoint's display name, URI, and URN.

Display name [?](#)

URI [?](#)

URN [?](#)

Cancel [Continue](#)

3. Seleccione HTTP básico como tipo de autenticación, introduzca el nombre de usuario `elastic` y la contraseña generada por el proceso de instalación de Elasticsearch. Para ir a la página siguiente, haga clic en continuar.

Authentication type [?](#)

Select the method used to authenticate connections to the endpoint.

Basic HTTP [v](#)

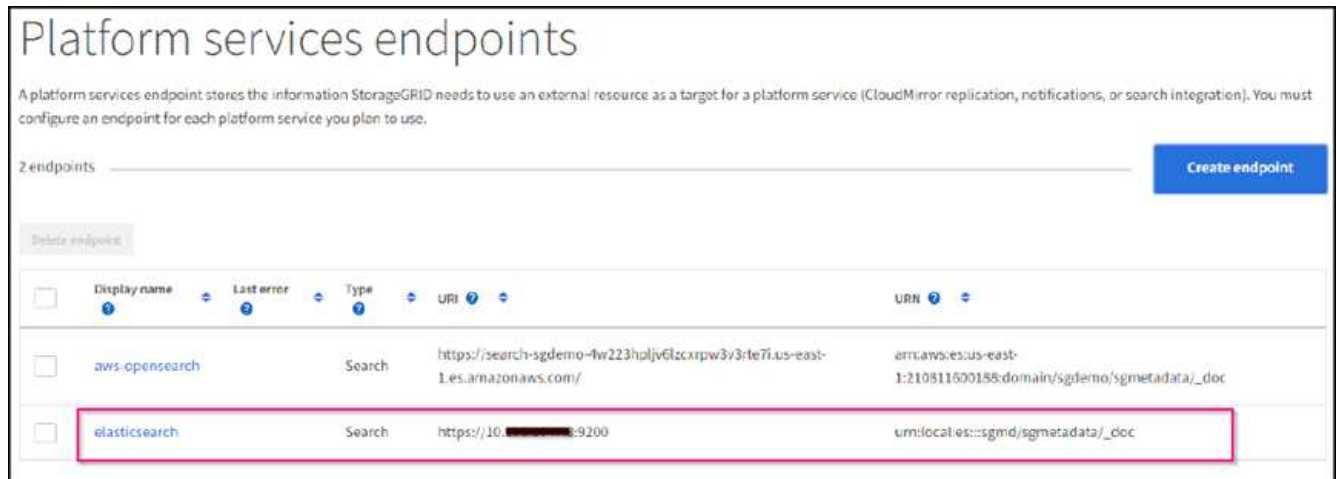
Username [?](#)

Password [?](#)

 [v](#)

Previous [Continue](#)

4. Seleccione no verificar certificado y probar y Crear extremo para verificar el extremo. Si la verificación se realiza correctamente, aparece una pantalla de punto final similar a la siguiente captura de pantalla. Si se produce un error en la verificación, compruebe que las entradas de URN, URI y nombre de usuario/contraseña sean correctas.



Configuración del servicio de integración de búsqueda de bloques

Una vez creado el extremo de servicio de la plataforma, el siguiente paso es configurar este servicio a nivel de bloque para enviar metadatos de objetos al extremo definido cada vez que se crea, se elimina o se actualizan sus metadatos o etiquetas.

Puede configurar la integración de búsqueda mediante el Administrador de inquilinos para aplicar un XML de configuración de StorageGRID personalizado a un bloque de la siguiente forma:

1. En el administrador de inquilinos, vaya a STORAGE(S3) > Buckets
2. Haga clic en Create Bucket, introduzca el nombre del bloque (por ejemplo, sgmetadata-test) y acepte el valor predeterminado us-east-1 región.
3. Haga clic en Continue > Create Bucket.
4. Para abrir la página bucket Overview, haga clic en el nombre del bloque y, a continuación, seleccione Platform Services.
5. Seleccione el cuadro de diálogo Habilitar integración de búsqueda. En el cuadro XML proporcionado, introduzca el XML de configuración mediante esta sintaxis.

El URN resaltado debe coincidir con el extremo de servicios de plataforma definido. Puede abrir otra pestaña del explorador para acceder al administrador de inquilinos y copiar el URN desde el extremo de servicios de plataforma definido.

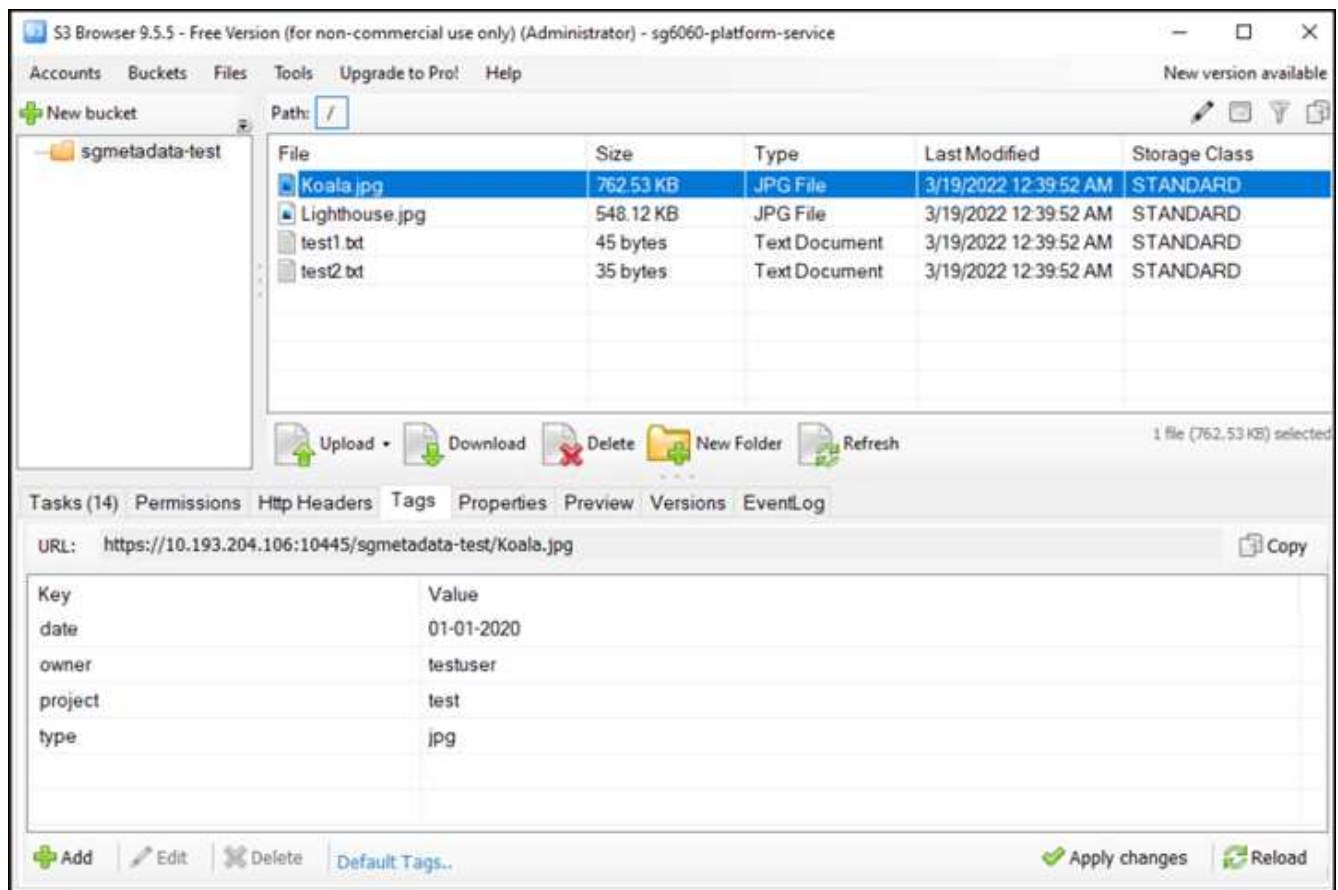
En este ejemplo, no hemos utilizado ningún prefijo, lo que significa que los metadatos de cada objeto de este bloque se envían al extremo de Elasticsearch definido previamente.

```

<MetadataNotificationConfiguration>
  <Rule>
    <ID>Rule-1</ID>
    <Status>Enabled</Status>
    <Prefix></Prefix>
    <Destination>
      <Urn> urn:local:es:::sgmd/sgmetadata/_doc</Urn>
    </Destination>
  </Rule>
</MetadataNotificationConfiguration>

```

6. Utilice el navegador S3 para conectarse a StorageGRID con la clave secreta/acceso de inquilino y cargar objetos de prueba a. sgmetadata-test agrupe y añada etiquetas o metadatos personalizados a los objetos.



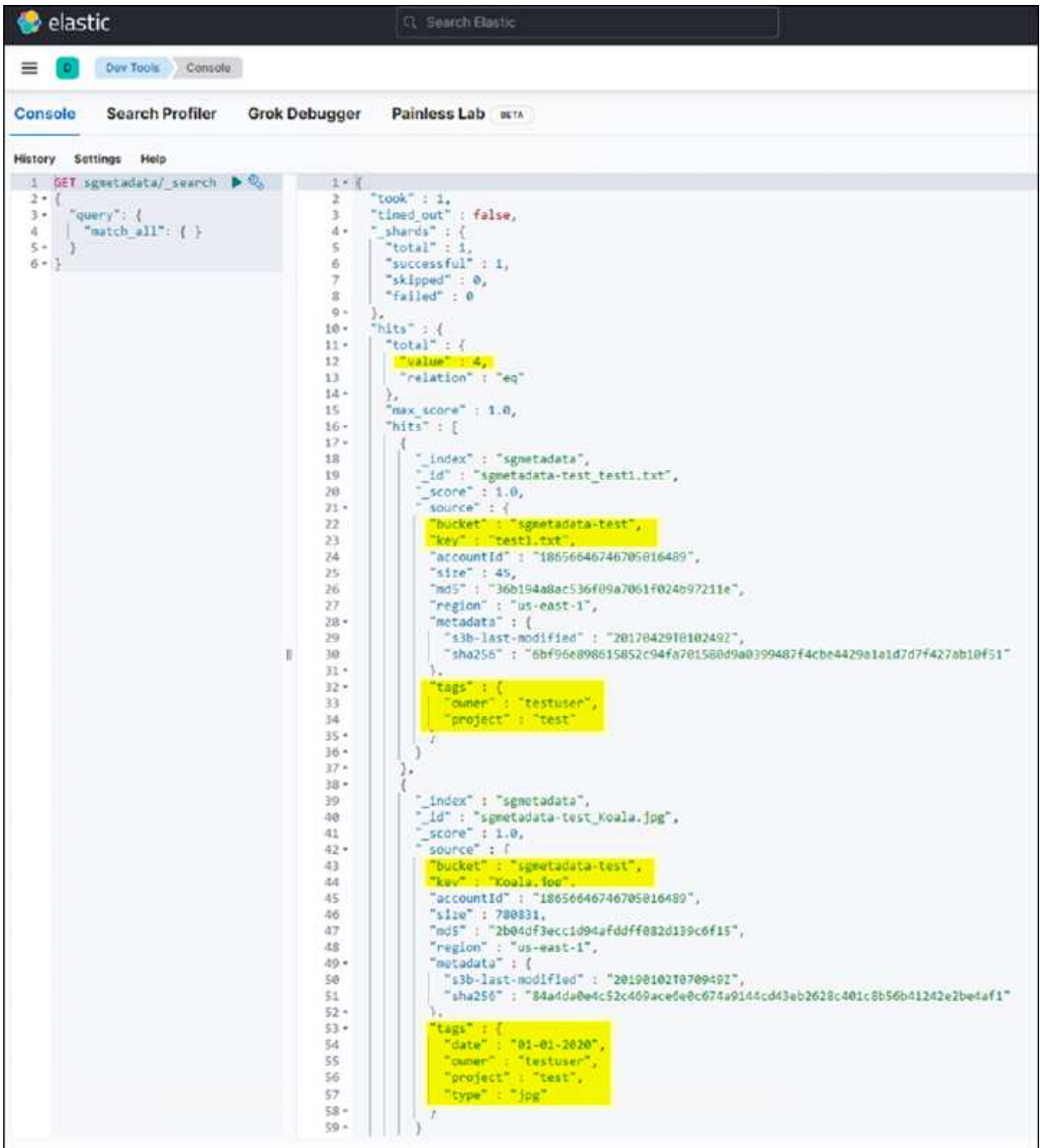
7. Utilice la interfaz de usuario de Kibana para verificar que los metadatos del objeto se cargaron en el índice de metadatos sg.
 - a. En el menú, seleccione Management > Dev Tools.
 - b. Pegue la consulta de ejemplo en el panel de la consola de la izquierda y haga clic en el símbolo de triángulo para ejecutarla.

El resultado de ejemplo de consulta 1 de la siguiente captura de pantalla de ejemplo muestra cuatro registros. Esto coincide con el número de objetos del segmento.

```

GET sgmetadata/_search
{
  "query": {
    "match_all": { }
  }
}

```



El resultado de ejemplo de la consulta 2 en la siguiente captura de pantalla muestra dos registros con el tipo de etiqueta jpg.

```

GET sgmetadata/_search
{
  "query": {
    "match": {
      "tags.type": {
        "query" : "jpg" }
      }
    }
  }
}

```

+

The screenshot shows the Elastic Search console interface. The top navigation bar includes 'elastic', 'Search Elastic', and various tool tabs like 'Dev Tools', 'Console', 'Search Profiler', 'Grok Debugger', and 'Painless Lab'. The main area is split into two panes: the left pane shows the search query, and the right pane shows the search results.

Search Query (Left Pane):

```

GET sgmetadata/_search
{
  "query": {
    "match": {
      "tags.type": {
        "query" : "jpg" }
      }
    }
  }
}

```

Search Results (Right Pane):

```

{
  "took": 1,
  "timed_out": false,
  "_shards": {
    "total": 1,
    "successful": 1,
    "skipped": 0,
    "failed": 0
  },
  "hits": {
    "total": 2,
    "value": 2,
    "relation": "eq"
  },
  "max_score": 0.18232156,
  "hits": [
    {
      "_index": "sgmetadata",
      "_id": "sgmetadata-test_koala.jpg",
      "_score": 0.18232156,
      "_source": {
        "bucket": "sgmetadata-test",
        "key": "Koala.jpg",
        "accountId": "18656646746705016489",
        "size": 788831,
        "md5": "2b84df3ecc1d94af0dff882d139c6f15",
        "region": "us-east-1",
        "metadata": {
          "s3b-last-modified": "20190102T070049Z",
          "sha256": "84a4da0e4c52c409ace6a0c674a9144cd43eb2628c001c0b56b41242e2be4af1"
        },
        "tags": [
          {
            "date": "01-01-2020",
            "owner": "testuser",
            "project": "test",
            "type": "jpg"
          }
        ]
      }
    },
    {
      "_index": "sgmetadata",
      "_id": "sgmetadata-test_lighthouse.jpg",
      "_score": 0.18232156,
      "_source": {
        "bucket": "sgmetadata-test",
        "key": "Lighthouse.jpg",
        "accountId": "18656646746705016489",
        "size": 561270,
        "md5": "8969288f4245120e7c3870287cce0ff3",
        "region": "us-east-1",
        "metadata": {
          "s3b-last-modified": "20090714T053221Z",
          "sha256": "ff86372ca435196075b8d8d29c98e9cbe905d400ba057c0544fa001fa4d0e73"
        },
        "tags": [
          {
            "date": "02-02-2022",
            "owner": "testuser",
            "project": "test",
            "type": "jpg"
          }
        ]
      }
    }
  ]
}

```

Dónde encontrar información adicional

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

- ["¿Qué son los servicios de plataforma"](#)
- ["Documentación de StorageGRID 11.6"](#)

Por Angela Cheng

Clon de nodo

Consideraciones y rendimiento del clon de nodos.

Consideraciones sobre el clon de nodo

El clon de nodo puede ser un método más rápido para reemplazar nodos de dispositivos existentes para realizar una actualización tecnológica, aumentar la capacidad o aumentar el rendimiento de su sistema StorageGRID. El clon de nodos también puede ser útil para convertir a cifrado de nodos con un KMS o cambiar un nodo de almacenamiento de DDP8 a DDP16.

- La capacidad utilizada del nodo de origen no es relevante al tiempo que se requiere para que se complete el proceso de clonado. El clon de nodo es una copia completa del nodo que incluye el espacio libre en el nodo.
- Los dispositivos de origen y destino deben tener la misma versión PGE
- El nodo de destino siempre debe tener una capacidad mayor que el origen
 - Asegúrese de que el nuevo dispositivo de destino tiene un tamaño de unidad mayor que la fuente
 - Si el dispositivo de destino tiene unidades del mismo tamaño y está configurado para DDP8, puede configurar el destino para DDP16. Si el origen ya está configurado para DDP16, el clon del nodo no será posible.
 - Al pasar de dispositivos SG5660 o SG5760 a dispositivos SG6060, tenga en cuenta que el SG6060 tiene 60 unidades de capacidad, en las que el SG6060 solo tiene 58.
- El proceso de clonado del nodo requiere que el nodo de origen esté desconectado al grid durante el proceso de clonado. Si se desconecta un nodo adicional durante esta ocasión, los servicios del cliente podrían verse afectados.
- Un nodo de almacenamiento solo puede estar desconectado durante 15 días. Si el cálculo del proceso de clonación está cerca de 15 días o supera los 15 días, utilice los procedimientos de expansión y retirada.
- Para un SG6060 con bandejas de expansión, es necesario añadir el tiempo para el tamaño de unidad de bandeja correcto a la hora del dispositivo de base para obtener el tiempo completo de los clones.
- La cantidad de volúmenes en un dispositivo de almacenamiento de destino debe ser mayor o igual que la cantidad de volúmenes en el nodo de origen. No se puede clonar un nodo de origen con volúmenes de almacenamiento de objetos 16 (rangedb) en un dispositivo de almacenamiento de destino con volúmenes de almacenamiento de objetos 12, incluso si el dispositivo de destino tiene más capacidad que el nodo de origen. La mayoría de los dispositivos de almacenamiento tienen volúmenes de almacenamiento de objetos de 16 TB, excepto el dispositivo de almacenamiento SGF6112 que solo tiene 12 volúmenes de almacenamiento de objetos. Por ejemplo, no puede clonar de un SG5760 a un SGF6112.

Estimaciones de rendimiento de clones de nodos

Las siguientes tablas contienen estimaciones calculadas para la duración del clon del nodo. Las condiciones varían de modo que las entradas en **NEGRITA** pueden correr el riesgo de superar el límite de 15 días para un nodo inactivo.

DDP8

SG5612 → cualquiera

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	1 día	2 días	2.5 días	3 días	4 días	4.5 días
25 GB	1 día	2 días	2.5 días	3 días	4 días	4.5 días

SG5712 → any

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	1 día	2 días	2.5 días	3 días	4 días	4.5 días
25 GB	1 día	2 días	2.5 días	3 días	4 días	4.5 días

SG5660 → SG5760

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	3 día	6 días	7 días	8.5 días	11.5 días	13 días
25 GB	3 día	6 días	7 días	8.5 días	11.5 días	13 días

SG5660 → SG6060

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	2.5 día	4.5 días	5.5 días	6.5 días	9 días	10 días
25 GB	2 días	4 días	5 días	6 días	8 días	9 días

SG5760 → SG5760

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	3 día	6 días	7 días	8.5 días	11.5 días	13 días
25 GB	3 día	6 días	7 días	8.5 días	11.5 días	13 días

SG5760 → SG6060

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	2.5 día	4.5 días	5.5 días	6.5 días	9 días	10 días
25 GB	1.5 día	3 días	3.5 días	4.5 días	6 días	6.5 días

SG6060 → SG6060

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	2.5 día	4.5 días	5.5 días	6.5 días	8.5 días	9.5 días
25 GB	1.5 día	3 días	3.5 días	4 días	5.5 días	6 días

DDP16

SG5760 → SG5760

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	3.5 día	6.5 días	8 días	9.5 días	12.5 días	14 días
25 GB	3.5 día	6.5 días	8 días	9.5 días	12.5 días	14 días

SG5760 → SG6060

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	2.5 día	5 días	6 días	7.5 días	10 días	11 días

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
25 GB	2 días	3.5 días	4 días	5 días	6.5 días	7 días

SG6060 → SG6060

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	3.5 día	5 días	6 días	7 días	9.5 días	10.5 días
25 GB	2 días	3 días	4 días	4.5 días	6 días	7 días

Bandeja de expansión (a partir de SG6060 para cada bandeja en el dispositivo de origen)

Velocidad de la interfaz de red	Tamaño de unidad de 4 TB	Tamaño de unidad de 8 TB	Tamaño de unidad de 10 TB	Tamaño de disco de 12 TB	Tamaño de la unidad de 16 TB	Tamaño de unidad de 18 TB
10 GB	3.5 día	5 días	6 días	7 días	9.5 días	10.5 días
25 GB	2 días	3 días	4 días	4.5 días	6 días	7 días

Por Aron Klein

Cómo utilizar el remap de puertos

Es posible que tenga que reasignar un puerto entrante o saliente por varias razones. Puede que esté pasando del servicio de equilibrador de carga CLB heredado al extremo de equilibrador de carga de servicio nginx actual y mantener el mismo puerto para reducir el impacto en los clientes, que desee utilizar el puerto 443 para el cliente S3 en una red de cliente de nodo de administración o para restricciones de firewall.

Migre clientes S3 de CLB A NGINX con la asignación de puertos

En versiones anteriores a StorageGRID 11.3, el servicio de equilibrador de carga incluido en los nodos de puerta de enlace es el equilibrador de carga de conexión (CLB). En StorageGRID 11.3, NetApp presenta el servicio NGINX como una completa solución integrada para el tráfico HTTP(s) de equilibrio de carga. Puesto que el servicio CLB sigue disponible en la versión actual de StorageGRID, no puede reutilizar el puerto 8082 en la nueva configuración del extremo del equilibrador de carga. Para evitar esto, el puerto de entrada 8082 se reasigna a 10443. Esto hace que todas las solicitudes HTTPS lleguen al puerto 8082 en la redirección de la puerta de enlace al puerto 10443, pasando por alto el servicio CLB y en su lugar conectándose al servicio NGINX. Aunque las siguientes instrucciones están disponibles para VMware, la funcionalidad PORT_REMAP existe para todos los métodos de instalación y puede utilizar un proceso similar para implementaciones sin configuración básica y dispositivos.

Puesta en marcha del nodo de puerta de enlace de máquinas virtuales de VMware

Los siguientes pasos son para una puesta en marcha de StorageGRID donde los nodos de puerta de enlace se ponen en marcha en VMware vSphere 7 como máquinas virtuales con el formato de virtualización abierta de StorageGRID (OVF). Este proceso implica la eliminación destructiva de la VM y la nueva puesta en marcha de la VM con el mismo nombre y configuración. Antes de encender la máquina virtual, cambie la propiedad VAPP para reasignar el puerto. A continuación, encienda la máquina virtual y siga el proceso de recuperación del nodo.

Requisitos previos

- Su sistema operativo es StorageGRID 11.3 o posterior
- Ha descargado y tiene acceso a los archivos de instalación de StorageGRID versión VMware instalados.
- Dispone de una cuenta de vCenter con permisos para encender y apagar equipos virtuales, cambiar la configuración de los equipos virtuales y vApps, eliminar equipos virtuales de vCenter y poner en marcha equipos virtuales por OVF.
- Ha creado un punto final de equilibrador de carga
 - El puerto está configurado para el puerto de redirección deseado
 - El certificado SSL de punto final es el mismo que el que se instala para el servicio CLB en el certificado de servidor certificados de configuración/certificados de servidor/extremos de servicio de API de almacenamiento de objetos o el cliente puede aceptar un cambio en el certificado.



If your existing certificate is self-signed, you cannot reuse it in the new endpoint. You must generate a new self-signed certificate when creating the endpoint and configure the clients to accept the new certificate.

Destruya el primer nodo de puerta de enlace

Para destruir el primer nodo de puerta de enlace, siga estos pasos:

1. Elija el nodo de puerta de enlace con el que empezar si la cuadrícula contiene más de uno.
2. Quite las IP de nodo de todas las entidades round-robin DNS o grupos de equilibradores de carga, si procede.
3. Espere a que caduque el tiempo de vida (TTL) y abra las sesiones.
4. Apague el nodo de máquina virtual.
5. Quite el nodo de máquina virtual del disco.

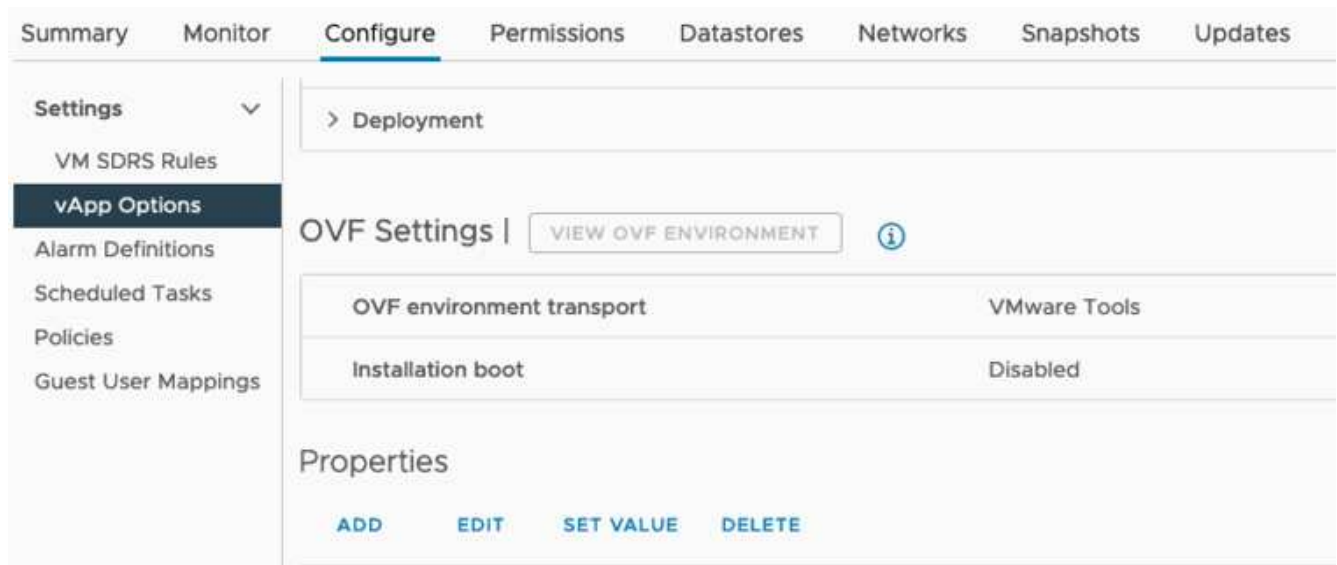
Implemente el nodo de puerta de enlace de reemplazo

Para implementar el nodo de puerta de enlace de repuesto, siga estos pasos:

1. Implemente el nuevo equipo virtual de OVF y seleccione los archivos .ovf, .mf y .vmdk en el paquete de instalación descargado del sitio de soporte:
 - vsphere-gateway.mf
 - vsphere-gateway.ovf

◦ NetApp-SG-11.4.0-20200721.1338.d3969b3.vmdk

- Una vez implementado el equipo virtual, selecciónelo en la lista de equipos virtuales, seleccione la pestaña Configurar vApp Options.



- Desplácese hasta la sección Propiedades y seleccione LA propiedad PORT_REMAP_INBOUND

Summary	Monitor	Configure	Permissions	Datastores	Networks	Snapshots	Updates
<input type="radio"/>	ADMIN_IP	Primary Admin IP	10.193.204.110		0.0.0.0	Grid Network (eth0)	ip
<input type="radio"/>	ADMIN_NETWORK_ESL	Admin network external subnet list				Admin Network (eth1)	string
<input type="radio"/>	ADMIN_NETWORK_IP	Admin network IP	10.193.174.112		0.0.0.0	Admin Network (eth1)	ip
<input type="radio"/>	NODE_TYPE	Node type			VM_API_Gateway	Grid Node Parameters	string["VM_Storage_Node", "VM_min_Node", "VM_API_Gateway", "_Archive_Node"]
<input type="radio"/>	CLIENT_NETWORK_CONFIG	Client network IP configuration	STATIC		DISABLED	Client Network (eth2)	string["DISABLED", "STATIC", "DHCP"]
<input checked="" type="radio"/>	PORT_REMAP_INBOUND	Inbound port remapping specification				Advanced	string
<input type="radio"/>	GRID_NETWORK	Grid network IP configuration	STATIC		STATIC	Grid Network	string["STATIC", "DHCP"]

- Desplácese hasta la parte superior de la lista Propiedades y haga clic en Editar



- Selecione la ficha Tipo, confirme que la casilla de verificación configurable por el usuario está seleccionada y, a continuación, haga clic en Guardar.

Edit property | Inbound port remapping specificati... X

General | **Type**

Static property

Type: String

User configurable:

Length: 0 - 65535

Default value: _____

Dynamic property

Macro: IP address

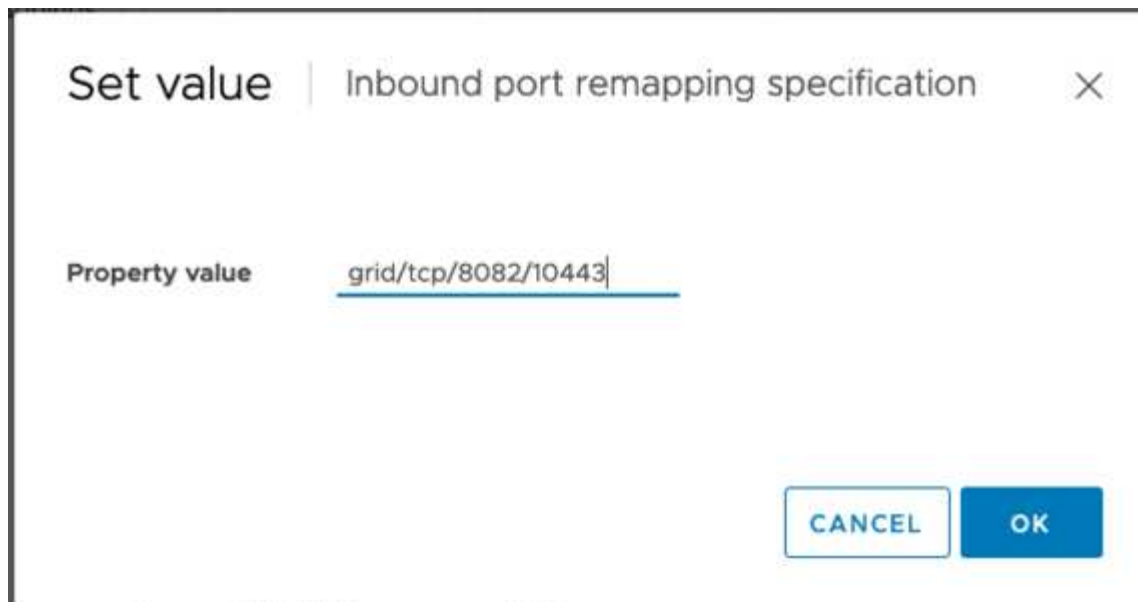
Network: MGMT_564

CANCEL SAVE

6. En la parte superior de la lista Propiedades, con la propiedad "PORT_REMAP_INBOUND" aún seleccionada, haga clic en establecer valor.



7. En el campo valor de propiedad, introduzca la red (grid, administrador o cliente), TCP, el puerto original (8082) y el puerto nuevo (10443) con "/" entre cada valor, como se describe a continuación.

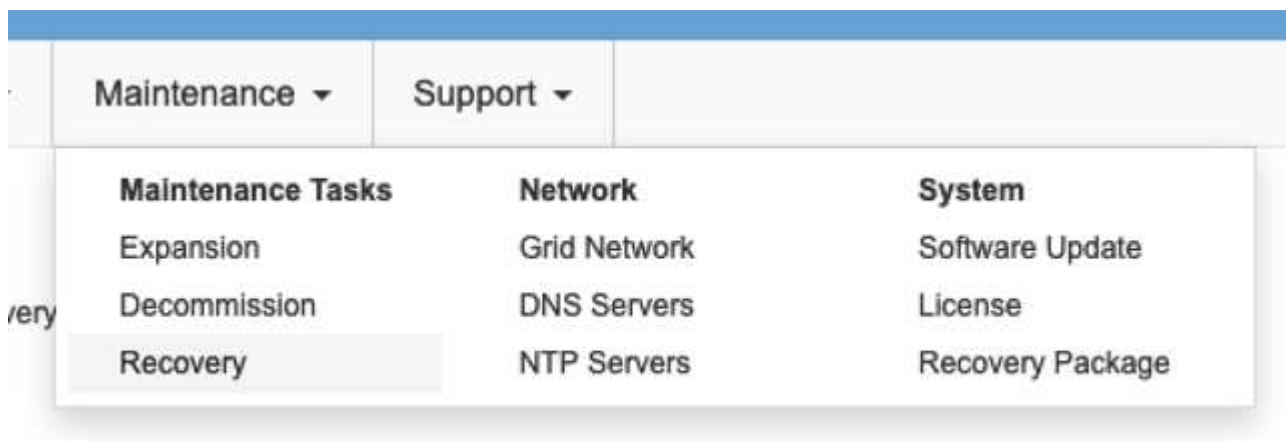


8. Si utiliza varias redes, utilice una coma (,) para separar las cadenas de red, por ejemplo, grid/tcp/8082/10443,admin/tcp/8082/10443,client/tcp/8082/10443

Recupere el nodo de puerta de enlace

Para recuperar el nodo Gateway, siga estos pasos:

1. Desplácese hasta la sección Mantenimiento/recuperación de la IU de gestión de grid.



2. Encienda el nodo de la máquina virtual y espere a que el nodo aparezca en la sección Maintenance/Recovery Pending Nodes de la interfaz de usuario de Grid Management.

Recovery

Select the failed grid node to recover, enter your provisioning passphrase, and then click Start Recovery to begin the recovery procedure.

Pending Nodes

Name	IPv4 Address	State	Recoverable
No results found.			



For information and directions for node recovery, see the <https://docs.netapp.com/sgws-114/topic/com.netapp.doc.sg-maint/GUID-7E22B1B9-4169-4800-8727-75F25FC0FFB1.html> [Recovery and Maintenance guide]

- Una vez recuperado el nodo, el IP se puede incluir en todas las entidades round-robin DNS o en los grupos de equilibradores de carga, si procede.

Ahora, cualquier sesión HTTPS en el puerto 8082 va al puerto 10443

Reasignar el puerto 443 para acceder S3 de cliente en un nodo de administración

La configuración predeterminada del sistema StorageGRID para un nodo de administración o un grupo de alta disponibilidad que contiene un nodo de administración es para que los puertos 443 y 80 se reserven para la interfaz de usuario del administrador de inquilinos y de gestión, y no se puede utilizar para extremos de equilibrio de carga. La solución a esto consiste en utilizar la operación de reasignación de puertos y redirigir el puerto de entrada 443 a un nuevo puerto que se configurará como punto final de equilibrio de carga. Una vez completado este tráfico de Client S3 podrá usar el puerto 443, la IU de administración de grid solo estará accesible a través del puerto 8443 y la IU de gestión de inquilinos solo estará accesible en el puerto 9443. La característica de reasignar puerto solo se puede configurar en el momento de instalación del nodo. Para implementar un remasterp de puertos de un nodo activo en la cuadrícula, se debe restablecer al estado preinstalado. Este es un procedimiento destructivo que incluye una recuperación de nodos una vez que se ha realizado el cambio de configuración.

Registros de backup y bases de datos

Los nodos de administración contienen registros de auditoría, métricas prometheus, así como información histórica sobre atributos, alarmas y alertas. Si tiene varios nodos de administrador, tendrá varias copias de estos datos. Si no tiene varios nodos de administrador en el grid, debe asegurarse de conservar estos datos para restaurar una vez que se haya recuperado el nodo al final de este proceso. Si tiene otro nodo de administrador en la cuadrícula, puede copiar los datos de ese nodo durante el proceso de recuperación. Si no tiene otro nodo de administrador en la cuadrícula, puede seguir estas instrucciones para copiar los datos antes de destruir el nodo.

Copiar registros de auditoría

- Inicie sesión en el nodo de administrador:
 - Introduzca el siguiente comando: `ssh admin@grid_node_IP`
 - Introduzca la contraseña que aparece en `Passwords.txt` archivo.

- c. Introduzca el siguiente comando para cambiar a la raíz: `su -`
- d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
- e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
- f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

2. Cree el directorio para copiar todos los archivos de registro de auditoría a una ubicación temporal en un nodo de cuadrícula independiente y permita utilizar `Storage_node_01`:
 - a. `ssh admin@storage_node_01_IP`
 - b. `mkdir -p /var/local/tmp/saved-audit-logs`
3. De nuevo en el nodo de administración, detenga el servicio AMS para evitar que cree un nuevo archivo de registro: `service ams stop`
4. Cambie el nombre del archivo `audit.log` para que no sobrescriba el archivo existente al copiarlo al nodo de administración recuperado.
 - a. Cambie el nombre de `audit.log` por un nombre de archivo numerado único como `aaaa-mm-dd.txt.1`. Por ejemplo, es posible cambiar el nombre del archivo de registro de auditoría a `2015-10-25.txt.1`

```
cd /var/local/audit/export
ls -l
mv audit.log 2015-10-25.txt.1
```

5. Reinicie el servicio AMS: `service ams start`
6. Copie todos los archivos del registro de auditoría: `scp * admin@storage_node_01_IP:/var/local/tmp/saved-audit-logs`

Copiar datos Prometheus



La copia de la base de datos Prometheus puede tardar una hora o más. Algunas funciones de Grid Manager no estarán disponibles mientras los servicios se detengan en el nodo de administración.

1. Cree el directorio para copiar los datos prometheus a una ubicación temporal en un nodo de cuadrícula independiente. De nuevo, utilizaremos `Storage_node_01`:
 - a. Inicie sesión en el nodo de almacenamiento:
 - i. Introduzca el siguiente comando: `ssh admin@storage_node_01_IP`
 - ii. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - iii. `mkdir -p /var/local/tmp/prometheus'`
2. Inicie sesión en el nodo de administrador:
 - a. Introduzca el siguiente comando: `ssh admin@admin_node_IP`

- b. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
- c. Introduzca el siguiente comando para cambiar a la raíz: `su -`
- d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
- e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
- f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

3. En el nodo de administración, detenga el servicio Prometheus: `service prometheus stop`
 - a. Copie la base de datos Prometheus del nodo de administración de origen en el nodo de ubicación del backup del nodo de almacenamiento: `/rsync -azh --stats "/var/local/mysql_ibdata/prometheus/data" "storage_node_01_IP:/var/local/tmp/prometheus/"`
4. Reinicie el servicio Prometheus en el nodo de administración de origen. `service prometheus start`

Información histórica de la copia de seguridad

La información histórica se almacena en una base de datos mysql. Para volcar una copia de la base de datos, necesitará el usuario y la contraseña de NetApp. Si posee otro nodo de administrador en la cuadrícula, este paso no es necesario y la base de datos se puede clonar a partir de un nodo de administrador restante durante el proceso de recuperación.

1. Inicie sesión en el nodo de administrador:
 - a. Introduzca el siguiente comando: `ssh admin@admin_node_IP`
 - b. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - c. Introduzca el siguiente comando para cambiar a la raíz: `su -`
 - d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
 - f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

2. Detenga los servicios de StorageGRID en el nodo de administración e inicie `ntp` y `mysql`
 - a. Detenga todos los servicios: `service servermanager stop`
 - b. reinicie el servicio `ntp`: `service ntp start..restart mysql servicio: service mysql start`
3. Volcar mi base de datos a `/var/local/tmp`
 - a. introduzca el siguiente comando: `mysqldump -u username -p password mi > /var/local/tmp/mysql-mi.sql`
4. Copie el archivo de volcado mysql en un nodo alternativo, usaremos `Storage_node_01`:
`scp /var/local/tmp/mysql-mi.sql _storage_node_01_IP:/var/local/tmp/mysql-mi.sql`

- a. Cuando ya no necesite un acceso sin contraseñas a otros servidores, quite la clave privada del agente SSH. Introduzca: `ssh-add -D`

Vuelva a crear el nodo Admin

Ahora que dispone de una copia de backup de todos los datos y registros deseados en otro nodo de administrador de la cuadrícula o almacenados en una ubicación temporal, es hora de restablecer el dispositivo para poder configurar el remapa de puertos.

1. El restablecimiento de un dispositivo vuelve al estado preinstalado y solo conserva el nombre de host, las IP y las configuraciones de red. Se perderán todos los datos, por lo que nos aseguramos de contar con una copia de seguridad de cualquier información importante.
 - a. introduzca el siguiente comando: `sgareinstall`

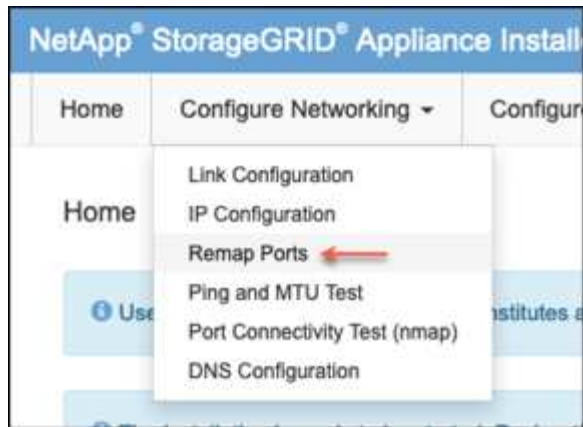
```
root@sg100-01:~ # sgareinstall
WARNING: All StorageGRID Webscale services on this node will be shut
down.
WARNING: Data stored on this node may be lost.
WARNING: You will have to reinstall StorageGRID Webscale to this
node.

After running this command and waiting a few minutes for the node to
reboot,
browse to one of the following URLs to reinstall StorageGRID Webscale
on
this node:

    https://10.193.174.192:8443
    https://10.193.204.192:8443
    https://169.254.0.1:8443

Are you sure you want to continue (y/n)? y
Renaming SG installation flag file.
Initiating a reboot to trigger the StorageGRID Webscale appliance
installation wizard.
```

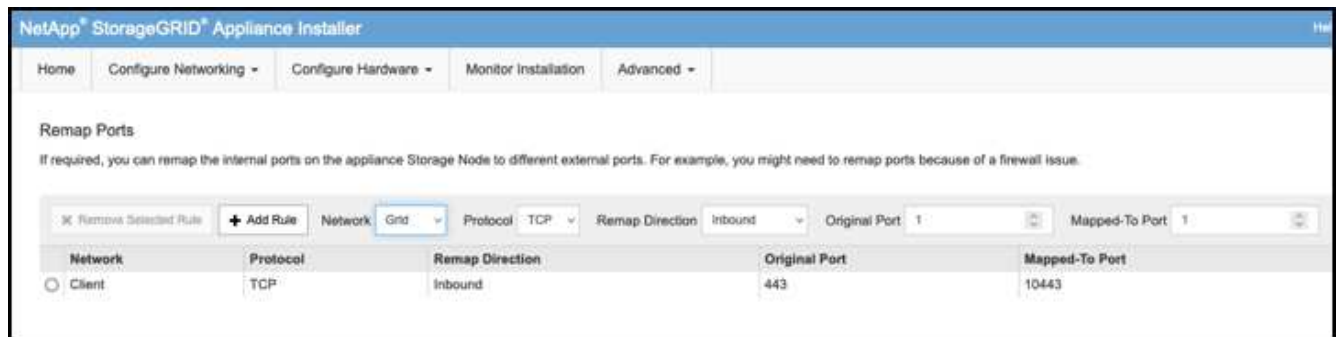
2. Cuando haya transcurrido un periodo de tiempo, el dispositivo se reiniciará y podrá acceder a la IU del nodo PGE.
3. Vaya a Configurar redes



4. Seleccione la red, el protocolo, la dirección y los puertos deseados y, a continuación, haga clic en el botón Agregar regla.



La reasignación del puerto de entrada 443 en LA red DE CUADRÍCULA interrumpirá la instalación y los procedimientos de expansión. No se recomienda reasignar el puerto 443 en la red DE RED.



5. Una de las reasignaciones de puerto deseadas se ha agregado, puede volver a la ficha de inicio y hacer clic en el botón Iniciar instalación.

Ahora puede seguir los procedimientos de recuperación del nodo de administrador en el "[documentación de productos](#)"

Restaurar bases de datos y registros

Ahora que el nodo de administrador se ha recuperado, podrá restaurar las métricas, los registros y la información histórica. Si tiene otro nodo de administrador en la cuadrícula, siga la "[documentación de productos](#)" utilizando los scripts *prometheus-clone-db.sh* y *mi-clone-db.sh*. Si este es el único nodo de administrador y decide realizar una copia de seguridad de estos datos, puede seguir los pasos que se indican a continuación para restaurar la información.

Vuelva a copiar los registros de auditoría

1. Inicie sesión en el nodo de administrador:
 - a. Introduzca el siguiente comando: `ssh admin@grid_node_IP`
 - b. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - c. Introduzca el siguiente comando para cambiar a la raíz: `su -`

- d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
- e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
- f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

2. Copie los archivos de registro de auditoría conservados en el nodo admin recuperado: `scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY* .`
3. Por motivos de seguridad, elimine los registros de auditoría del nodo de grid con errores después de verificar que se han copiado correctamente al nodo de administrador recuperado.
4. Actualice la configuración de usuario y grupo de los archivos de registro de auditoría en el nodo de administración recuperado: `chown ams-user:bycast *`

También debe restaurar cualquier acceso de cliente preexistente al recurso compartido de auditoría. Para obtener más información, consulte las instrucciones para administrar StorageGRID.

Restaurar métricas de Prometheus



La copia de la base de datos Prometheus puede tardar una hora o más. Algunas funciones de Grid Manager no estarán disponibles mientras los servicios se detengan en el nodo de administración.

1. Inicie sesión en el nodo de administrador:
 - a. Introduzca el siguiente comando: `ssh admin@grid_node_IP`
 - b. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - c. Introduzca el siguiente comando para cambiar a la raíz: `su -`
 - d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
 - f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

2. En el nodo de administración, detenga el servicio Prometheus: `service prometheus stop`
 - a. Copie la base de datos Prometheus de la ubicación temporal del backup al nodo de administración:
`/rsync -azh --stats "backup_node:/var/local/tmp/prometheus/"
"/var/local/mysql_ibdata/prometheus/"`
 - b. compruebe que los datos están en la ruta correcta y que han finalizado `ls /var/local/mysql_ibdata/prometheus/data/`
3. Reinicie el servicio Prometheus en el nodo de administración de origen. `service prometheus start`

Restaurar información histórica

1. Inicie sesión en el nodo de administrador:
 - a. Introduzca el siguiente comando: `ssh admin@grid_node_IP`
 - b. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - c. Introduzca el siguiente comando para cambiar a la raíz: `su -`
 - d. Introduzca la contraseña que aparece en `Passwords.txt` archivo.
 - e. Añada la clave privada SSH al agente SSH. Introduzca: `ssh-add`
 - f. Introduzca la contraseña de acceso SSH que aparece en la `Passwords.txt` archivo.

When you are logged in as root, the prompt changes from ``$`` to ``#``.

2. Copie el archivo de volcado mysql del nodo alternativo: `scp grid_node_IP_:/var/local/tmp/mysql-mi.sql /var/local/tmp/mysql-mi.sql`
3. Detenga los servicios de StorageGRID en el nodo de administración e inicie ntp y mysql
 - a. Detenga todos los servicios: `service servermanager stop`
 - b. reinicie el servicio ntp: `service ntp start..restart mysql servicio: service mysql start`
4. Borre la base de datos mi y cree una nueva base de datos vacía: `mysql -u username -p password -A mi -e "drop database mi; create database mi;"`
5. restaure la base de datos mysql desde el volcado de la base de datos: `mysql -u username -p password -A mi < /var/local/tmp/mysql-mi.sql`
6. Reinicie todos los demás servicios `service servermanager start`

Por Aron Klein

Reubicación del sitio de grid y procedimiento de cambio de red en todo el sitio

Esta guía describe la preparación y el procedimiento para la reubicación del sitio StorageGRID en una cuadrícula de varios sitios. Usted debe tener una comprensión completa de este procedimiento y planificar con anticipación para garantizar un proceso sin problemas y minimizar la interrupción a los clientes.

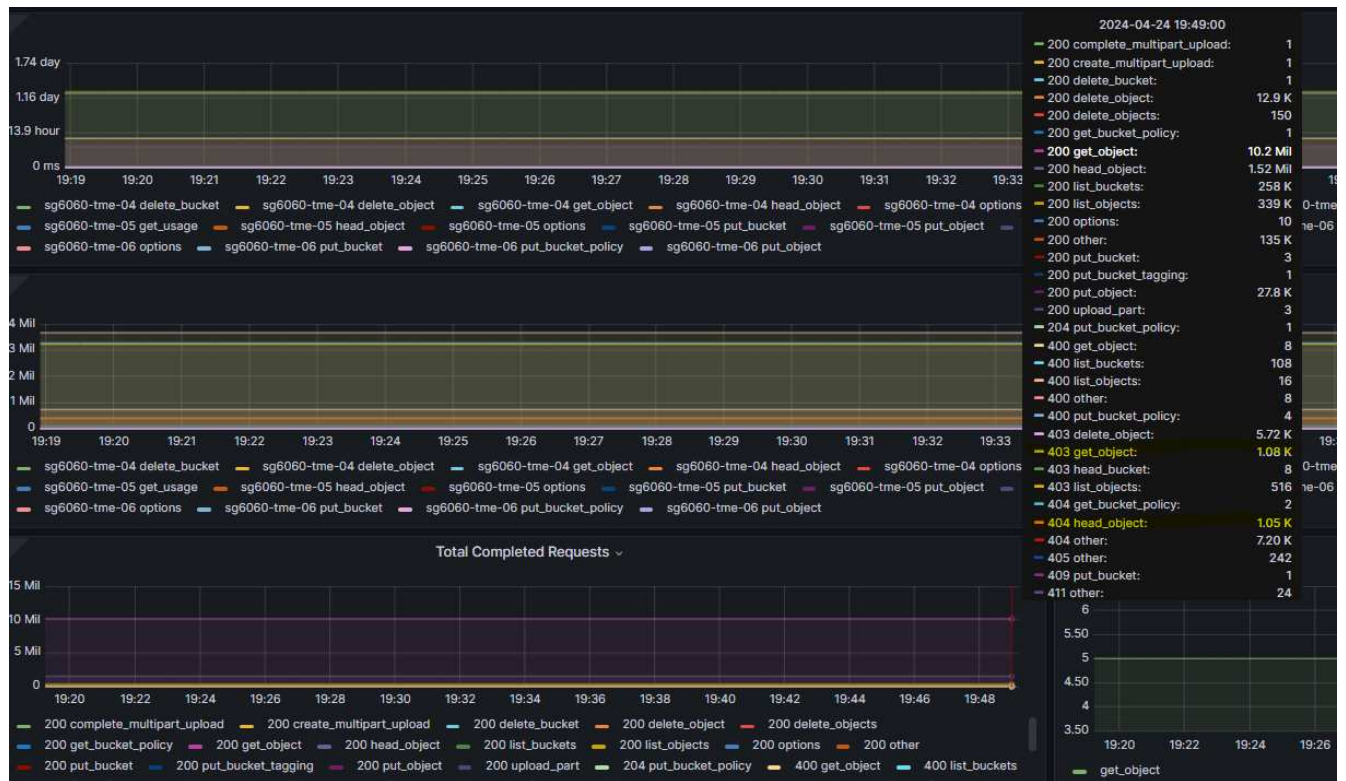
Si necesita cambiar la red de Grid de toda la Grid, consulte ["Cambie las direcciones IP para todos los nodos de la cuadrícula"](#).

Consideraciones antes de la reubicación del sitio

- El movimiento del sitio debe completarse y todos los nodos en línea en 15 días para evitar la reconstrucción de la base de datos Cassandra.
["Recupere el nodo de almacenamiento en más de 15 días"](#)
- Si alguna regla de ILM en una política activa utiliza un comportamiento de ingesta estricto, considere cambiarla al equilibrio o a la confirmación doble si el cliente desea seguir PONIENDO objetos en el Grid

durante la reubicación del sitio.

- Para los dispositivos de almacenamiento con unidades 60 o más, no mueva nunca la bandeja con unidades de disco instaladas. Etiquete cada unidad de disco y quítelas del compartimento de almacenamiento antes de empacar/mover.
- La VLAN de la red de grid del dispositivo StorageGRID se puede realizar de forma remota a través de la red de administración o la red de cliente. O bien, planifique encontrarse en las instalaciones para realizar el cambio antes o después de la reubicación.
- Compruebe si la aplicación del cliente está utilizando HEAD u OBJECT OPTIMITY antes de PUT. En caso afirmativo, cambie la consistencia del bloque a un sitio seguro para evitar el error de HTTP 500. Si no está seguro, consulte la descripción general de S3 Gráficos Grafana * Administrador de cuadrícula > Soporte > Métricas * y pase el ratón sobre el gráfico 'Total de solicitudes completadas'. Si hay un recuento muy alto de 404 objetos Get Object o 404 objetos head, es probable que una o más aplicaciones estén usando objetos head o get nonexistence. El recuento es acumulativo, pasa el ratón sobre una línea de tiempo diferente para ver la diferencia.



Procedimiento para cambiar la dirección IP de la cuadrícula antes de la reubicación del sitio

Pasos

1. Si se va a utilizar una nueva subred de red de Grid en la nueva ubicación, ["Agregue la subred a la lista de subred de red de cuadrícula"](#)
2. Inicie sesión en el nodo de administración principal, use CHANGE-ip para hacer el cambio de Grid IP, debe * almacenar en zona intermedia * el cambio antes de cerrar el nodo para su reubicación.
 - a. Seleccione 2 y, a continuación, 1 para el cambio de IP de cuadrícula

Editing: Node IP/subnet and gateway

Use up arrow to recall a previously typed value, which you can then edit
Use d or 0.0.0.0/0 as the IP/mask to delete the network from the node
Use q to complete the editing session early and return to the previous menu
Press <enter> to use the value shown in square brackets

```
=====
Site: LONDON
=====
LONDON-ADM1 Grid IP/mask [ 10.45.74.14/26 ]: 10.45.74.24/26
LONDON-S1   Grid IP/mask [ 10.45.74.16/26 ]: 10.45.74.26/26
LONDON-S2   Grid IP/mask [ 10.45.74.17/26 ]: 10.45.74.27/26
LONDON-S3   Grid IP/mask [ 10.45.74.18/26 ]: 10.45.74.28/26
=====
LONDON-ADM1 Grid Gateway [ 10.45.74.1 ]:
LONDON-S1   Grid Gateway [ 10.45.74.1 ]:
LONDON-S2   Grid Gateway [ 10.45.74.1 ]:
LONDON-S3   Grid Gateway [ 10.45.74.1 ]:
=====
Site: OXFORD
=====
OXFORD-ADM1 Grid IP/mask [ 10.45.75.14/26 ]:
OXFORD-S1   Grid IP/mask [ 10.45.75.16/26 ]:
OXFORD-S2   Grid IP/mask [ 10.45.75.17/26 ]:
OXFORD-S3   Grid IP/mask [ 10.45.75.18/26 ]:
=====
OXFORD-ADM1 Grid Gateway [ 10.45.75.1 ]:
OXFORD-S1   Grid Gateway [ 10.45.75.1 ]:
OXFORD-S2   Grid Gateway [ 10.45.75.1 ]:
OXFORD-S3   Grid Gateway [ 10.45.75.1 ]:
=====
Finished editing. Press Enter to return to menu. █
```

b. seleccione 5 para mostrar los cambios

```
=====
Site: LONDON
=====
LONDON-ADM1 Grid IP [ 10.45.74.14/26 ]: 10.45.74.24/26
LONDON-S1   Grid IP [ 10.45.74.16/26 ]: 10.45.74.26/26
LONDON-S2   Grid IP [ 10.45.74.17/26 ]: 10.45.74.27/26
LONDON-S3   Grid IP [ 10.45.74.18/26 ]: 10.45.74.28/26
Press Enter to continue █
```

c. seleccione 10 para validar y aplicar el cambio.

```

Welcome to the StorageGRID IP Change Tool.

Selected nodes: all

1:  SELECT NODES to edit
2:  EDIT IP/mask and gateway
3:  EDIT admin network subnet lists
4:  EDIT grid network subnet list
5:  SHOW changes
6:  SHOW full configuration, with changes highlighted
7:  VALIDATE changes
8:  SAVE changes, so you can resume later
9:  CLEAR all changes, to start fresh
10: APPLY changes to the grid
0:  Exit

Selection: 10

```

- d. Debe seleccionar **stage** en este paso.

```

Validating new networking configuration... PASSED.
Checking for Grid Network IP address swaps... PASSED.

Applying these changes will update the following nodes:

LONDON-ADM1
LONDON-S1
LONDON-S2
LONDON-S3

The following nodes will also require restarting:

LONDON-ADM1
LONDON-S1
LONDON-S2
LONDON-S3

Select one of the following options:

apply:  apply all changes and automatically restart nodes (if necessary)
stage:  stage the changes; no changes will take effect until the nodes are restarted
cancel: do not make any network changes at this time

[apply/stage/cancel]> stage

```

- e. Si el nodo de administración principal está incluido en el cambio anterior, introduzca 'a' para reiniciar manualmente el nodo de administración principal

```

10.45.74.14 - PuTTY
Validating new networking configuration... PASSED.
Checking for Grid Network IP address swaps... PASSED.

Applying these changes will update the following nodes:

LONDON-ADM1
LONDON-S1
LONDON-S2
LONDON-S3

The following nodes will also require restarting:

LONDON-ADM1
LONDON-S1
LONDON-S2
LONDON-S3

Select one of the following options:

  apply:  apply all changes and automatically restart nodes (if necessary)
  stage:  stage the changes; no changes will take effect until the nodes are restarted
  cancel: do not make any network changes at this time

[apply/stage/cancel]> stage

Generating new grid networking description file... PASSED.
Running provisioning... PASSED.
Updating network configuration on LONDON-S1... PASSED.
Updating network configuration on LONDON-S2... PASSED.
Updating network configuration on LONDON-S3... PASSED.
Updating network configuration on LONDON-ADM1... PASSED.
Finished staging network changes. You must manually restart these nodes for the changes to take effect:

LONDON-ADM1 (has IP 10.45.74.14 until restart)
LONDON-S1 (has IP 10.45.74.16 until restart)
LONDON-S2 (has IP 10.45.74.17 until restart)
LONDON-S3 (has IP 10.45.74.18 until restart)

Importing bundles... PASSED.
*****
*                               *
*             IMPORTANT         *
*                               *
* A new recovery package has been generated as a result of the *
* configuration change. Select Maintenance > Recovery Package *
* in the Grid Manager to download it.                          *
*****

Network Update Complete. Primary admin restart required. Select 'continue' to restart this node immediately, 'abort' to restart manually.
Enter a to abort, c to continue [a/c]>

```

f. Pulse ENTER para volver al menú anterior y salir de la interfaz CHANGE-ip.

```

Network Update Complete. Primary admin restart required. Select 'continue' to restart this node immediately, 'abort' to restart manually.
Enter a to abort, c to continue [a/c]> a
Restart aborted. You must manually restart this node as soon as possible
Press Enter to return to the previous menu.

```

3. En Grid Manager, descargue el nuevo paquete de recuperación. **Grid manager > Mantenimiento > Paquete de recuperación**
4. Si se requiere un cambio de VLAN en el dispositivo StorageGRID, consulte la sección [Cambio de VLAN del dispositivo](#).
5. Apague todos los nodos o dispositivos en el sitio, etiquete/quite las unidades de disco si es necesario, desmonte el rack, empaquete y mueva.
6. Si tiene pensado cambiar la ip de la red de administración y/o la VLAN y la dirección ip de cliente, puede realizar el cambio después de la reubicación.

Cambio de VLAN del dispositivo

En el siguiente procedimiento se asume que tiene acceso remoto a la red de clientes o administradores del dispositivo StorageGRID para realizar el cambio de forma remota.

Pasos

1. Antes de apagar el aparato, ["coloque el aparato en modo de mantenimiento"](#).

2. Utilice un explorador para acceder a la GUI del instalador del dispositivo StorageGRID mediante <https://<admin-or-client-network-ip>:8443>. No se puede utilizar Grid IP debido a que la nueva IP de Grid ya está en su lugar una vez que el dispositivo se arranca en modo de mantenimiento.
3. Cambie la VLAN para la red de grid. Si accede al dispositivo a través de red de cliente, no puede cambiar la VLAN de cliente en este momento; puede cambiarlo después del movimiento.
4. ssh en el dispositivo y apague el nodo mediante 'hutdown -h now'
5. Una vez que los dispositivos estén listos en el sitio nuevo, acceda a la interfaz gráfica de usuario del instalador del dispositivo StorageGRID mediante <https://<grid-network-ip>:8443>. Confirme que el almacenamiento se encuentre en estado óptimo y que la conectividad de red a otros nodos de Grid mediante las herramientas ping/nmap en la GUI.
6. Si planea cambiar la IP de red del cliente, puede cambiar la VLAN del cliente en este momento. La red cliente no está lista hasta que actualice la ip de la red cliente mediante la herramienta CHANGE-ip en el paso posterior.
7. Salga del modo de mantenimiento. En el instalador del dispositivo StorageGRID, seleccione **Avanzado > Reiniciar controlador** y, a continuación, seleccione **Reiniciar en StorageGRID**.
8. Una vez que todos los nodos estén activos y Grid no muestre ningún problema de conectividad, utilice change-ip para actualizar la red de administración del dispositivo y la red de cliente si es necesario.

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.

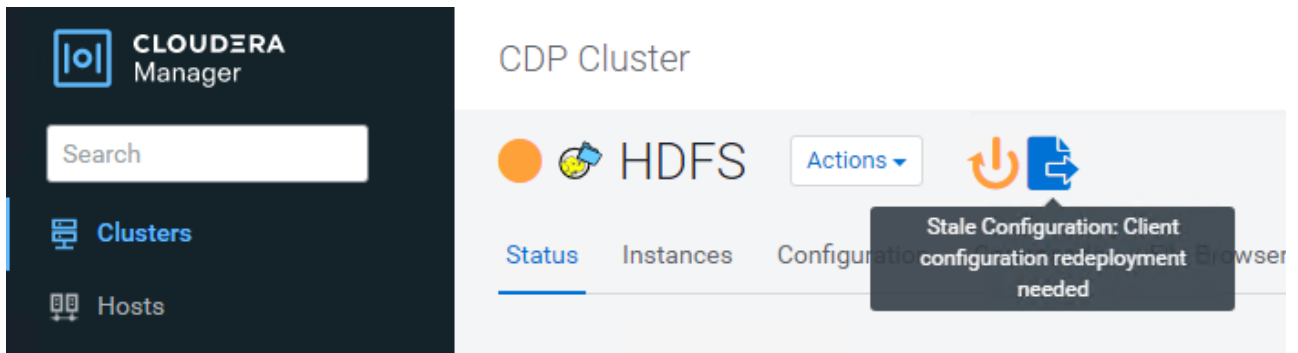
Nombre	Valor
fs.s3a.access.key	<clave de acceso de inquilino s3 de StorageGRID>
fs.s3a.secret.key	<clave secreta de inquilino s3 de StorageGRID>
fs.s3a.connection.ssl.enabled	[verdadero o falso] (el valor predeterminado es https si falta esta entrada)
fs.s3a.endpoint	<StorageGRID S3 endpoint:Port>

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

Name	fs.s3a.endpoint	 
Value	sgdemo.netapp.com:10443	
Description	StorageGRID s3 load balancer endpoint	
	<input checked="" type="checkbox"/> Final	
Name	fs.s3a.access.key	 
Value	OMC...BAN	
Description	SG CDP S3 access key	
	<input checked="" type="checkbox"/> Final	
Name	fs.s3a.secret.key	 
Value	mapz...Qfc	
Description	SG CDP S3 secret key	
	<input checked="" type="checkbox"/> Final	
Name	fs.s3a.impl	 
Value	org.apache.hadoop.fs.s3a.S3AFileSystem	
Description		
	<input checked="" type="checkbox"/> Final	
Name	fs.s3a.path.style.access	 
Value	true	
Description		
	<input checked="" type="checkbox"/> Final	

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 filesystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClietIOException:
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 filesystem
s3a://hdfs-test/: org.apache.hadoop.fs.s3a.AWSClietIOException:
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.

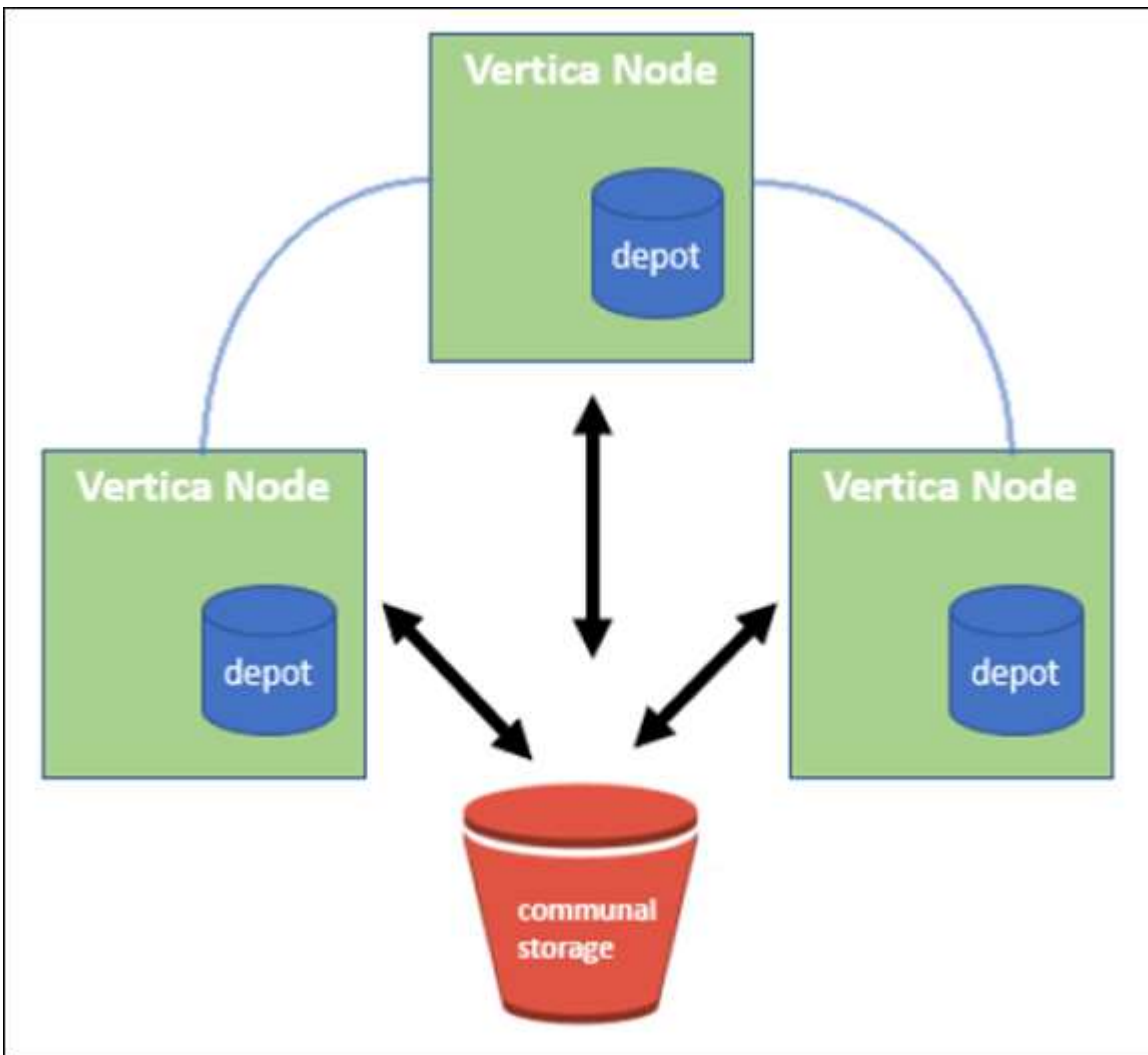
Los modos Eon y Enterprise difieren principalmente en el lugar en el que almacenan datos:

- 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 = MNVU40YFAY2xyz123:03vu04M4KmdfwffT8nqnBmnMVTr78Gu9wANabcxyz
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:03vu04M4KmdfwffT8nqnBmnMVTr78Gu9wAN+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
'xxxxxxxxx'
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/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602.tar
204217344	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610.tar
16109056	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0.tar
12853248	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800.tar

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
8937984	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a.tar
56260608	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2.tar
53947904	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba.tar
44932608	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de.tar
256306688	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar
8062464	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar
20024832	s3://vertica/metadatos/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.tar
10444	s3://vertica/metadatos/VMart/cluster_config.json
823266	s3://vertica/metadatos/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/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/completed
2958	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/chkpt_1.cat.gz
231	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/completed
822521	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/chkpt_1.cat.gz
231	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/completed
746513	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g14.cat
2596	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_3_g3.cat.gz
821065	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_4_g4.cat.gz
6440	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_5_g5.cat
8518	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_8_g8.cat

Tamaño del objeto (byte)	Ruta completa de clave de bloque/objeto
0	s3://vertica/metadatos/VMart/nodos/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadatos/VMart/nodos/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadatos/VMart/nodos/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadatos/VMart/nodos/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadatos/VMart/nodos/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadatos/VMart/nodos/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadatos/VMart/nodos/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadatos/VMart/nodos/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadatos/VMart/nodos/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadatos/VMart/nodos/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadatos/VMart/nodos/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 `vsq1` 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 ~]$ vsq1
Password:
Welcome to vsq1, the Vertica Analytic Database interactive terminal.
Type:      \h or \? for help with vsq1 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.

```
vsq1 -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. `vsq1 < vmart_define_schema.sql`
- d. `vsq1 < 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: `vsq1 < 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**

Actions	Field	Value
...	_id	yuh0iIEBVP6KX4EwqcyU
...	_index	sglog-2022.06.21
...	_score	-
...	@timestamp	Jun 21, 2022 @ 18:07:45.444
...	event.original	<28>Jun 21 22:07:45 SITE2-S3 ADE: syslog messages being dropped
...	host.hostname	SITE2-S3
...	msg-details	syslog messages being dropped
...	process.name	ADE
...	syslog_pri	28
...	timestamp	Jun 21 22:07:45

• 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.

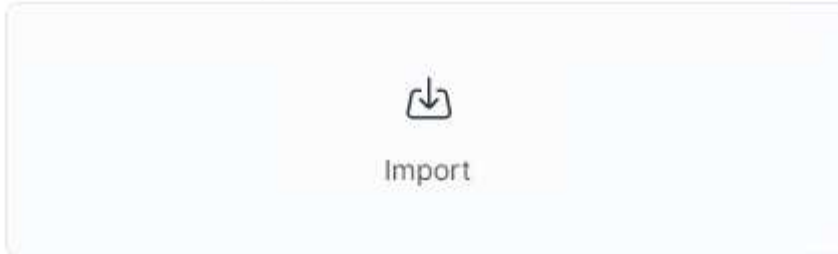
```
GET /_cat/indices/*?v=true&s=index
```

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"

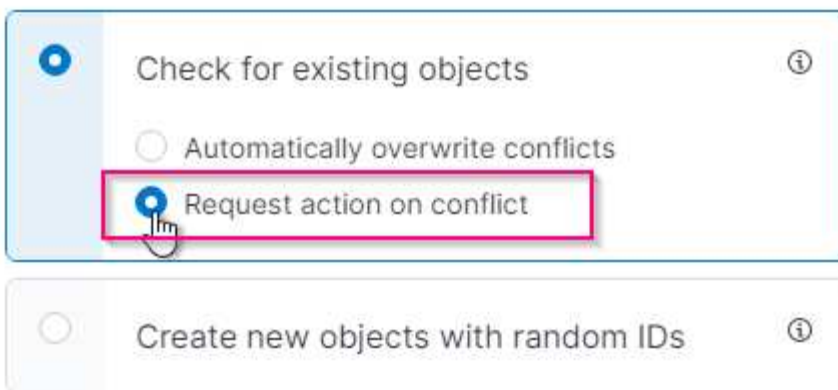
Import saved objects



Select a file to import



Import options



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.

Import saved objects ×

Data Views Conflicts

The following saved objects use data views that do not exist. Please select the data views you'd like re-associated with them. You can [create a new data view](#) if necessary.

ID	Count	Sample of aff...	New data view
594f91a0-d192-11ec-b30f-09f67aedd1d9	2		sglog ▾
60cf3620-e5fa-11ec-af71-8f6e980d6eb0	1		sglog ▾

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"](#)

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 instalarse 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.

Add a client certificate

1

Enter details

2

Enter details

Certificate details

Certificate name [?](#)

Permissions

Allow prometheus [?](#)

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".

Add a client certificate



Enter details



Enter details

Certificate type



Upload certificate



Generate certificate

Domain name

prometheus.grid.local

[Add another domain](#)

IP

192.168.0.10

[Add another IP address](#)

Subject

/CN=Prometheus

Days valid

730

[Generate](#)

[Previous](#)

[Create](#)



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.

Generate

Certificate details

Download certificate Copy certificate PEM

Subject DN: /CN=Prometheus
 Serial Number: 72:D9:6E:D7:04:CC:4F:29:66:0A:CA:53:24:79:1B:09:49:3A:BC:56
 Issuer DN: /CN=Prometheus
 Issued On: 2022-08-22T17:54:33.000Z
 Expires On: 2024-08-21T17:54:33.000Z
 SHA-1 Fingerprint: 10:47:6E:FD:67:D8:53:E7:6E:E5:D8:8A:DF:BD:45:94:04:53:47:1E
 SHA-256 Fingerprint: 74:23:C2:02:3A:D9:08:C0:EE:C1:F8:59:8A:7C:AE:18:AB:80:7D:21:31:F3:EB:AF:BF:4F:9E:C7:90:C9:FA:E7
 Alternative Names: DNS:prometheus.grid.local
 IP Address:192.168.0.10

Certificate private key

⚠ You will not be able to view the certificate private key after you close this dialog. To save the keys for future reference, copy and paste the values to another location.

Download private key Copy private key

```
-----BEGIN RSA PRIVATE KEY-----
MIIEpAIBAAKCAQEA3bIcyIEpMWPk5ritVpMkmIDKLIjaTM3ertq23VcAALwxziaU
...
```



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
```

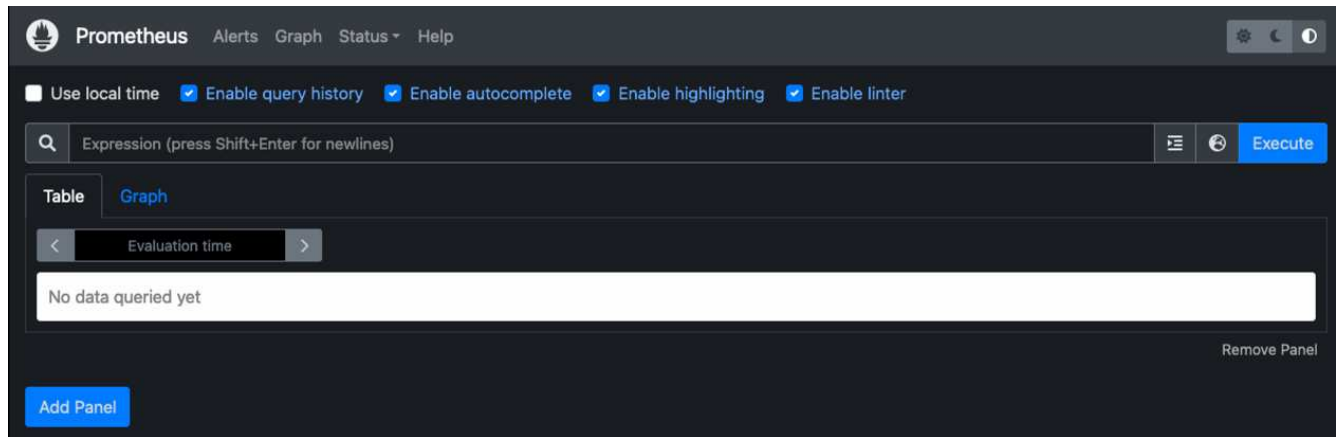
```

• 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>

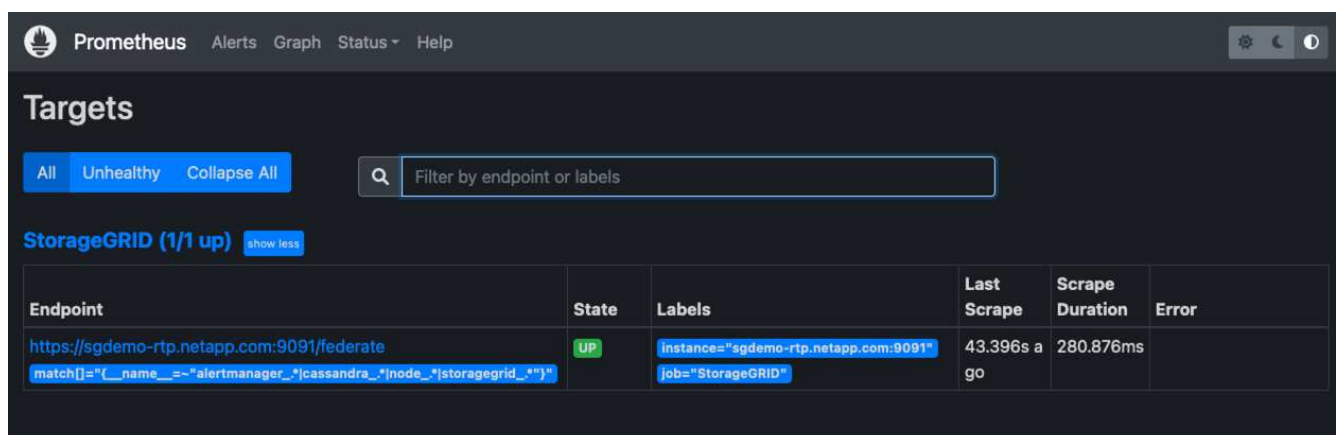
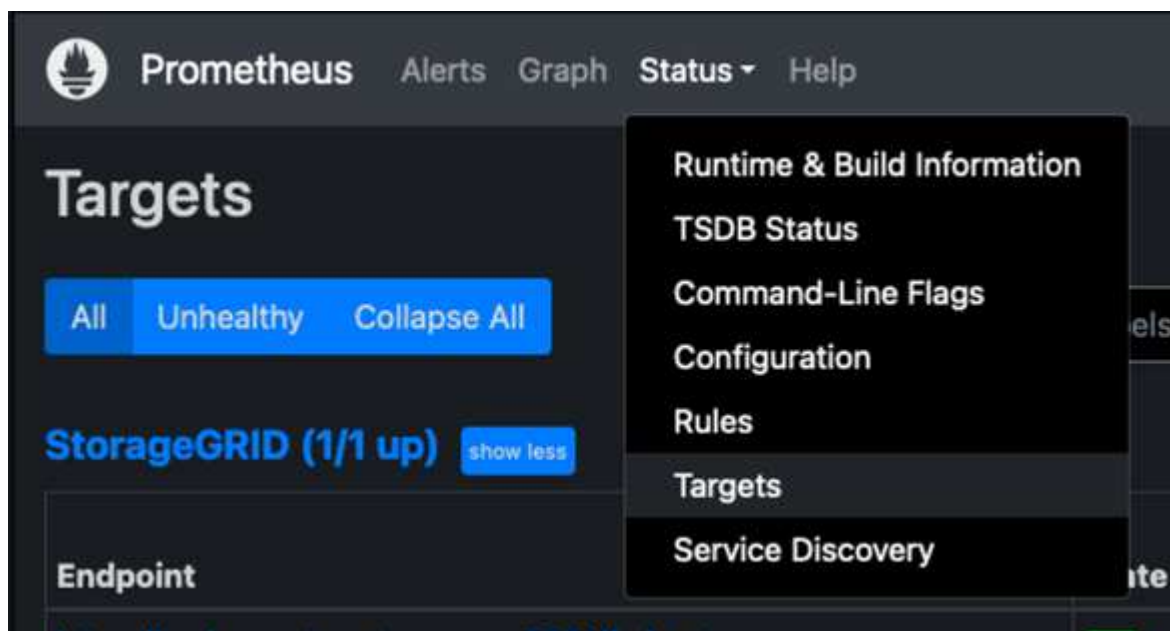
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.510Z caller=head.go:544 level=info component=tsdb
msg="Replaying WAL, this may take a while"
Aug 22 15:14:24 aj-deb-prom01 prometheus[6498]: ts=2022-08-
22T19:14:24.816Z caller=head.go:615 level=info component=tsdb msg="WAL
segment 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..."

```

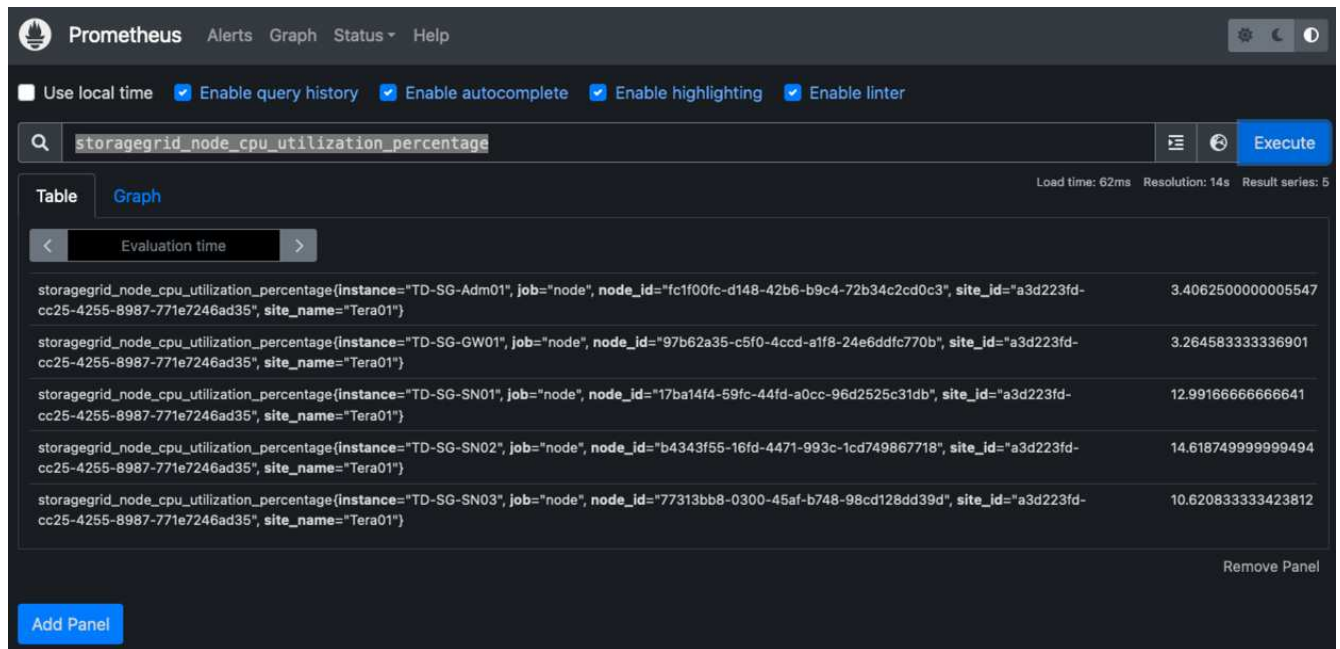
6. Ahora debe poder acceder a la interfaz de usuario de su servidor prometheus <http://Prometheus-server:9090> Y ver la interfaz de usuario



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.

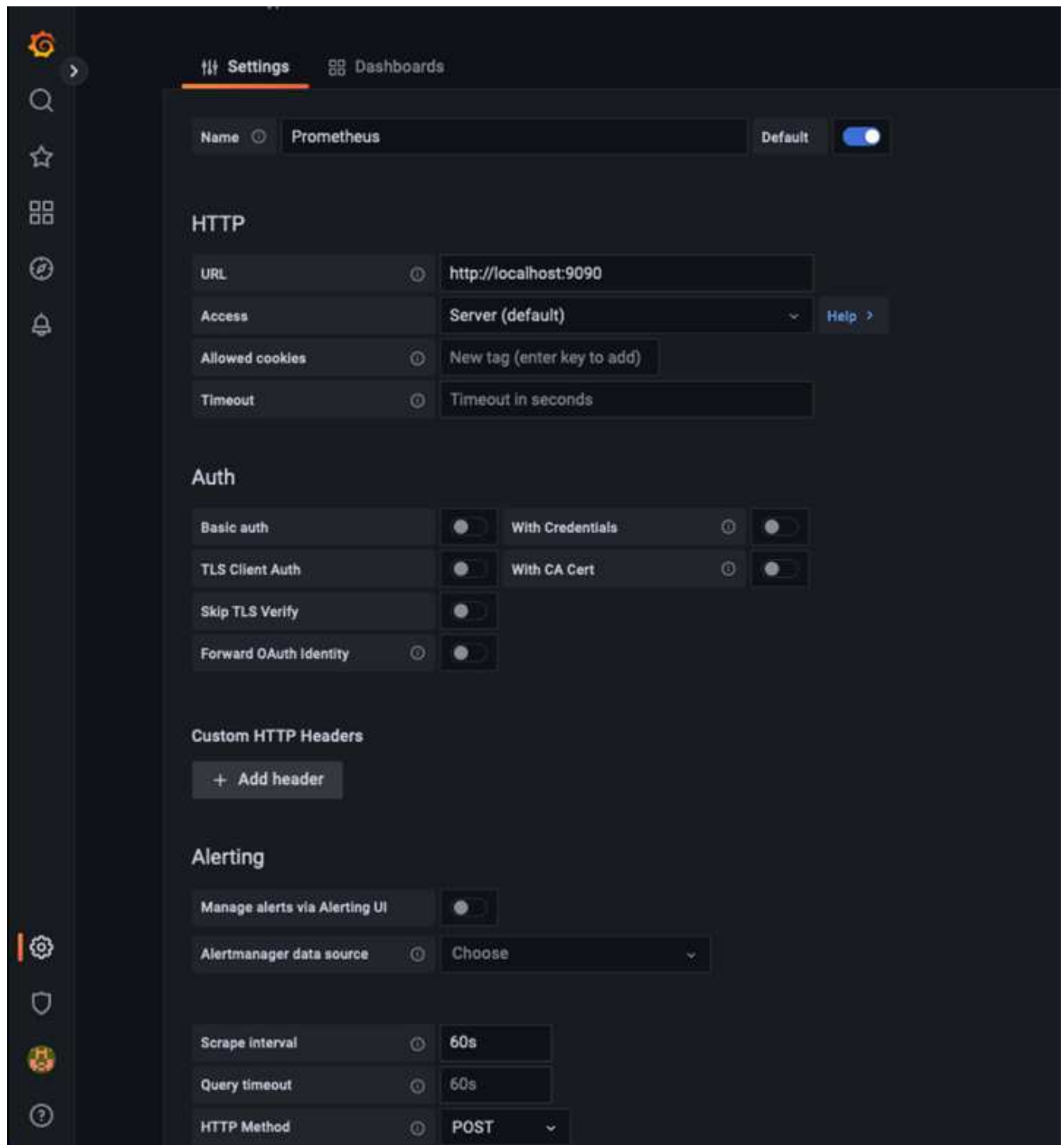
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

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"

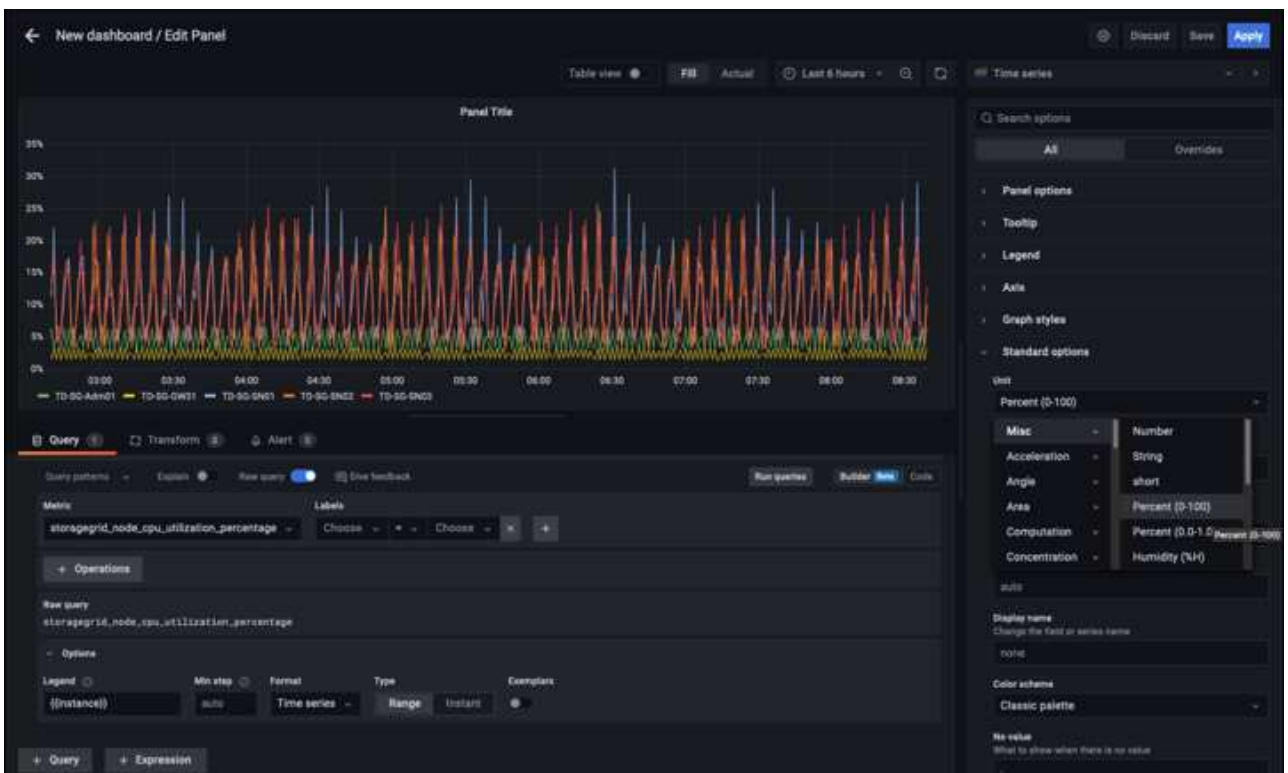


6. Ahora que tenemos configurado el origen de datos, podemos crear un panel de control.

a. En el panel izquierdo, expanda "Paneles" y seleccione "+ New Dashboard".

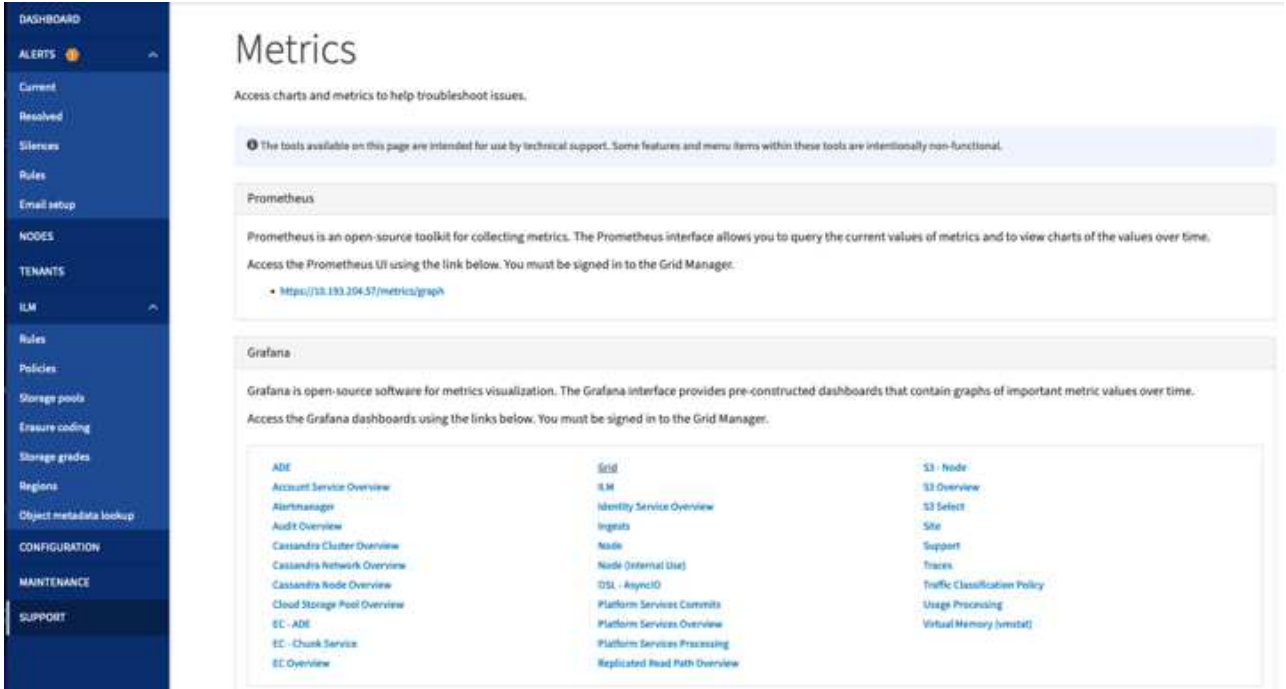
b. Seleccione "Añadir un nuevo panel"

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.

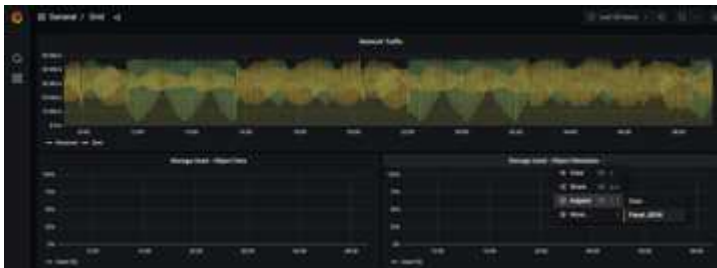


7. Podríamos seguir desarrollando nuestro panel de control como este para cada métrica que deseamos, pero por suerte StorageGRID ya dispone de paneles con paneles que podemos copiar en nuestros paneles personalizados.

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



- 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".



- Copie el código JSON y cierre la ventana.

Inspect: Storage Used - Object Metadata

4 queries with total query time of 549 ms

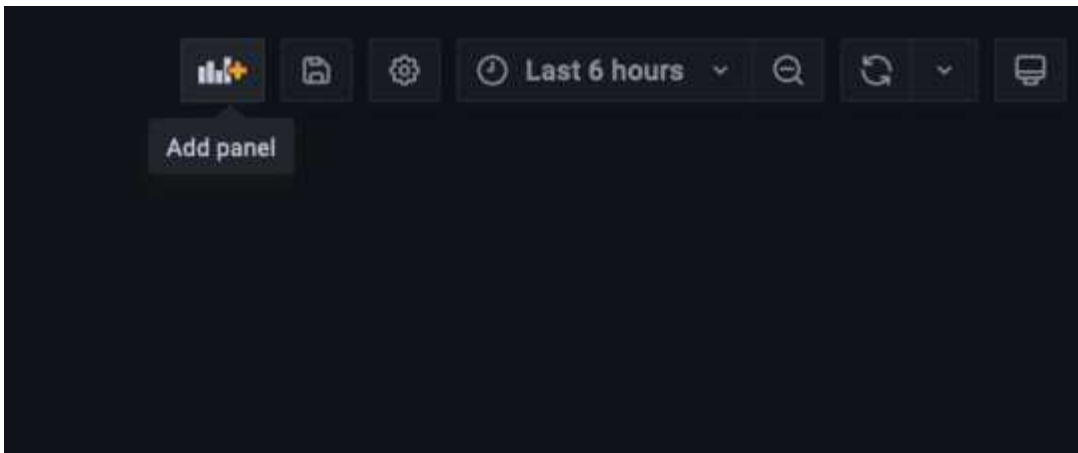
Data Stats **JSON**

Select source

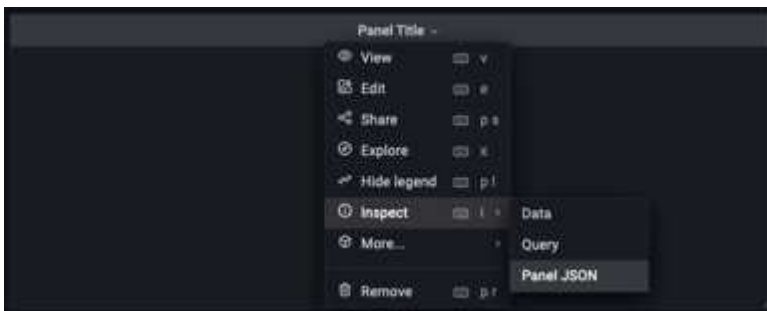
Panel JSON

```
1  [
2  "aliasColors": {},
3  "bars": false,
4  "dashLength": 10,
5  "dashes": false,
6  "datasource": "Prometheus",
7  "decimals": 2,
8  "fill": 1,
9  "fillGradient": 0,
10 "gridPos": {
11   "h": 7,
12   "w": 12,
13   "x": 12,
14   "y": 7
15 },
16 "id": 6,
17 "legend": {
18   "avg": false,
19   "current": false,
20   "max": false,
21   "min": false,
22   "show": true,
23   "total": false,
24   "values": false
25 },
26 "lines": true,
27 "linewidth": 1,
28 "links": [],
29 "nullPointMode": "null",
30 "options": {
31   "alertThreshold": true
32 },
33 "percentage": false,
34 "pointradius": 5,
35 "points": false,
36 "renderer": "flot",
37 "seriesOverrides": [
38   {
39     "alias": "Used",
```

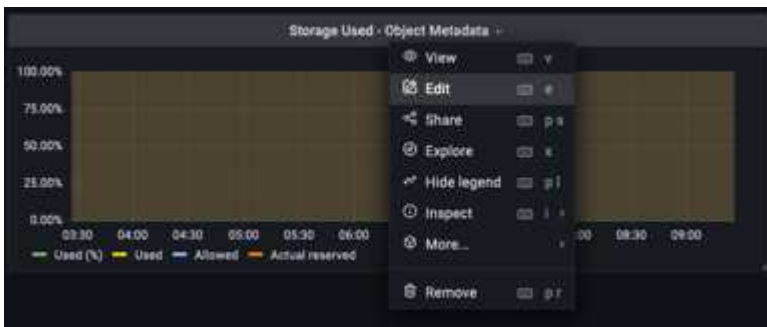
e. En nuestro nuevo panel, haga clic en el icono para añadir un nuevo panel.

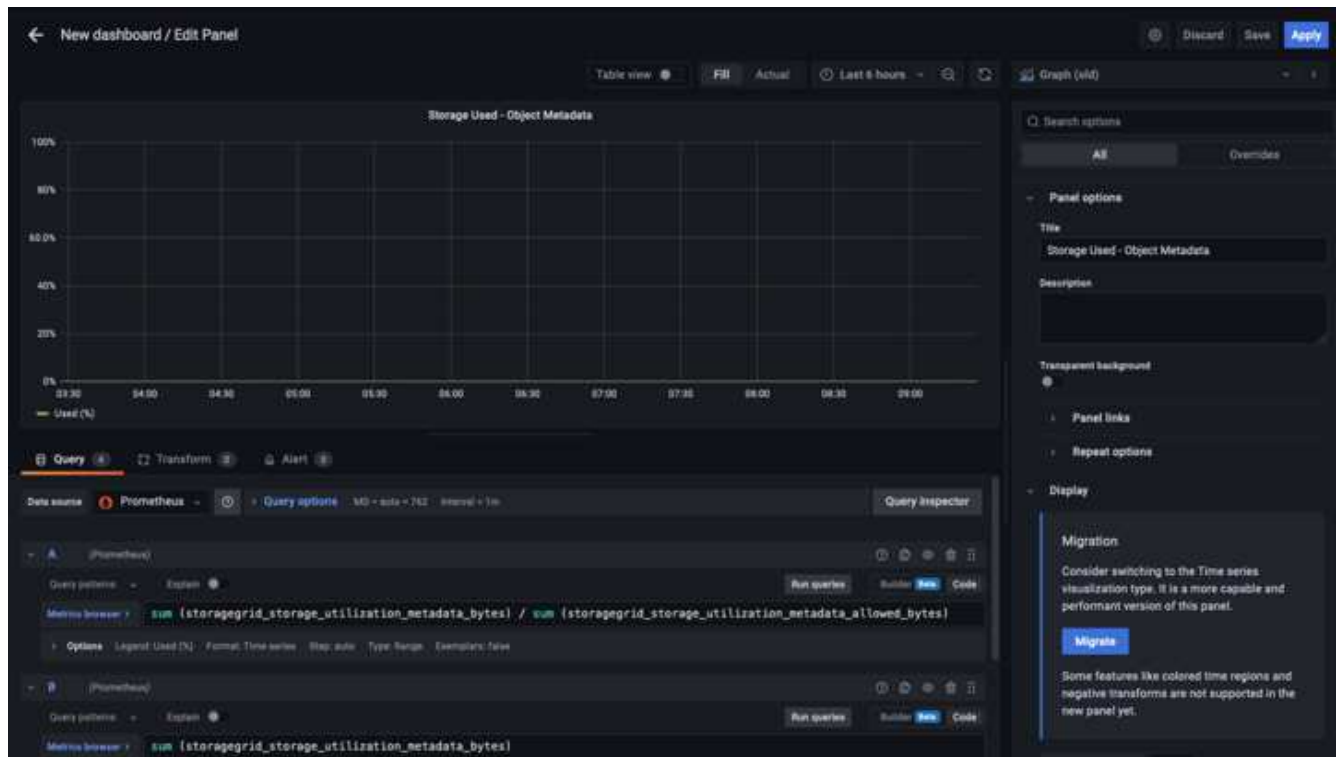


- f. Aplique el nuevo panel sin realizar cambios
- g. Al igual que con el panel StorageGRID, inspeccione el JSON. Quite todo el código JSON y sustitúyalo por el código copiado del panel StorageGRID.



- 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/datadog-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 datadog. 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/datadog-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.yml`

```

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.

SNMP Agent

You can configure SNMP for read-only MIB access and notifications. SNMPv1, SNMPv2c, SNMPv3 are supported. For SNMPv3, only User Security Model (USM) authentication is supported. All nodes in the grid share the same SNMP configuration.

Enable SNMP

System Contact

System Location

Enable SNMP Agent Notifications

Enable Authentication Traps

Community Strings

Default Trap Community

Read-Only Community

String 1 +

Other Configurations

Agent Addresses (0)

USM Users (0)

Trap Destinations (1)

+ Create Edit Remove

Version	Type	Host	Port	Protocol	Community/USM User
<input type="radio"/> SNMPv2C	Inform	10.193.92.241	9162	UDP	Default Community: st0r@gegrid

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"](https://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, Loadbalancer.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.

The screenshot shows the 'S3 Object Lock' configuration page in StorageGRID. At the top, it says 'Configuration > S3 Object Lock'. The main heading is 'S3 Object Lock'. Below the heading, there is a blue information banner that reads: 'S3 Object Lock has been enabled for the grid and cannot be disabled.' Underneath, there is a paragraph explaining that S3 Object Lock is enabled for the entire system and cannot be disabled. This is followed by a list of requirements for enabling S3 Object Lock, such as creating replicated copies and ensuring placement instructions are compliant. At the bottom, there is a checkbox labeled 'Enable S3 Object Lock' which is checked, and an 'Apply' button.


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.

Manage object settings Optional

Object versioning

Enable object versioning if you want to store every version of each object in this bucket. You can then retrieve previous versions of an object as needed.

 Object versioning has been enabled automatically because this bucket has S3 Object Lock enabled.


Enable object versioning

S3 Object Lock

S3 Object Lock allows you to specify retention and legal hold settings for the objects ingested into a bucket. If you want to use S3 Object Lock, you must enable this setting when you create the bucket. You cannot add or disable S3 Object Lock after a bucket is created.

If S3 Object Lock is enabled, object versioning is enabled for the bucket automatically and cannot be suspended.

Enable S3 Object Lock

Default retention 

Automatically protect new objects put into this bucket from being deleted or overwritten.

Disable

Enable

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

Buckets > veeam12

veeam12

Region: us-east-1
 S3 Object Lock: Enabled
 Date created: 2023-09-21 08:01:38 GMT
 Object count: 0

[View bucket contents in Experimental S3 Console](#)

[Delete objects in bucket](#) [Delete bucket](#)

Bucket options | [Bucket access](#) | [Platform services](#)

Consistency level	Read-after-new-write (default)	▼
Last access time updates	Disabled	▼
Object versioning	Enabled	▼
S3 Object Lock	Enabled	▼

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.

Bucket options | [Bucket access](#) | [Platform services](#)

Consistency level Read-after-new-write (default) ▲

Change the consistency control for operations performed on the objects in the bucket. Consistency levels provide a balance between the availability of objects and the consistency of those objects across different Storage Nodes and sites.

In general, use the **Read-after-new-write** consistency level for your buckets. Then, if objects do not meet availability or consistency requirements, change the client application's behavior, or set the Consistency-Control header for an individual API request, which overrides the bucket setting.

- All
Provides the highest guarantee of consistency. All nodes receive the data immediately, or the request will fail.
- Strong-global**
Guarantees read-after-write consistency for all client requests across all sites.
- Strong-site
Guarantees read-after-write consistency for all client requests within a site.
- Read-after-new-write (default)
Provides read-after-write consistency for new objects and eventual consistency for object updates. Offers high availability and data protection guarantees. Recommended for most cases.
- Available
Provides eventual consistency for both new objects and object updates. For S3 buckets, use only as required (for example, for a bucket that contains log values that are rarely read, or for HEAD or GET operations on keys that do not exist). Not supported for FabricPool buckets.

[Save changes](#)

Last access time updates Disabled ▼

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.

Create a traffic classification policy

You can create traffic classification policies to monitor the network traffic for specific buckets, tenants, IP addresses, subnets, or load balancer endpoints. You can optionally limit this traffic based on bandwidth, number of concurrent requests, or the request rate.

✓ Enter policy name — ✓ Add matching rules — ✓ Set limits — 4 Review the policy

Review the policy

Policy name: Veeam

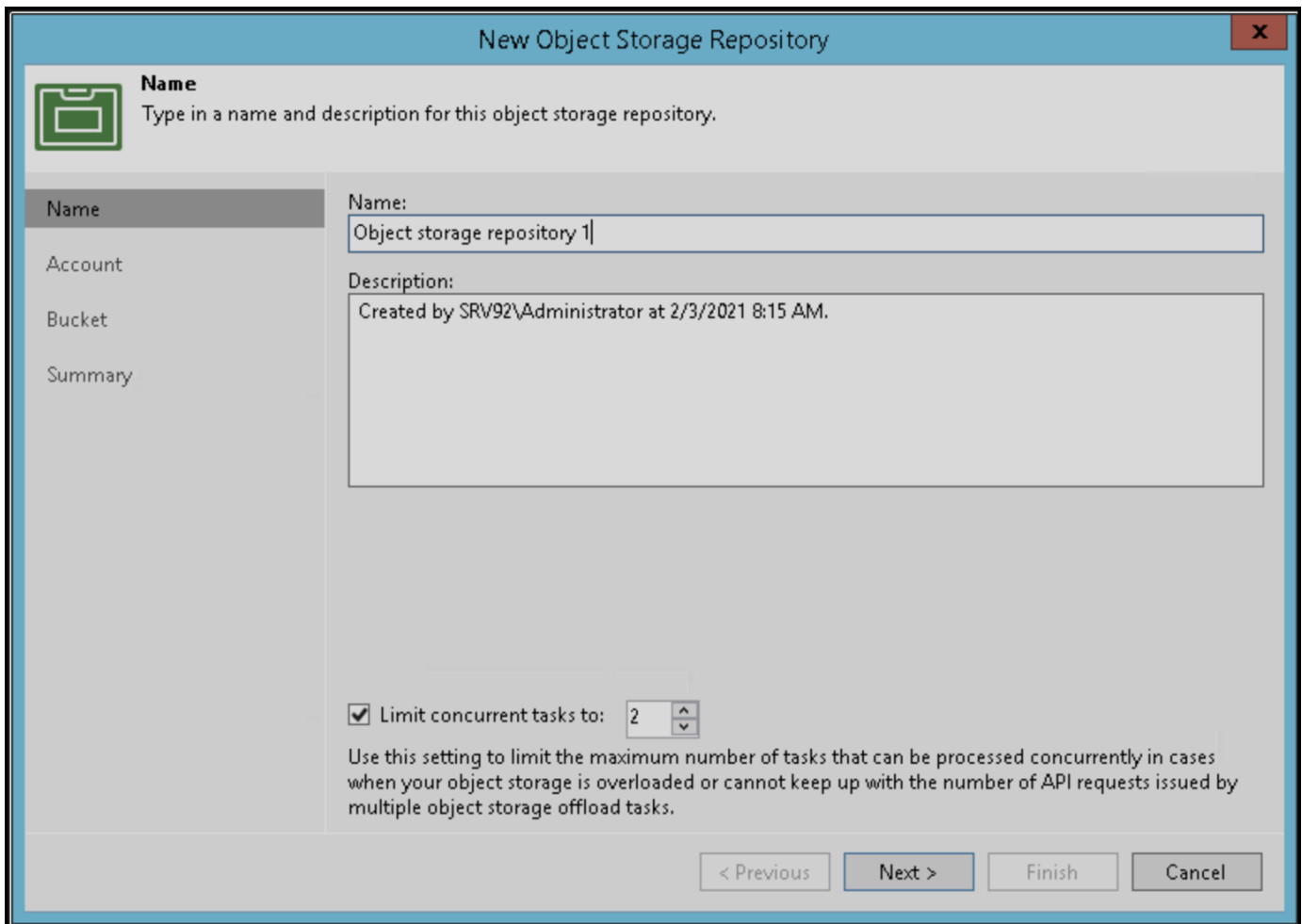
Description: Policy to monitor Veeam bucket traffic

Matching rules

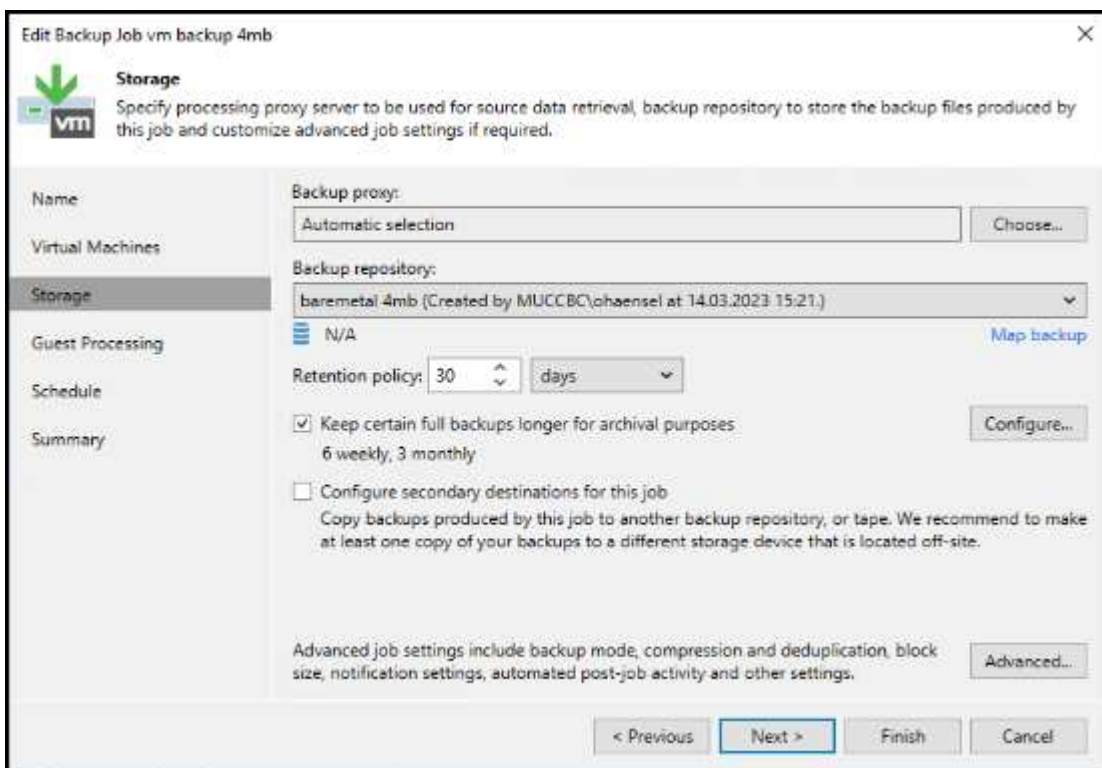
Type ?	Match value ?	Inverse match ?
Bucket	test	No

Veeam

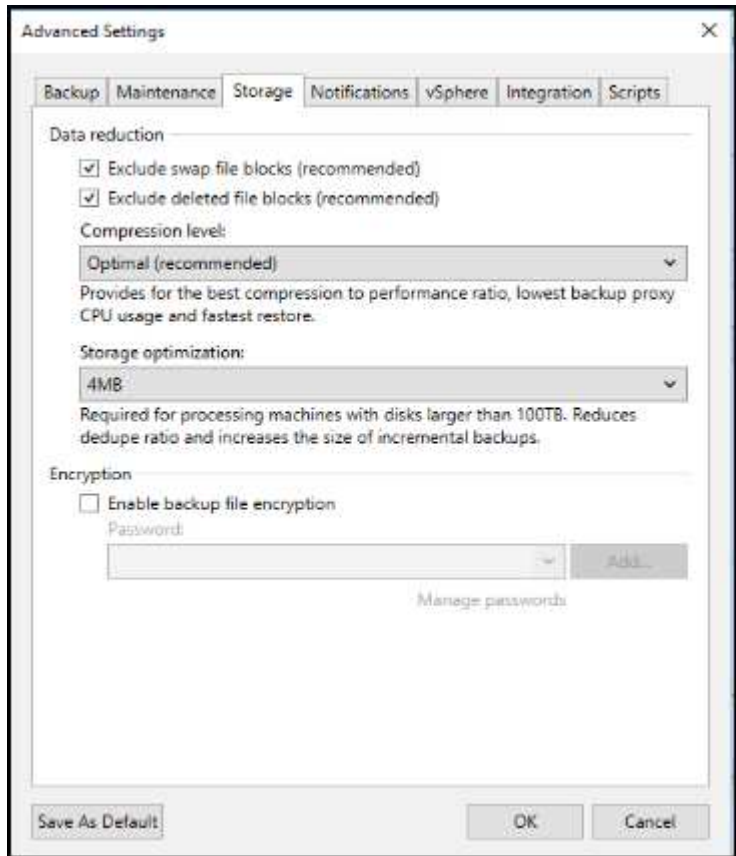
En función del modelo y la cantidad de dispositivos StorageGRID, puede que sea necesario seleccionar y configurar un límite para el número de operaciones simultáneas en el bloque.



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.

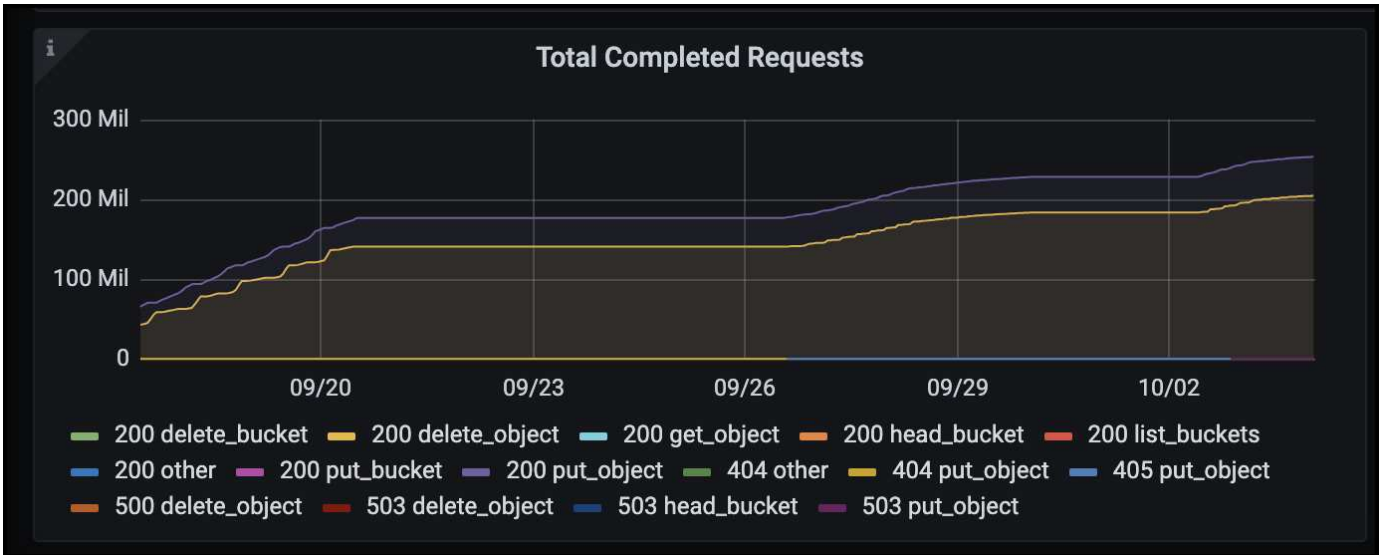
Si observa las tasas de S3 y las solicitudes activas, puede ver cuánta carga está manejando cada nodo y el número general de solicitudes por tipo.



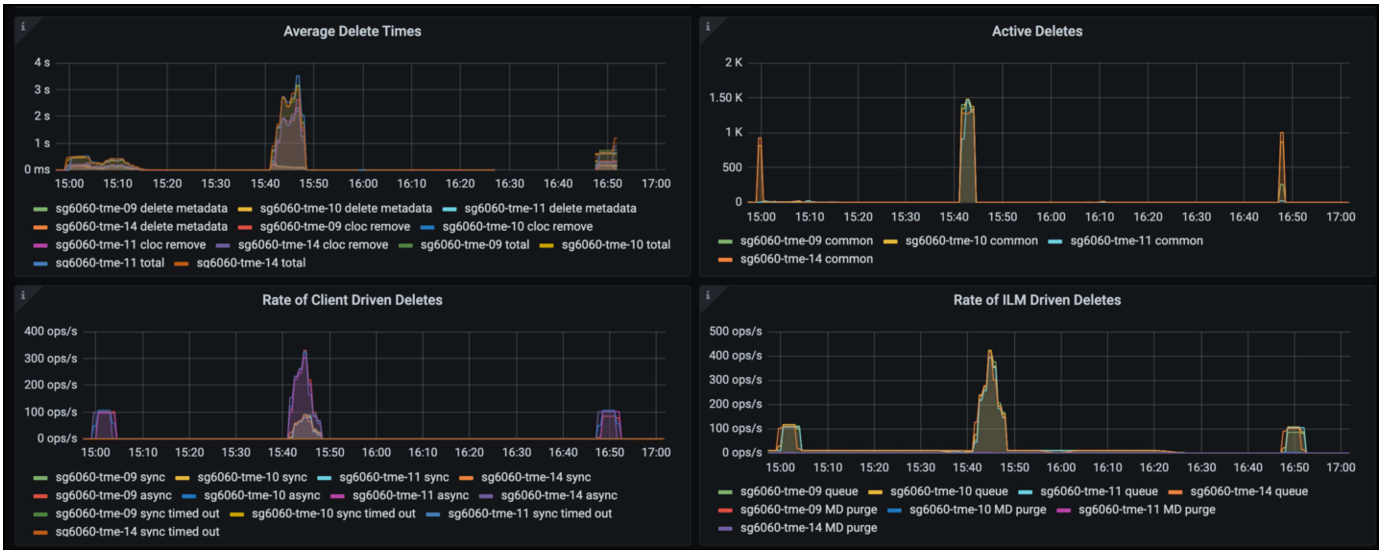
El gráfico Duración Media muestra el tiempo medio que cada nodo está tomando para cada tipo de solicitud. Esta es la latencia media de la solicitud y puede ser un buen indicador de que se puede requerir un ajuste adicional, o hay espacio para que el sistema StorageGRID asuma más carga.



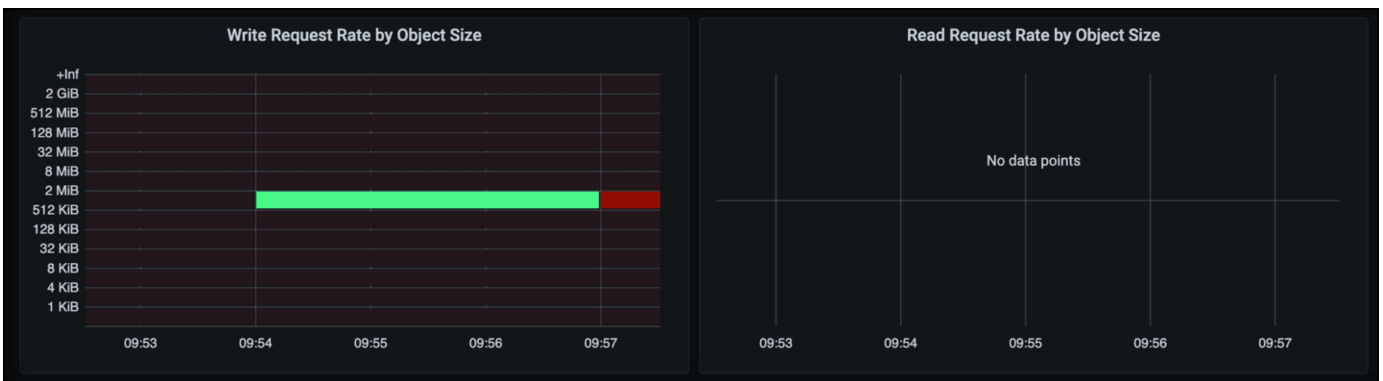
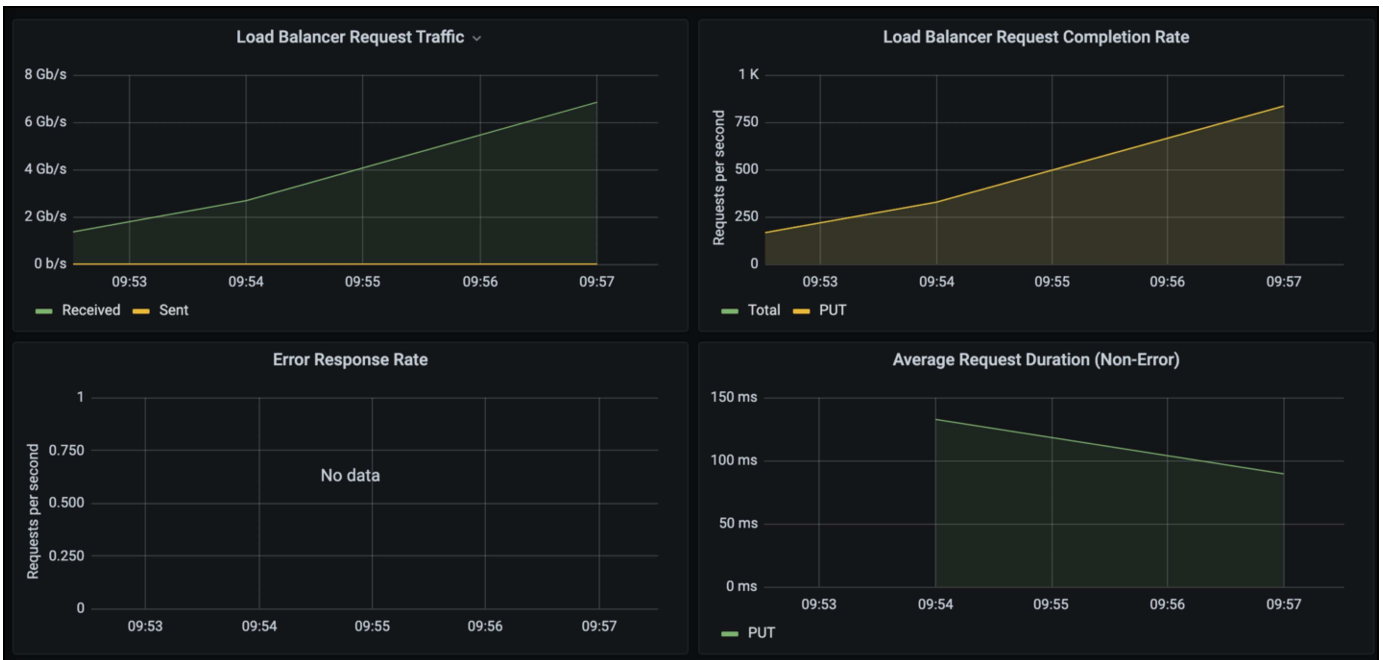
En el gráfico Total de Solicitudes Completadas, puede ver las solicitudes por tipo y códigos de respuesta. Si ve respuestas distintas de 200 (OK) para las respuestas, esto puede indicar un problema como que el sistema StorageGRID está recibiendo una carga elevada enviando respuestas 503 (Lento) y puede 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

General

Advanced Options

Reflection Refresh

Metadata

Privileges

- Enable asynchronous access when possible
- Enable compatibility mode
- Apply requester-pays to S3 requests
- Enable file status check
- Enable partition column inference

Root Path

Server side encryption key ARN

Default CTAS Format

Connection Properties

Name	Value
fs.s3a.path.style.access	true
fs.s3a.endpoint	sgdemo.netapp.com
fs.s3a.connection.maximum	1000

[+ Add property](#)

Allowlisted buckets

No allowlisted buckets added

[+ Add bucket](#)

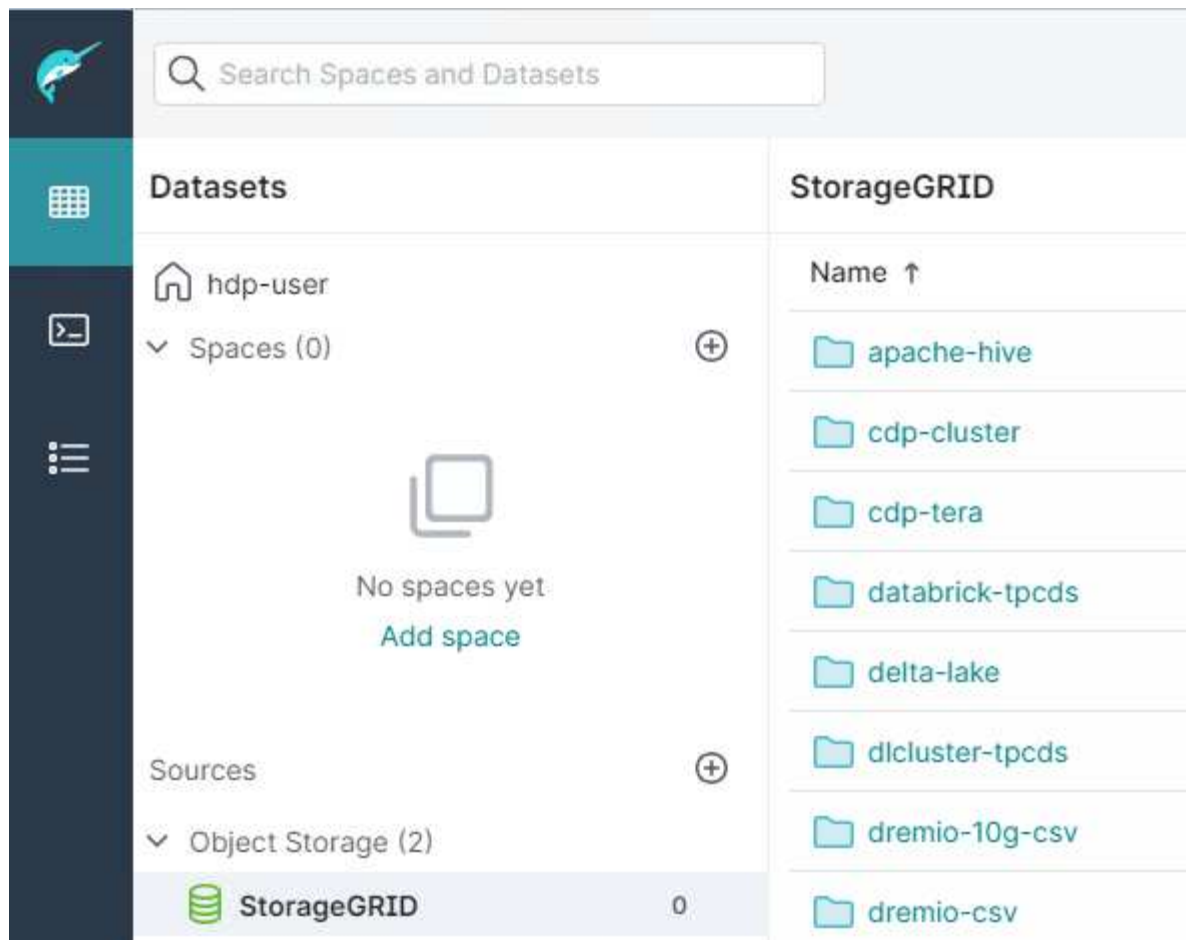
Cache Options

- Enable local caching when possible

Max percent of total available cache space to use when possible

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_style' => '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'
```

Procedimientos y ejemplos de API

Pruebe y muestre opciones de cifrado de S3 en StorageGRID

StorageGRID y la API de S3 ofrecen varias formas diferentes de cifrar sus datos en reposo. Para obtener más información, consulte ["Consulte los métodos de cifrado de StorageGRID"](#).

En esta guía se mostrarán los métodos de cifrado de la API de S3.

Cifrado del servidor (SSE)

SSE permite al cliente almacenar un objeto y cifrarlo con una clave única gestionada por StorageGRID. Cuando se solicita el objeto, la clave almacenada en StorageGRID descifra el objeto.

Ejemplo de SSE

- PONGA un objeto con SSE

```
aws s3api put-object --bucket <bucket> --key <file> --body "<file>"  
--server-side-encryption AES256 --endpoint-url https://s3.example.com
```

- DIRÍJASE al objeto para verificar el cifrado

```
aws s3api head-object --bucket <bucket> --key <file> --endpoint-url  
https://s3.example.com
```

```
{  
  "AcceptRanges": "bytes",  
  "LastModified": "2022-05-02T19:03:03+00:00",  
  "ContentLength": 47,  
  "ETag": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",  
  "ContentType": "text/plain",  
  "ServerSideEncryption": "AES256",  
  "Metadata": {}  
}
```

- OBTENGA el objeto

```
aws s3api get-object --bucket <bucket> --key <file> <file> --endpoint  
-url https://s3.example.com
```

Cifrado del servidor con claves proporcionadas por el cliente (SSE-C)

SSE permite al cliente almacenar un objeto y cifrarlo con una clave única proporcionada por el cliente con el objeto. Cuando se solicita el objeto, se debe proporcionar la misma clave para descifrar y devolver el objeto.

Ejemplo de SSE-C.

- Con fines de prueba o demostración, puede crear una clave de cifrado
 - Cree una clave de cifrado

```
openssl enc -aes-128-cbc -pass pass:secret -P`
```

```
salt=E9DBB6603C7B3D2A  
key=23832BAC16516152E560F933F261BF03  
iv =71E87C0F6EC3C45921C2754BA131A315
```

- Coloque un objeto con la clave generada

```
aws s3api put-object --bucket <bucket> --key <file> --body "file" --sse  
-customer-algorithm AES256 --sse-customer-key  
23832BAC16516152E560F933F261BF03 --endpoint-url https://s3.example.com
```

- Dirigir el objeto

```
aws s3api head-object --bucket <bucket> --key <file> --sse-customer  
-algorithm AES256 --sse-customer-key 23832BAC16516152E560F933F261BF03  
--endpoint-url https://s3.example.com
```

```
{  
  "AcceptRanges": "bytes",  
  "LastModified": "2022-05-02T19:20:02+00:00",  
  "ContentLength": 47,  
  "ETag": "\"f92ef20ab87e0e13951d9bee862e9f9a\"",  
  "ContentType": "binary/octet-stream",  
  "Metadata": {},  
  "SSECustomerAlgorithm": "AES256",  
  "SSECustomerKeyMD5": "rjGuMdjLpPV1eRuotNaPMQ=="  
}
```



Si no proporciona la clave de cifrado, recibirá un error "se ha producido un error (404) al llamar a la operación HeadObject: Not found"

- Obtenga el objeto

```
aws s3api get-object --bucket <bucket> --key <file> <file> --sse
-customer-algorithm AES256 --sse-customer-key
23832BAC16516152E560F933F261BF03 --endpoint-url https://s3.example.com
```



Si no proporciona la clave de cifrado, recibirá un error que indica que se ha producido un error (InvalidRequest) al llamar a la operación GetObject: El objeto se ha almacenado utilizando un formulario de cifrado del lado del servidor. Se deben proporcionar los parámetros correctos para recuperar el objeto."

Cifrado del servidor de bloques (SSE-S3)

SSE-S3 permite al cliente definir un comportamiento de cifrado predeterminado para todos los objetos almacenados en un bloque. Los objetos se cifran con una clave única gestionada por StorageGRID. Cuando se solicita el objeto, éste se descifra mediante una clave almacenada en StorageGRID.

Ejemplo de bloque SSE-S3

- Crear un bloque nuevo y establecer una política de cifrado predeterminada
 - Crear nuevo bloque

```
aws s3api create-bucket --bucket <bucket> --region us-east-1
--endpoint-url https://s3.example.com
```

- Put bucket Encryption

```
aws s3api put-bucket-encryption --bucket <bucket> --server-side
-encryption-configuration '{"Rules":
[{"ApplyServerSideEncryptionByDefault": {"SSEAlgorithm":
"AES256"}}]}' --endpoint-url https://s3.example.com
```

- Coloque un objeto en el bloque

```
aws s3api put-object --bucket <bucket> --key <file> --body "file"
--endpoint-url https://s3.example.com
```

- Dirigir el objeto

```
aws s3api head-object --bucket <bucket> --key <file> --endpoint-url
https://s3.example.com
```

```
{
  "AcceptRanges": "bytes",
  "LastModified": "2022-05-02T20:16:23+00:00",
  "ContentLength": 47,
  "ETag": "\"82e8fb872e778a4687a26e6c0b36bc1\"",
  "ContentType": "binary/octet-stream",
  "ServerSideEncryption": "AES256",
  "Metadata": {}
}
```

- OBTENGA el objeto

```
aws s3api get-object --bucket <bucket> --key <file> <file> --endpoint
--url https://s3.example.com
```

Por Aron Klein

Pruebe y muestre el bloqueo de objetos de S3 en StorageGRID

El bloqueo de objetos proporciona un modelo WORM para evitar que los objetos se eliminen o se sobrescriban. La implementación de StorageGRID de un bloqueo de objetos es un activo de valor empresarial que se evalúa para ayudar a cumplir los requisitos normativos, respalda la conservación legal y el modo de cumplimiento de normativas para la retención de objetos y las políticas de retención de bloques predeterminadas.

En esta guía se demostrará la API de bloqueo de objetos S3.

Conservación legal

- La retención legal de bloqueo de objetos es un estado de activación/desactivación simple aplicado a un objeto.

```
aws s3api put-object-legal-hold --bucket <bucket> --key <file> --legal
--hold Status=ON --endpoint-url https://s3.company.com
```

- Verifiquelo mediante una OPERACIÓN DE OBTENER.

```
aws s3api get-object-legal-hold --bucket <bucket> --key <file>
--endpoint-url https://s3.company.com
```



```
{
  "LegalHold": {
    "Status": "ON"
  }
}
```

- Desactivar la conservación legal

```
aws s3api put-object-legal-hold --bucket <bucket> --key <file> --legal
--hold Status=OFF --endpoint-url https://s3.company.com
```

- Verifiquelo mediante una OPERACIÓN DE OBTENER.

```
aws s3api get-object-legal-hold --bucket <bucket> --key <file>
--endpoint-url https://s3.company.com
```

```
{
  "LegalHold": {
    "Status": "OFF"
  }
}
```

Modo de cumplimiento de normativas

- La retención de objetos se realiza con una Marca de tiempo de retención hasta.

```
aws s3api put-object-retention --bucket <bucket> --key <file>
--retention '{"Mode":"COMPLIANCE", "RetainUntilDate": "2025-06-
10T16:00:00"}' --endpoint-url https://s3.company.com
```

- Compruebe el estado de retención

```
aws s3api get-object-retention --bucket <bucket> --key <file> --endpoint
-url https://s3.company.com
+
```

```
{
  "Retention": {
    "Mode": "COMPLIANCE",
    "RetainUntilDate": "2025-06-10T16:00:00+00:00"
  }
}
```

Retención predeterminada

- Establezca el período de retención en días y años en lugar de una fecha de retención hasta definida con la api por objeto.

```
aws s3api put-object-lock-configuration --bucket <bucket> --object-lock
-configuration '{"ObjectLockEnabled": "Enabled", "Rule": {
"DefaultRetention": { "Mode": "COMPLIANCE", "Days": 10 } }' --endpoint
-url https://s3.company.com
```

- Compruebe el estado de retención

```
aws s3api get-object-lock-configuration --bucket <bucket> --endpoint-url
https://s3.company.com
```

```
{
  "ObjectLockConfiguration": {
    "ObjectLockEnabled": "Enabled",
    "Rule": {
      "DefaultRetention": {
        "Mode": "COMPLIANCE",
        "Days": 10
      }
    }
  }
}
```

- Coloque un objeto en el bloque

```
aws s3api put-object --bucket <bucket> --key <file> --body "file"
--endpoint-url https://s3.example.com
```

- La duración de retención establecida en el bloque se convierte en una Marca de tiempo de retención en el objeto.

```
aws s3api get-object-retention --bucket <bucket> --key <file> --endpoint-url https://s3.company.com
```

```
{  
  "Retention": {  
    "Mode": "COMPLIANCE",  
    "RetainUntilDate": "2022-03-02T15:22:47.202000+00:00"  
  }  
}
```

Pruebe a eliminar un objeto con una retención definida

El bloqueo de objetos se crea sobre el control de versiones. La retención se define en una versión del objeto. Si se intenta eliminar un objeto con una retención definida y no se especifica ninguna versión, se crea un marcador de borrado como la versión actual del objeto.

- Elimine el objeto con la retención definida

```
aws s3api delete-object --bucket <bucket> --key <file> --endpoint-url https://s3.example.com
```

- Enumere los objetos del bloque

```
aws s3api list-objects --bucket <bucket> --endpoint-url https://s3.example.com
```

- Observe que el objeto no aparece en la lista.
- Enumere las versiones para ver el marcador de borrado y la versión original bloqueada

```
aws s3api list-object-versions --bucket <bucket> --prefix <file> --endpoint-url https://s3.example.com
```

```

{
  "Versions": [
    {
      "ETag": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",
      "Size": 47,
      "StorageClass": "STANDARD",
      "Key": "file.txt",
      "VersionId":
"RDVDMjYwMTQtQkNDQS0xMUVDLThGOEUtNjQ3NTAwQzAxQTk1",
      "IsLatest": false,
      "LastModified": "2022-04-15T14:46:29.734000+00:00",
      "Owner": {
        "DisplayName": "Tenant01",
        "ID": "56622399308951294926"
      }
    }
  ],
  "DeleteMarkers": [
    {
      "Owner": {
        "DisplayName": "Tenant01",
        "ID": "56622399308951294926"
      },
      "Key": "file01.txt",
      "VersionId":
"QjVDQzgzOTAtQ0FGNi0xMUVDLThFMzgtQ0RGMjAwQjk0MjM1",
      "IsLatest": true,
      "LastModified": "2022-05-03T15:35:50.248000+00:00"
    }
  ]
}

```

- Elimine la versión bloqueada del objeto

```

aws s3api delete-object --bucket <bucket> --key <file> --version-id
"<VersionId>" --endpoint-url https://s3.example.com

```

```

An error occurred (AccessDenied) when calling the DeleteObject
operation: Access Denied

```

Por Aron Klein

Ejemplo de políticas de bloque y grupo(IAM)

A continuación se muestran ejemplos de políticas de bloques y políticas de grupo (políticas IAM).

Políticas de grupo (IAM)

Acceso a bloque de estilo de directorio de casa

Esta política de grupo sólo permitirá a los usuarios acceder a los objetos del depósito denominado nombre de usuario de los usuarios.

```
"Statement": [
  {
    "Sid": "AllowListBucketOfASpecificUserPrefix",
    "Effect": "Allow",
    "Action": "s3:ListBucket",
    "Resource": "arn:aws:s3:::home",
    "Condition": {
      "StringLike": {
        "s3:prefix": "${aws:username}/*"
      }
    }
  },
  {
    "Sid": "AllowUserSpecificActionsOnlyInTheSpecificUserPrefix",
    "Effect": "Allow",
    "Action": "s3:*Object",
    "Resource": "arn:aws:s3:::home/?/?/${aws:username}/*"
  }
]
```

Denegar creación de bloque de bloqueo de objetos

Esta política de grupo restringirá a los usuarios a crear un bloque con el bloqueo de objeto habilitado en el bloque.



Esta política no se aplica en la interfaz de usuario de StorageGRID, sino que solo se aplica mediante la API de S3.

```

{
  "Statement": [
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::*"
    },
    {
      "Action": [
        "s3:PutBucketObjectLockConfiguration",
        "s3:PutBucketVersioning"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::*"
    }
  ]
}

```

Límite de retención de bloqueo de objetos

Esta política de depósito restringirá la duración de la retención de bloqueo de objetos a 10 días o menos

```

{
  "Version": "2012-10-17",
  "Id": "CustSetRetentionLimits",
  "Statement": [
    {
      "Sid": "CustSetRetentionPeriod",
      "Effect": "Deny",
      "Principal": "*",
      "Action": [
        "s3:PutObjectRetention"
      ],
      "Resource": "arn:aws:s3:::testlock-01/*",
      "Condition": {
        "NumericGreaterThan": {
          "s3:object-lock-remaining-retention-days": "10"
        }
      }
    }
  ]
}

```

Restringir a los usuarios la supresión de objetos por versionID

Esta política de grupo restringirá a los usuarios la supresión de objetos versionados por versionID

```
{
  "Statement": [
    {
      "Action": [
        "s3:DeleteObjectVersion"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::*"
    },
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::*"
    }
  ]
}
```

Esta política de depósito restringirá a un usuario (identificado por el ID de usuario «56622399308951294926») de eliminar objetos versionados por versionID

```

{
  "Statement": [
    {
      "Action": [
        "s3:DeleteObjectVersion"
      ],
      "Effect": "Deny",
      "Resource": "arn:aws:s3:::verdeny/*",
      "Principal": {
        "AWS": [
          "56622399308951294926"
        ]
      }
    },
    {
      "Action": "s3:*",
      "Effect": "Allow",
      "Resource": "arn:aws:s3:::verdeny/*",
      "Principal": {
        "AWS": [
          "56622399308951294926"
        ]
      }
    }
  ]
}

```

Restringir bloque a un solo usuario con acceso de sólo lectura

Esta directiva permite a un solo usuario tener acceso de sólo lectura a un bloque y denys explícitamente acceso a todos los demás usuarios. La agrupación de las declaraciones denegadas en la parte superior de la directiva es una buena práctica para una evaluación más rápida.


```

{
  "Statement": [
    {
      "Sid": "Deny non user1",
      "Effect": "Deny",
      "NotPrincipal": {
        "AWS":
"urn:sgws:identity::34921514133002833665:user/user1"
      },
      "Action": [
        "s3:*"
      ],
      "Resource": [
        "urn:sgws:s3:::bucket1",
        "urn:sgws:s3:::bucket1/*"
      ]
    },
    {
      "Sid": "Allow user1 read access to bucket bucket1",
      "Effect": "Allow",
      "Principal": {
        "AWS":
"urn:sgws:identity::34921514133002833665:user/user1"
      },
      "Action": [
        "s3:GetObject",
        "s3:ListBucket"
      ],
      "Resource": [
        "urn:sgws:s3:::bucket1",
        "urn:sgws:s3:::bucket1/*"
      ]
    }
  ]
}

```

Restrinja un grupo a un único subdirectorio (prefijo) con acceso de solo lectura

Esta política permite a los miembros del grupo tener acceso de solo lectura a un subdirectorio (prefijo) dentro de un bloque. El nombre del depósito es «study» y el subdirectorio es «study01».

```

{
  "Statement": [
    {
      "Sid": "AllowUserToSeeBucketListInTheConsole",

```

```

    "Action": [
        "s3:ListAllMyBuckets"
    ],
    "Effect": "Allow",
    "Resource": [
        "arn:aws:s3:::*"
    ]
},
{
    "Sid": "AllowRootAndstudyListingOfBucket",
    "Action": [
        "s3:ListBucket"
    ],
    "Effect": "Allow",
    "Resource": [
        "arn:aws:s3::: study"
    ],
    "Condition": {
        "StringEquals": {
            "s3:prefix": [
                "",
                "study01/"
            ],
            "s3:delimiter": [
                "/"
            ]
        }
    }
},
{
    "Sid": "AllowListingOfstudy01",
    "Action": [
        "s3:ListBucket"
    ],
    "Effect": "Allow",
    "Resource": [
        "arn:aws:s3:::study"
    ],
    "Condition": {
        "StringLike": {
            "s3:prefix": [
                "study01/*"
            ]
        }
    }
},

```

```
{
  "Sid": "AllowAllS3ActionsInstudy01Folder",
  "Effect": "Allow",
  "Action": [
    "s3:Getobject"
  ],
  "Resource": [
    "arn:aws:s3:::study/study01/*"
  ]
}
]
```

Informes técnicos

NetApp StorageGRID y análisis de Big Data

Casos de uso de NetApp StorageGRID

La solución de almacenamiento de objetos NetApp StorageGRID ofrece escalabilidad, disponibilidad de datos, seguridad y alto rendimiento. Organizaciones de todos los tamaños y sectores utilizan StorageGRID S3 para una amplia gama de casos de uso. Vamos a explorar algunos escenarios típicos:

- **Análisis de grandes volúmenes de datos:** * StorageGRID S3 se utiliza con frecuencia como un lago de datos, donde las empresas almacenan grandes cantidades de datos estructurados y no estructurados para el análisis utilizando herramientas como Apache Spark, Splunk Smartstore y Dremio.
- **Organización en niveles de datos*** Los clientes de NetApp utilizan la función FabricPool de ONTAP para mover datos automáticamente entre un nivel local de alto rendimiento a StorageGRID. La organización en niveles reserva el costoso almacenamiento flash para los datos calientes y mantiene los datos fríos disponibles en el almacenamiento de objetos de bajo coste. Esto maximiza el rendimiento y el ahorro.

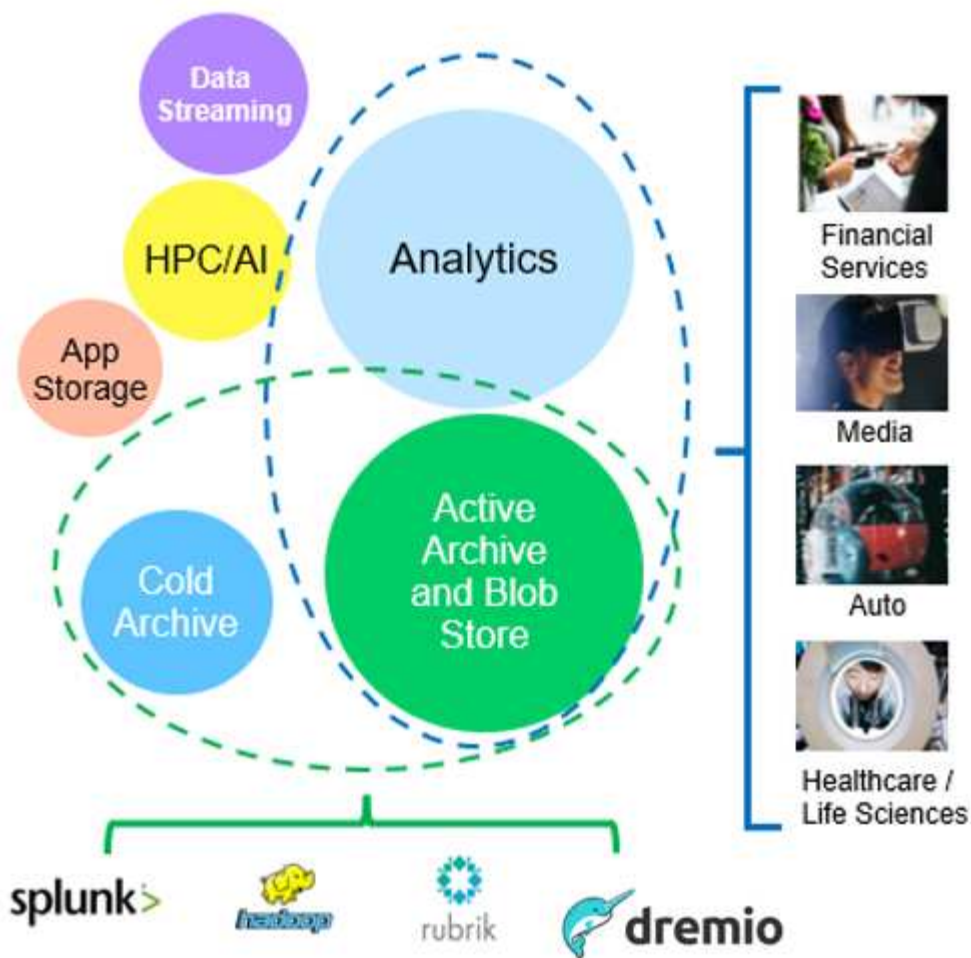
Respaldo de datos y recuperación de desastres: Las empresas pueden utilizar StorageGRID S3 como una solución confiable y rentable para hacer copias de seguridad de datos críticos y recuperarlos en caso de desastre.

Almacenamiento de datos para aplicaciones: StorageGRID S3 se puede utilizar como backend de almacenamiento para aplicaciones, lo que permite a los desarrolladores almacenar y recuperar fácilmente archivos, imágenes, videos y otros tipos de datos.

Entrega de contenido: StorageGRID S3 se puede utilizar para almacenar y entregar contenido estático del sitio web, archivos multimedia y descargas de software a usuarios de todo el mundo, aprovechando la distribución geográfica y el espacio de nombres global de StorageGRID para una entrega de contenido rápida y confiable.

- **Nivel de datos:*** Los clientes de NetApp utilizan la función ONTAP FabricPool para mover automáticamente datos entre un nivel local de alto rendimiento a StorageGRID. La organización en niveles reserva el costoso almacenamiento flash para datos calientes y mantiene los datos fríos disponibles en el almacenamiento de objetos de bajo coste. Esto maximiza el rendimiento y el ahorro.
- **Archivo de datos:*** StorageGRID ofrece diferentes tipos de almacenamiento y admite la clasificación por niveles en opciones de almacenamiento público de bajo costo a largo plazo, lo que lo convierte en una solución ideal para el archivado y la retención a largo plazo de datos que deben conservarse para fines de cumplimiento o históricos.

Casos de uso de almacenamiento de objetos

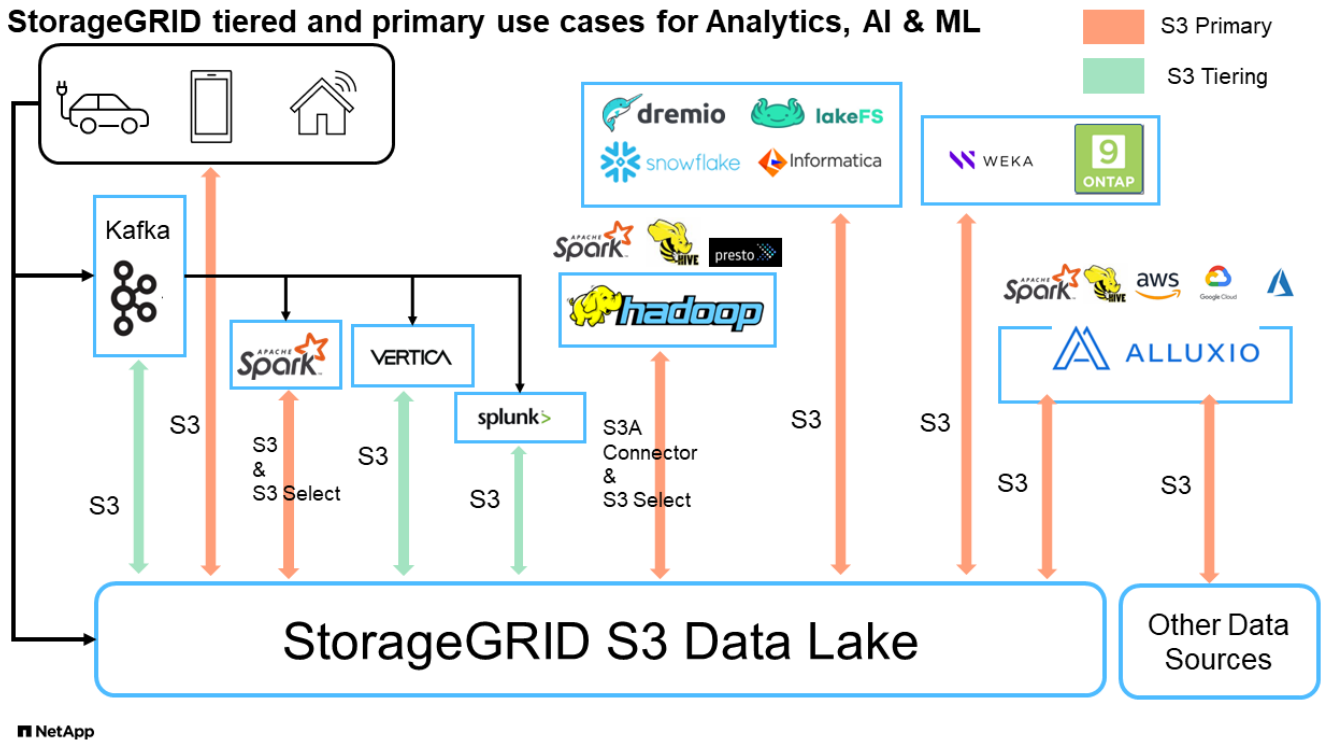


Entre los anteriores, el análisis de Big Data es uno de los casos de uso más importantes y su uso es tendencia al alza.

¿Por qué elegir StorageGRID para lagos de datos?

- Mayor colaboración: Multi-tenancy compartido masivo con acceso a API estándar del sector
- Costes operativos reducidos: Sencillez operativa de una única arquitectura de escalado horizontal automatizada y de reparación automática
- Escalabilidad: A diferencia de las soluciones tradicionales Hadoop y almacenes de datos, el almacenamiento de objetos S3 de StorageGRID separa el almacenamiento de la computación y los datos, lo que permite a la empresa escalar sus necesidades de almacenamiento a medida que crecían.
- Durabilidad y fiabilidad: StorageGRID proporciona una durabilidad del 99,999999999 %, lo que significa que los datos almacenados son muy resistentes a la pérdida de datos. También ofrece una alta disponibilidad, lo que garantiza la accesibilidad de los datos en todo momento.
- Seguridad: StorageGRID ofrece distintas funciones de seguridad, como el cifrado, la política de control de acceso, la gestión del ciclo de vida de los datos, el bloqueo de objetos y el control de versiones para proteger los datos almacenados en bloques de S3

StorageGRID S3 lagos de datos



Qué almacén de datos o lago de datos funcionan mejor con el almacenamiento de objetos S3

NetApp comparó a StorageGRID con tres ecosistemas de almacenes de datos/casas de lago - Colmena, Lago Delta y Dremio. "[Apache Iceberg: La guía definitiva](#)" incluye una breve introducción del almacén de datos y la casa del lago de datos y pro/cons de estas dos arquitecturas.

- Herramienta de referencia - TPC-DS - <https://www.tpc.org/tpcds/>
- Ecosistemas de Big Data
 - Un clúster de 5 equipos virtuales, cada uno con 128G GB de RAM y 24 vCPU, almacenamiento SSD para el disco del sistema
 - Hadoop 3.3.5 con Hive 3.1.3 (1 nodo de nombres + 4 nodos de datos)
 - Delta Lake con Spark 3.2.0 (1 maestro + 4 trabajadores) y Hadoop 3.3.5
 - Dremio V23 (1 master + 4 ejecutores)
- Almacenamiento de objetos
 - NetApp® StorageGRID® 11,6 con equilibrador de carga 3 x SG6060 + 1x SG1000
 - Protección de objetos: 2 copias
- Tamaño de base de datos 1000GB
- Se ha desactivado la caché en los 3 ecosistemas para obtener un resultado consistente para cada prueba de consulta.

TPC-DS incluye 99 consultas SQL complejas para la evaluación comparativa de consultas. Medimos los minutos totales para completar las 99 consultas y profundizamos desglosando el tipo y el número de solicitudes S3 para analizar el resultado. La primera tabla a continuación muestra la duración total de las 99 consultas y la segunda tabla resume el número y los tipos de S3 solicitudes enviadas a StorageGRID por cada ecosistema.

Resultado de consulta TPC-DS

Ecosistema	Subárbol	Lago Delta	Dremio
Capa de almacenamiento	NetApp® StorageGRID®	NetApp® StorageGRID®	NetApp® StorageGRID®
Tipo de unidad	HDD	HDD	HDD
Formato de tabla	Parquet	Parquet	Parquet ¹
Tamaño de la base de datos	1000G	1000G	1000G
TPCDS 99 consultas total de minutos	1084 ²	55	47

¹ Probado el formato de tabla Parquet e Iceberg, el resultado es similar.

² Hive no ha podido completar la consulta número 72.

Consultas TPC-DS - S3 solicitudes de desglose

S3 000 solicitudes	Subárbol	Lago Delta	Dremio
OBTENGA	1.117.184	2.074.610	4.414.227
observación: Todas las gamas GET	80% rango de obtención de 2KB a 2MB de 32MB objetos, 50 - 100 solicitudes/seg	El rango del 73% se obtiene por debajo de 100KB de 32MB objetos, de 1000 a 1400 solicitudes por segundo	90% 1M bytes de rango de obtención de 256MB objetos, 2000 - 2300 solicitudes/seg
Mostrar objetos	312.053	24.158	240
CABEZAL (objeto inexistente)	156.027	12.103	192
CABEZAL (objeto existente)	982.126	922.732	1.845
Total de solicitudes	2.567.390	3.033.603	4.416.504

Desde la primera mesa, podemos ver Delta Lake y Dremio son mucho más rápidos que Hive. Desde la segunda tabla, notamos que Hive envió muchas solicitudes de objetos de lista S3, lo cual suele ser lento en todas las plataformas de almacenamiento de objetos, especialmente si se trata de un cubo que contiene muchos objetos. Esto aumenta significativamente la duración general de la consulta. Otra observación es que Dremio fue capaz de enviar un gran número de SOLICITUDES GET en paralelo, de 2.000 a 2.300 solicitudes por segundo frente a 50 - 100 solicitudes por segundo en Hive. El sistema de archivos estándar mimic de Hive y Hadoop S3A contribuye a la lentitud de Hive para el almacenamiento de objetos S3.

El uso de Hadoop (ya sea en HDFS o en el almacenamiento de objetos S3) con Hive o Spark requiere un amplio conocimiento de Hadoop y Hive/Spark y cómo interactúan los ajustes de cada servicio; juntos tienen más de 1000 ajustes. Muy a menudo, los ajustes están interrelacionados y no se pueden cambiar solos. Se necesita una gran cantidad de tiempo y esfuerzo para encontrar la combinación óptima de ajustes y valores para usar.

Dremio es un motor de lago de datos que utiliza Apache Arrow de extremo a extremo para aumentar drásticamente el rendimiento de las consultas. Apache Arrow proporciona un formato de memoria columnar

estandarizado para compartir datos de forma eficiente y realizar análisis rápidos. Arrow emplea un enfoque independiente del lenguaje, diseñado para eliminar la necesidad de serialización y deserialización de datos, mejorando el rendimiento y la interoperabilidad entre los sistemas y procesos de datos complejos.

El rendimiento de Dremio se basa principalmente en la potencia de cálculo en el clúster Dremio. Aunque Dremio utiliza el conector S3A de Hadoop para la conexión de almacenamiento de objetos S3, no se requiere Hadoop y Dremio no utiliza la mayoría de los ajustes fs.S3A de Hadoop. Esto hace que el ajuste del rendimiento de Dremio sea fácil sin perder tiempo para aprender y probar varios ajustes de Hadoop S3A.

A partir de este resultado de las pruebas de rendimiento, podemos concluir que el sistema de análisis de Big Data optimizado para cargas de trabajo basadas en S3 es un factor de rendimiento principal. Dremio optimiza la ejecución de consultas, utiliza metadatos de manera eficiente y proporciona un acceso fluido a datos S3, lo que resulta en un mejor rendimiento en comparación con Hive cuando se trabaja con almacenamiento S3. Consulte este apartado "[página](#)" Para configurar el origen de datos Dremio S3 con StorageGRID.

Visite los enlaces siguientes para obtener más información sobre cómo StorageGRID y Dremio trabajan juntos para proporcionar una infraestructura de lago de datos moderna y eficiente y cómo NetApp migró de Hive + HDFS a Dremio + StorageGRID para mejorar drásticamente la eficiencia del análisis de Big Data.

- ["Impulse el rendimiento de sus Big Data con NetApp StorageGRID"](#)
- ["Infraestructura de lago de datos moderna, potente y eficiente con StorageGRID y Dremio"](#)
- ["Cómo NetApp está redefiniendo la experiencia del cliente con el análisis de productos"](#)

Ajuste Hadoop S3A

El conector S3A de Hadoop facilita la interacción fluida entre aplicaciones basadas en Hadoop y almacenamiento de objetos S3. El ajuste del conector Hadoop S3A es esencial para optimizar el rendimiento cuando se trabaja con el almacenamiento de objetos S3. Antes de entrar en el ajuste de detalles, entendamos lo básico de Hadoop y sus componentes.

¿Qué es Hadoop?

- Hadoop * es un potente marco de código abierto diseñado para gestionar el procesamiento y almacenamiento de datos a gran escala. Permite el almacenamiento distribuido y el procesamiento paralelo entre clústeres de equipos.

Los tres componentes principales de Hadoop son:

- **Hadoop HDFS (Hadoop Distributed File System):** Se encarga del almacenamiento, dividiendo los datos en bloques y distribuyéndolos a través de los nodos.
- **Hadoop MapReduce:** Responsable del procesamiento de datos dividiendo las tareas en fragmentos más pequeños y ejecutándolas en paralelo.
- * Hadoop YARN (Otro Negociador de Recursos):* ["Gestiona los recursos y programa las tareas de forma eficiente"](#)

Hadoop HDFS y conector S3A

HDFS es un componente vital del ecosistema de Hadoop, y tiene un papel crucial en el procesamiento eficiente de Big Data. HDFS permite un almacenamiento y una gestión fiables. Garantiza el procesamiento paralelo y un almacenamiento de datos optimizado, lo que acelera el acceso y el análisis de los datos.

En el procesamiento de Big Data, HDFS ofrece almacenamiento con tolerancia a fallos para grandes

conjuntos de datos. Y todo ello gracias a la replicación de datos. Puede almacenar y gestionar grandes volúmenes de datos estructurados y no estructurados en un entorno de almacén de datos. Además, se integra sin problemas con los principales marcos de procesamiento de Big Data, como Apache Spark, Hive, Pig y Flink, lo que permite un procesamiento de datos escalable y eficiente. Es compatible con sistemas operativos basados en Unix (Linux), por lo que es una opción ideal para las organizaciones que prefieren utilizar entornos basados en Linux para su procesamiento de Big Data.

A medida que ha ido creciendo el volumen de datos con el tiempo, el enfoque de añadir nuevas máquinas al clúster Hadoop con sus propios recursos informáticos y de almacenamiento se ha vuelto ineficiente. Escalar de forma lineal crea retos para usar los recursos de forma eficiente y gestionar la infraestructura.

Para abordar estos retos, el conector Hadoop S3A ofrece I/O de alto rendimiento frente al almacenamiento de objetos de 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 para 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.

Ajuste de conector Hadoop S3A

S3 se comporta de forma diferente a HDFS, y algunos intentos de preservar la apariencia de un sistema de archivos están excesivamente subóptimos. Es necesario realizar ajustes, pruebas y experimentos cuidadosos para hacer el uso más eficiente de los recursos de S3.

Las opciones de Hadoop incluidas en este documento se basan en Hadoop 3,3.5, consulte "[Hadoop 3.3.5 core-site.xml](#)" para todas las opciones disponibles.

Nota: El valor predeterminado de algunas configuraciones de Hadoop fs.S3A es diferente en cada versión de Hadoop. Asegúrese de consultar el valor predeterminado específico de su versión actual de Hadoop. Si no se especifica esta configuración en Hadoop core-site.xml, se utilizará el valor predeterminado. Puede anular el valor en tiempo de ejecución con las opciones de configuración de Spark o Hive.

Tienes que ir a esto "[Página de Apache Hadoop](#)" para entender cada opción fs.s3a. Si es posible, pruébalos en un clúster Hadoop que no sea de producción para encontrar los valores óptimos.

Deberías leer "[Maximizar el rendimiento cuando se trabaja con el conector S3A](#)" para otras recomendaciones de ajuste.

Veamos algunas consideraciones clave:

1. Compresión de datos

No active la compresión StorageGRID. La mayoría de los sistemas de Big Data utilizan un rango de bytes GET en lugar de recuperar todo el objeto. El uso de un rango de bytes GET con objetos comprimidos reduce considerablemente el rendimiento GET.

2. S3A comités

En general, Magic S3A committer se recomienda. Consulte este apartado "[Página de opciones comunes de S3A committer](#)" para obtener una mejor comprensión de magic committer y sus configuraciones s3a relacionadas.

Responsable de Magic:

El Magic committer confía específicamente en S3Guard para ofrecer listados de directorios consistentes en el almacén de objetos de S3.

Con S3 consistente (que ahora es el caso), el comensur Magic se puede usar de forma segura con cualquier cubo S3.

Opciones y experimentación:

En función de su caso de uso, puede elegir entre el comité de almacenamiento provisional (que se basa en un sistema de archivos HDFS del clúster) y el comité mágico.

Experimente con ambos para determinar cuál se adapta mejor a su carga de trabajo y sus requisitos.

En resumen, los Comités S3A ofrecen una solución al desafío fundamental de un compromiso de producción consistente, de alto rendimiento y fiable con S3. Su diseño interno garantiza una transferencia de datos eficiente al tiempo que mantiene la integridad de los datos.

Option	Meaning	Default (Hadoop 3.3.5)
fs.s3a.committer.name	Committer to create for output to S3A, one of: "file", "directory", "partitioned", "magic".	file
fs.s3a.buffer.dir	Local filesystem directory for data being written and/or staged.	\${env.LOCAL_DIRS:-\${hadoop.tmp.dir}}/s3a
fs.s3a.committer.magic.enabled	Enable "magic committer" support in the filesystem.	true
fs.s3a.committer.abort.pending.uploads	list and abort all pending uploads under the destination path when the job is committed or aborted.	true
fs.s3a.committer.threads	Number of threads in committers for parallel operations on files.	8
fs.s3a.committer.generate.uuid	Generate a Job UUID if none is passed down from Spark	false
fs.s3a.committer.require.uuid	Require the Job UUID to be passed down from Spark	false
mapreduce.fileoutputcommitter.marksuccessfuljobs	Write a _SUCCESS file on the successful completion of the job.	true
mapreduce.outputcommitter.factory.scheme.s3a	The committer factory to use when writing data to S3A filesystems. If mapreduce.outputcommitter.factory.class is set, it will override this property. (This property is set in mapred-default.xml)	org.apache.hadoop.fs.s3a.commit.S3ACommitterFactory

3. Thread, tamaños de pool de conexiones y tamaño de bloque

- Cada cliente **S3A** que interactúa con un solo depósito tiene su propio conjunto dedicado de conexiones HTTP 1,1 abiertas e hilos para las operaciones de carga y copia.
- ["Puede ajustar estos tamaños de pool para lograr un equilibrio entre el rendimiento y el uso de memoria/thread"](#).
- Al cargar datos a S3, se divide en bloques. El tamaño de bloque predeterminado es de 32 MB. Puede personalizar este valor configurando la propiedad fs.S3A.block.size.
- Los bloques mayores pueden mejorar el rendimiento de las cargas de datos grandes al reducir la sobrecarga que supone gestionar piezas de varias partes durante la carga. El valor recomendado es de 256 MB o superior para un conjunto de datos grande.

Option	Meaning	Default (Hadoop 3.3.5)
fs.s3a.threads.max	The total number of threads available in the filesystem for data uploads *or any other queued filesystem operation*.	64
fs.s3a.connection.maximum	Controls the maximum number of simultaneous connections to S3. This must be bigger than the value of fs.s3a.threads.max so as to stop threads being blocked waiting for new HTTPS connections. Why not equal? The AWS SDK transfer manager also uses these connections.	96
fs.s3a.max.total.tasks	The number of operations which can be queued for execution. This is in addition to the number of active threads in fs.s3a.threads.max.	32
fs.s3a.committer.threads	Number of threads in committers for parallel operations on files (upload, commit, abort, delete...)	8
fs.s3a.executor.capacity	The maximum number of submitted tasks which is a single operation (e.g. rename(), delete()) may submit simultaneously for execution -excluding the IO-heavy block uploads, whose capacity is set in "fs.s3a.fast.upload.active.blocks" All tasks are submitted to the shared thread pool whose size is set in "fs.s3a.threads.max"; the value of capacity should be less than that of the thread pool itself, as the goal is to stop a single operation from overloading that thread pool.	16
fs.s3a.fast.upload.active.blocks (see also related fs.s3a.fast.upload.buffer option)	Maximum Number of blocks a single output stream can have active (uploading, or queued to the central FileSystem instance's pool of queued operations. This stops a single stream overloading the shared thread pool.	4
fs.s3a.block.size	Block size to use when reading files using s3a: file system. A suffix from the set {K,M,G,T,P} may be used to scale the numeric value.	32MB (tested 1TB data set with 256MB and 512MB block size shows significant improvement in both read and write)

4. Carga multiparte

Los responsables de S3A **SIEMPRE** usan MPU (carga multiparte) para cargar datos al cubo S3. Esto es necesario para permitir: Fallo de tarea, ejecución especulativa de tareas y abortos de trabajo antes de la confirmación. A continuación se indican algunas especificaciones clave relacionadas con las cargas de varias partes:

- Tamaño máximo de objeto: 5 TiB (terabytes).
- Número máximo de piezas por carga: 10.000.
- Números de referencia: Desde 1 hasta 10.000 (inclusive).
- Tamaño de la pieza: Entre 5 MiB y 5 GiB. Cabe destacar que no hay límite de tamaño mínimo para la última parte de la carga de varias partes.

El uso de un tamaño de pieza más pequeño para cargas de varias partes S3 tiene ventajas y desventajas.

Ventajas:

- Recuperación rápida de problemas de red: Al cargar piezas más pequeñas, se minimiza el impacto de reiniciar una carga fallida debido a un error de red. Si una pieza falla, solo necesita volver a cargar esa pieza específica en lugar de todo el objeto.

- **Mejor Paralelización:** Se pueden subir más partes en paralelo, aprovechando las conexiones multi-threading o concurrentes. Esta paralelización mejora el rendimiento, sobre todo cuando se trata de archivos grandes.

Desventaja:

- **Sobrecarga de red:** El tamaño de la pieza más pequeño significa más partes para cargar, cada parte requiere su propia solicitud HTTP. Más solicitudes HTTP aumentan la sobrecarga de iniciar y completar solicitudes individuales. La gestión de un gran número de piezas pequeñas puede afectar al rendimiento.
- **Complejidad:** Gestionar el pedido, realizar un seguimiento de las piezas y garantizar que las cargas sean satisfactorias puede resultar engorroso. Si es necesario anular la carga, se debe realizar un seguimiento y depurar todos los artículos que ya se han cargado.

Para Hadoop, se recomienda un tamaño de pieza de 256MB o superior para `fs.S3A.multipart.size`. Defina siempre el valor `fs.S3A.multipart.threshold` en $2 \times fs.S3A.multipart.size$. Por ejemplo, si `fs.S3A.multipart.size = 256M`, `fs.S3A.multipart.threshold` debe ser 512M.

Utilice un tamaño de pieza más grande para un conjunto de datos grande. Es importante elegir un tamaño de pieza que equilibre estos factores en función de su caso de uso específico y las condiciones de red.

Una carga de varias partes es un "[proceso de tres pasos](#)":

1. Se inicia la carga, StorageGRID devuelve un ID de carga.
2. Las partes del objeto se cargan mediante el identificador de carga.
3. Una vez que se han cargado todas las partes del objeto, envía una solicitud de carga completa de varias partes con `upload-id`. StorageGRID construye el objeto a partir de las piezas cargadas, y el cliente puede acceder al objeto.

Si la solicitud completa de carga de varias partes no se envía correctamente, las piezas permanecen en StorageGRID y no crearán ningún objeto. Esto ocurre cuando los trabajos se interrumpen, fallan o se anulan. Los artículos permanecen en la cuadrícula hasta que la carga de varias partes se completa o se anula o StorageGRID depura estos artículos si han transcurrido 15 días desde que se inició la carga. Si hay muchas (unos pocos cientos de miles o millones) cargas multiparte en curso en un depósito, cuando Hadoop envía «lista-multiparte-cargas» (esta solicitud no filtra por identificador de carga), la solicitud puede tardar mucho tiempo en completarse o eventualmente en agotarse. Puede considerar establecer `fs.S3A.multipart.purge` en `true` con un valor `fs.S3A.multipart.purge.age` apropiado (por ejemplo, 5 a 7 días, no utilice el valor predeterminado de 86400, es decir, 1 día). O póngase en contacto con el servicio de soporte de NetApp para investigar la situación.

Option	Meaning	Default (Hadoop 3.3.5)
fs.s3a.multipart.size	How big (in bytes) to split upload or copy operations up into. A suffix from the set {K,M,G,T,P} may be used to scale the numeric value.	64M
fs.s3a.multipart.threshold	How big (in bytes) to split upload or copy operations up into. This also controls the partition size in renamed files, as rename() involves copying the source file(s). A suffix from the set {K,M,G,T,P} may be used to scale the numeric value.	128M
fs.s3a.multipart.purge	True if you want to purge existing multipart uploads that may not have been completed/aborted correctly. The corresponding purge age is defined in fs.s3a.multipart.purge.age. If set, when the filesystem is instantiated then all outstanding uploads older than the purge age will be terminated -across the entire bucket. This will impact multipart uploads by other applications and users. so should be used sparingly, with an age value chosen to stop failed uploads, without breaking ongoing operations.	false
fs.s3a.multipart.purge.age	Minimum age in seconds of multipart uploads to purge on startup if "fs.s3a.multipart.purge" is true	86400

5. Buffer escribe datos en la memoria

Para mejorar el rendimiento, puede almacenar en búfer los datos de escritura en la memoria antes de cargarlos en S3. Esto puede reducir el número de escrituras pequeñas y mejorar la eficiencia.

Option	Meaning	Default (Hadoop 3.3.5)
fs.s3a.fast.upload.buffer	The buffering mechanism to for data being written. Values: disk, array, bytearray. "disk" will use the directories listed in fs.s3a.buffer.dir as the location(s) to save data prior to being uploaded. "array" uses arrays in the JVM heap "bytebuffer" uses off-heap memory within the JVM. Both "array" and "bytebuffer" will consume memory in a single stream up to the number of blocks set by: fs.s3a.multipart.size * fs.s3a.fast.upload.active.blocks. If using either of these mechanisms, keep this value low The total number of threads performing work across all threads is set by fs.s3a.threads.max, with fs.s3a.max.total.tasks values setting the number of queued work items.	disk

Recuerda que S3 y HDFS funcionan de distintas maneras. Es necesario realizar un ajuste/prueba/experimento

cuidadoso para hacer el uso más eficiente de los recursos de S3.

Blogs de StorageGRID de NetApp

Puede encontrar algunos de los grandes blogs de StorageGRID de NetApp aquí:

- Mayo de 10: ["Lab on Demand es su mejor herramienta de venta para StorageGRID"](#)
- Mayo de 24: ["Modernice sus cargas de trabajo de análisis con NetApp y Alluxio"](#)
- Mayo de 26: ["StorageGRID: Almacenar y gestionar los datos de backup y replicación on-premises"](#)
- 9 de junio: ["Utilice el conector Hadoop S3A de Cloudera con StorageGRID"](#)
- 26 de julio: ["Consulte la creciente lista de soluciones de partners validadas para StorageGRID"](#)
- Agosto de 5: ["StorageGRID de NetApp obtiene una certificación de seguridad de criterios comunes"](#)
- Agosto de 16: ["Integración de StorageGRID con la pila ELK de código abierto para mejorar la experiencia del cliente"](#)
- Agosto de 17: ["Todo comienza con Object Locking... Creación de un ecosistema de almacenamiento S3 para aplicaciones de backup esenciales"](#)
- Agosto de 23: ["Cree su lago de datos en StorageGRID"](#)
- Septiembre de 1: ["Tome estas métricas y gráficas"](#)
- Septiembre de 19: ["Soporte de bloqueo de datos y protección frente a ransomware para StorageGRID"](#)
- Septiembre de 26: ["StorageGRID de NetApp para proveedores de servicios"](#)
- Octubre de 5: ["Descongele sus datos en StorageGRID para copo de nieve"](#)
- Octubre de 5: ["Cloud Insights de NetApp añade consolas a la galería de StorageGRID"](#)
- Noviembre de 7: ["Compatibilidad con StorageGRID y ONTAP S3: Diferencias, similitudes e integración"](#)
- Noviembre de 23: ["IA explicable con MLOps impulsado por NetApp y Modzy"](#)
- Dic 6: ["StorageGRID obtiene la certificación de cumplimiento de KPMG"](#)
- Ene 16: ["StorageGRID renueva la certificación de cumplimiento NF203 e ISO/IEC 25051"](#)
- Ene 18: ["StorageGRID S3 Object Lock validado para Veritas NetBackup"](#)
- 14 de febrero: ["¿Qué tienen en común el chocolate, el esquí, los relojes y los mainframes?"](#)
- Marzo de 14: ["Cómo realizar backups de las bases de datos de EHR de Epic Systems con un comando en una arquitectura compatible con 3:2:1"](#)
- Marzo de 30: ["Utilice BlueXP para proteger Epic EHR con una política de backup conforme a las normativas 3:2:1"](#)
- Marzo de 30: ["Punto de montaje para la versión alfa de Amazon S3 con StorageGRID"](#)
- Mayo de 16: ["Novedades de la familia de almacenamiento de objetos de StorageGRID"](#)
- Mayo de 16: ["Presentamos StorageGRID 11,7 y el nuevo dispositivo de almacenamiento de objetos all-flash SGF6112"](#)
- Agosto de 30: ["Mountpoint para Amazon S3 File System ahora es GA"](#)
- Septiembre de 1: ["Aprovechamiento de Cloud Insights para supervisar y recopilar registros mediante bits fluentes"](#)
- Octubre de 17: ["Pasando de Hadoop: Modernización del análisis de datos con Dremio y StorageGRID"](#)
- Noviembre de 7: ["Spectra Logic Glaciar en las instalaciones con StorageGRID"](#)
- Dic 12: ["Análisis de Big Data en StorageGRID: Dremio se ejecuta 23 veces más rápido que Apache Hive"](#)

- 2 de febrero: "Presentación del resumen de la solución StorageGRID + lakeFS"
- 16 de febrero: "Presentación de StorageGRID 11,8: Mayor seguridad, simplicidad y experiencia de usuario"
- 16 de febrero: "Presentamos 11,8 de StorageGRID"

Documentación de StorageGRID de NetApp

Puede encontrar la documentación completa de cada versión de StorageGRID de NetApp aquí:

- ["Dispositivos StorageGRID"](#)
- ["StorageGRID 11,8"](#)
- ["StorageGRID 11,7"](#)
- ["StorageGRID 11.6"](#)
- ["StorageGRID 11.5"](#)
- ["StorageGRID 11.4"](#)
- ["StorageGRID 11.3"](#)
- ["StorageGRID 11.2"](#)

Avisos legales

Los avisos legales proporcionan acceso a las declaraciones de copyright, marcas comerciales, patentes y mucho más.

Derechos de autor

["https://www.netapp.com/company/legal/copyright/"](https://www.netapp.com/company/legal/copyright/)

Marcas comerciales

NETAPP, el logotipo de NETAPP y las marcas enumeradas en la página de marcas comerciales de NetApp son marcas comerciales de NetApp, Inc. Los demás nombres de empresas y productos son marcas comerciales de sus respectivos propietarios.

["https://www.netapp.com/company/legal/trademarks/"](https://www.netapp.com/company/legal/trademarks/)

Estadounidenses

Puede encontrar una lista actual de las patentes propiedad de NetApp en:

<https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf>

Política de privacidad

["https://www.netapp.com/company/legal/privacy-policy/"](https://www.netapp.com/company/legal/privacy-policy/)

Código abierto

Los archivos de notificación proporcionan información sobre los derechos de autor y las licencias de terceros que se utilizan en software de NetApp.

https://library.netapp.com/ecm/ecm_download_file/2879263

https://library.netapp.com/ecm/ecm_download_file/2881511

Información de copyright

Copyright © 2024 NetApp, Inc. Todos los derechos reservados. Imprimido en EE. UU. No se puede reproducir este documento protegido por copyright ni parte del mismo de ninguna forma ni por ningún medio (gráfico, electrónico o mecánico, incluidas fotocopias, grabaciones o almacenamiento en un sistema de recuperación electrónico) sin la autorización previa y por escrito del propietario del copyright.

El software derivado del material de NetApp con copyright está sujeto a la siguiente licencia y exención de responsabilidad:

ESTE SOFTWARE LO PROPORCIONA NETAPP «TAL CUAL» Y SIN NINGUNA GARANTÍA EXPRESA O IMPLÍCITA, INCLUYENDO, SIN LIMITAR, LAS GARANTÍAS IMPLÍCITAS DE COMERCIALIZACIÓN O IDONEIDAD PARA UN FIN CONCRETO, CUYA RESPONSABILIDAD QUEDA EXIMIDA POR EL PRESENTE DOCUMENTO. EN NINGÚN CASO NETAPP SERÁ RESPONSABLE DE NINGÚN DAÑO DIRECTO, INDIRECTO, ESPECIAL, EJEMPLAR O RESULTANTE (INCLUYENDO, ENTRE OTROS, LA OBTENCIÓN DE BIENES O SERVICIOS SUSTITUTIVOS, PÉRDIDA DE USO, DE DATOS O DE BENEFICIOS, O INTERRUPTIÓN DE LA ACTIVIDAD EMPRESARIAL) CUALQUIERA SEA EL MODO EN EL QUE SE PRODUJERON Y LA TEORÍA DE RESPONSABILIDAD QUE SE APLIQUE, YA SEA EN CONTRATO, RESPONSABILIDAD OBJETIVA O AGRAVIO (INCLUIDA LA NEGLIGENCIA U OTRO TIPO), QUE SURJAN DE ALGÚN MODO DEL USO DE ESTE SOFTWARE, INCLUSO SI HUBIEREN SIDO ADVERTIDOS DE LA POSIBILIDAD DE TALES DAÑOS.

NetApp se reserva el derecho de modificar cualquiera de los productos aquí descritos en cualquier momento y sin aviso previo. NetApp no asume ningún tipo de responsabilidad que surja del uso de los productos aquí descritos, excepto aquello expresamente acordado por escrito por parte de NetApp. El uso o adquisición de este producto no lleva implícita ninguna licencia con derechos de patente, de marcas comerciales o cualquier otro derecho de propiedad intelectual de NetApp.

Es posible que el producto que se describe en este manual esté protegido por una o más patentes de EE. UU., patentes extranjeras o solicitudes pendientes.

LEYENDA DE DERECHOS LIMITADOS: el uso, la copia o la divulgación por parte del gobierno están sujetos a las restricciones establecidas en el subpárrafo (b)(3) de los derechos de datos técnicos y productos no comerciales de DFARS 252.227-7013 (FEB de 2014) y FAR 52.227-19 (DIC de 2007).

Los datos aquí contenidos pertenecen a un producto comercial o servicio comercial (como se define en FAR 2.101) y son propiedad de NetApp, Inc. Todos los datos técnicos y el software informático de NetApp que se proporcionan en este Acuerdo tienen una naturaleza comercial y se han desarrollado exclusivamente con fondos privados. El Gobierno de EE. UU. tiene una licencia limitada, irrevocable, no exclusiva, no transferible, no sublicenciable y de alcance mundial para utilizar los Datos en relación con el contrato del Gobierno de los Estados Unidos bajo el cual se proporcionaron los Datos. Excepto que aquí se disponga lo contrario, los Datos no se pueden utilizar, desvelar, reproducir, modificar, interpretar o mostrar sin la previa aprobación por escrito de NetApp, Inc. Los derechos de licencia del Gobierno de los Estados Unidos de América y su Departamento de Defensa se limitan a los derechos identificados en la cláusula 252.227-7015(b) de la sección DFARS (FEB de 2014).

Información de la marca comercial

NETAPP, el logotipo de NETAPP y las marcas que constan en <http://www.netapp.com/TM> son marcas comerciales de NetApp, Inc. El resto de nombres de empresa y de producto pueden ser marcas comerciales de sus respectivos propietarios.