StorageGRID in Ihrer Umgebung aktivieren

How to enable StorageGRID in your environment

NetApp
April 26, 2024

This PDF was generated from https://docs.netapp.com/de-de/storagegrid-enable/index.html on April 26, 2024. Always check docs.netapp.com for the latest.

Inhalt

StorageGRID in Ihrer Umgebung aktivieren

Validierte Lösungen von Drittanbietern

Validierte Lösungen von Drittanbietern: Überblick

In Zusammenarbeit mit unseren Partnern hat NetApp diese Lösungen zur Verwendung mit StorageGRID validiert. Lesen Sie die Informationen in diesem Abschnitt, um zu erfahren, welche Lösungen validiert wurden, und um weitere Anweisungen, falls zutreffend, zu erhalten.

Im Team mit NetApp erweitern Sie Ihr Portfolio schneller, erzielen einen höheren Bekanntheitsgrad und generieren Umsätze, indem Sie getestete, erstklassige NetApp Lösungen entwickeln. "Werden Sie noch heute Alliance-Partner".

StorageGRID 11.8 validierte Lösungen von Drittanbietern

Die folgenden Drittanbieterlösungen wurden für die Verwendung mit StorageGRID 11.8 validiert.
Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Alluxio
- Apache Kafka
- AWS Bereitstellungspunkt
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Collibra (Minimum Collibra Data Quality Version 2024.02)
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover-Daten
- Dremio
- Emam
- Fujifilm Object Archive
- GitHub Enterprise Server

- IBM FileNet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Microsoft SQL Server Big Data-Cluster
- Model9
- Modzy
- Moonwalk Universal
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- PixitMedia ngenea
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 Build 220706 oder höher
- Rubrik CDM
- s3a
- Signiant
- Schneeflocke
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Leicht Gemacht
- Trino
- Lack Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric
- WEKA v3.10 oder höher

Lösungen von Drittanbietern, die auf StorageGRID mit Objektsperre validiert wurden

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Commvault 11, Feature Release 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 und höher

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.7 validierte Lösungen von Drittanbietern

Die folgenden Drittanbieterlösungen wurden für die Verwendung mit StorageGRID 11.7 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Alluxio
- Apache Kafka
- AWS Bereitstellungspunkt
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Collibra (Minimum Collibra Data Quality Version 2024.02)
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover-Daten
- Dremio
- Emam
- Fujifilm Object Archive
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Microsoft SQL Server Big Data-Cluster
- Model9
- Modzy
- Moonwalk Universal
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura

- PixitMedia ngenea
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 Build 220706 oder höher
- Rubrik CDM
- s3a
- Signiant
- Schneeflocke
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Leicht Gemacht
- Trino
- Lack Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric
- WEKA v3.10 oder höher

Lösungen von Drittanbietern, die auf StorageGRID mit Objektsperre validiert wurden

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Commvault 11, Feature Release 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 und höher

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications

- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.6 validierte Drittanbieterlösungen

Die folgenden Drittanbieterlösungen wurden zur Verwendung mit StorageGRID 11.6 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Alluxio
- Apache Kafka
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Diskover-Daten

- Dremio
- Emam
- Fujifilm Object Archive
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect Plus
- Interica
- Komprise
- Microsoft SQL Server Big Data-Cluster
- Model9
- Modzy
- Moonwalk Universal
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- PixitMedia ngenea
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Reveille v10 Build 220706 oder höher
- Rubrik CDM
- s3a
- Signiant
- Schneeflocke
- Spectra Logic On-Prem Glacier
- Splunk Smartstore
- Storage Leicht Gemacht
- Trino
- Lack Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0

- Vertica 10.x
- Vidispine
- Virtalica StorageFabric
- WEKA v3.10 oder höher

Lösungen von Drittanbietern, die auf StorageGRID mit Objektsperre validiert wurden

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Commvault 11, Feature Release 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 und höher

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.5 validierte Drittanbieterlösungen

Die folgenden Drittanbieterlösungen wurden zur Verwendung mit StorageGRID 11.5 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Alluxio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- Moonwalk Universal
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM
- s3a
- Signiant
- Splunk Smartstore

- Trino
- Lack Enterprise 6.0.4
- Veeam 11
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric

Lösungen von Drittanbietern, die auf StorageGRID mit Objektsperre validiert wurden

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- OpenText Documentum 21.4
- Veeam 11

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Gitlab
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.4 validierte Drittanbieterlösungen

Die folgenden Drittanbieterlösungen wurden zur Verwendung mit StorageGRID 11.4 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM
- Signiant
- Splunk Smartstore
- Lack Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12

- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.3 validierte Drittanbieterlösungen

Die folgenden Drittanbieterlösungen wurden zur Verwendung mit StorageGRID 11.3 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6

- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- Splunk Smartstore
- Lack Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera

- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

StorageGRID 11.2 validierte Drittanbieterlösungen

Die folgenden Drittanbieterlösungen wurden zur Verwendung mit StorageGRID 11.2 validiert. + Wenn die von Ihnen gesuchte Lösung nicht aufgeführt ist, wenden Sie sich bitte an Ihren NetApp Ansprechpartner.

Drittanbieterlösungen, die auf StorageGRID validiert sind

Diese Lösungen wurden in Zusammenarbeit mit den jeweiligen Partnern getestet.

- Actifio
- Bridgestor
- Cantemo
- Citrix Content Collaboration
- Commvault 11
- Ctera Portal 6
- Dalet
- Datadobi
- Data Dynamics StorageX
- DefendX
- Interica
- Komprise
- SCHÖN
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 mit CyanGate Cloud
- Panzura
- Gateway für die Archivierung 2.0
- Punkt Storage Manager 6.4
- Primestream
- Quantum StorNext 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342

- Signiant
- Splunk Smartstore
- Lack Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

Von StorageGRID unterstützte Lösungen von Drittanbietern

Diese Lösungen wurden getestet.

- Archiware
- Axis Communications
- Kongruation360
- DataFrameworks
- EcoDigital DIVA-Plattform
- Encoding.com
- Fujifilm Object Archive
- GE Centricity Enterprise Archive
- Hyland Acuo
- IBM Aspera
- Milestone Systems
- OnSSI
- Schubmotor
- SilverTrak
- SoftNAS
- QStar
- Velasea

Produktfunktionshandbücher

Cloud Storage Pool für AWS oder Google Cloud erstellen

Sie können einen Cloud Storage Pool verwenden, wenn Sie StorageGRID-Objekte in einen externen S3-Bucket verschieben möchten. Der externe Bucket kann zu Amazon S3 (AWS) oder Google Cloud gehören.

Was Sie benötigen

- StorageGRID 11.6 wurde konfiguriert.
- Sie haben bereits einen externen S3-Bucket auf AWS oder Google Cloud eingerichtet.

Schritte

1. Navigieren Sie im Grid Manager zu **ILM** > **Storage Pools**.
2. Wählen Sie im Abschnitt Cloud-Speicherpools der Seite **Erstellen** aus.

   Das Popup-Fenster „Cloud-Speicherpool erstellen“ wird angezeigt.

3. Geben Sie einen Anzeigenamen ein.
4. Wählen Sie in der Dropdown-Liste Provider Type * Amazon S3* aus.

   Dieser Provider-Typ funktioniert für AWS S3 oder Google Cloud.

5. Geben Sie den URI für den S3-Bucket ein, der für den Cloud-Storage-Pool verwendet werden soll.

   Es sind zwei Formate zulässig:

   ```
   https://host:port
   ```

   ```
   http://host:port
   ```

6. Geben Sie den S3-Bucket-Namen ein.

   Der angegebene Name muss exakt mit dem Namen des S3-Buckets übereinstimmen. Andernfalls schlägt die Erstellung von Cloud-Storage-Pool fehl. Sie können diesen Wert nicht ändern, nachdem der Cloud-Speicherpool gespeichert wurde.

7. Geben Sie optional die Zugriffsschlüssel-ID und den geheimen Zugriffsschlüssel ein.
8. Wählen Sie in der Dropdown-Liste * Zertifikat nicht überprüfen* aus.
9. Klicken Sie Auf **Speichern**.

Erwartetes Ergebnis

Es muss sichergestellt werden, dass ein Cloud-Storage-Pool für Amazon S3 oder Google Cloud erstellt wurde.

Von Jonathan Wong

Cloud Storage Pool für Azure Blob Storage erstellen

Sie können einen Cloud Storage Pool verwenden, wenn Sie StorageGRID-Objekte in

einen externen Azure Container verschieben möchten.

Was Sie benötigen

- StorageGRID 11.6 wurde konfiguriert.
- Sie haben bereits einen externen Azure-Container eingerichtet.

Schritte

1. Navigieren Sie im Grid Manager zu **ILM** > **Storage Pools**.
2. Wählen Sie im Abschnitt Cloud-Speicherpools der Seite **Erstellen** aus.

   Das Popup-Fenster „Cloud-Speicherpool erstellen“ wird angezeigt.

3. Geben Sie einen Anzeigenamen ein.
4. Wählen Sie in der Dropdown-Liste Provider Type * Azure Blob Storage* aus.
5. Geben Sie den URI für den S3-Bucket ein, der für den Cloud-Storage-Pool verwendet werden soll.

   Es sind zwei Formate zulässig:

   ```
   https://host:port
   ```

   ```
   http://host:port
   ```

6. Geben Sie den Azure-Containernamen ein.

   Der angegebene Name muss exakt mit dem Azure-Containernamen übereinstimmen. Andernfalls schlägt die Erstellung des Cloud-Storage-Pools fehl. Sie können diesen Wert nicht ändern, nachdem der Cloud-Speicherpool gespeichert wurde.

7. Geben Sie optional den zugeordneten Kontonamen und den Kontokschlüssel des Azure-Containers für die Authentifizierung ein.
8. Wählen Sie in der Dropdown-Liste * Zertifikat nicht überprüfen* aus.
9. Klicken Sie Auf **Speichern**.

Erwartetes Ergebnis

Erstellen eines Cloud-Storage-Pools für Azure Blob Storage bestätigen

Von Jonathan Wong

Verwenden Sie einen Cloud Storage Pool für Backups

Sie können eine ILM-Regel erstellen, um Objekte für Backups in einen Cloud Storage-Pool zu verschieben.

Was Sie benötigen

- StorageGRID 11.6 wurde konfiguriert.
- Sie haben bereits einen externen Azure-Container eingerichtet.

Schritte

1. Navigieren Sie im Grid Manager zu **ILM** > **Regeln** > **Erstellen**.

2. Geben Sie eine Beschreibung ein.
3. Geben Sie ein Kriterium ein, um die Regel auszulösen.
4. Klicken Sie Auf **Weiter**.
5. Replizieren Sie das Objekt auf Storage Nodes.
6. Fügen Sie eine Platzierungsregel hinzu.
7. Replizieren des Objekts in den Cloud Storage Pool
8. Klicken Sie Auf **Weiter**.
9. Klicken Sie Auf **Speichern**.

Erwartetes Ergebnis

Vergewissern Sie sich, dass im Aufbewahrungsdiagramm die lokal in StorageGRID gespeicherten Objekte und in einem Cloud-Speicherpool für Backups angezeigt werden.

Vergewissern Sie sich, dass bei Auslösung der ILM-Regel im Cloud Storage Pool eine Kopie vorhanden ist und Sie das Objekt lokal abrufen können, ohne ein Objekt wiederherstellen zu müssen.

Von Jonathan Wong

Konfigurieren Sie den Integrationsservice für die StorageGRID Suche

Dieses Handbuch enthält detaillierte Anweisungen zur Konfiguration des NetApp StorageGRID 11.6 Suchintegrationsservice mit Amazon OpenSearch Service oder On-Premises-Elasticsearch.

Einführung

StorageGRID unterstützt drei Arten von Plattform-Services.

- **StorageGRID CloudMirror Replikation**. Spiegeln bestimmter Objekte aus einem StorageGRID-Bucket auf ein angegebenes externes Ziel
- **Benachrichtigungen**. Bucket-spezifische Ereignisbenachrichtigungen senden Benachrichtigungen über bestimmte Aktionen, die an Objekten durchgeführt werden, an einen bestimmten externen Amazon Simple Notification Service (Amazon SNS).
- **Integrationsservice suchen**. Senden von einfachen Storage Service (S3) Objektmetadaten in einen angegebenen Elasticsearch-Index, wo Sie die Metadaten mithilfe des externen Service durchsuchen oder analysieren können.

Plattform-Services werden vom S3-Mandanten über die Mandanten-Manager-UI konfiguriert. Weitere Informationen finden Sie unter "Überlegungen bei der Verwendung von Plattform-Services".

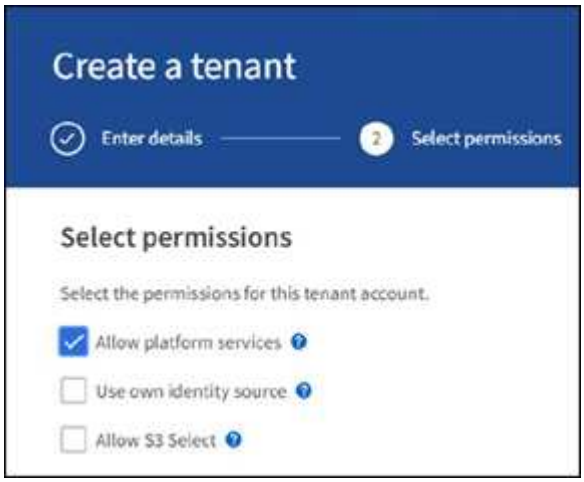
Dieses Dokument dient als Ergänzung zum "StorageGRID 11.6 Mandantenleitfaden" Und enthält Schritt-für-Schritt-Anleitungen und Beispiele für die Endpunkt- und Bucket-Konfiguration für Suchintegrations-Services. Die hier enthaltene Anleitung zur Einrichtung von Amazon Web Services (AWS) oder lokalen Elasticsearch-Services dienen nur zu Test- oder Demonstrationszwecken.

Zielgruppen sollten mit Grid Manager, Mandanten-Manager vertraut sein und über den S3-Browser Zugang verfügen, um grundlegende Vorgänge zum Hochladen (PUT) und Herunterladen (GET) für StorageGRID-

Suchintegrationstests durchzuführen.

Erstellung von Mandanten und Aktivierung von Plattform-Services

1. Erstellen Sie einen S3-Mandanten mithilfe von Grid Manager, geben Sie einen Anzeigenamen ein und wählen Sie das S3-Protokoll aus.
2. Wählen Sie auf der Berechtigungsseite die Option Plattformdienste zulassen. Wählen Sie ggf. andere Berechtigungen aus.



3. Richten Sie das ursprüngliche Kennwort des Mandanten-Root-Benutzers ein, oder wählen Sie, falls im Raster der Identifikationsverbund aktiviert ist, welche föderierte Gruppe Root-Zugriffsberechtigungen hat, um das Mandantenkonto zu konfigurieren.
4. Klicken Sie auf als Stamm anmelden und wählen Sie Bucket: Erstellen und verwalten.

   Dies führt Sie zur Seite Tenant Manager.

5. Wählen Sie im Tenant Manager My Access Keys aus, um den S3-Zugriffsschlüssel für spätere Tests zu erstellen und herunterzuladen.

Integrationsservices mit Amazon OpenSearch suchen

Einrichtung des Amazon OpenSearch Service (ehemals Elasticsearch)

Verwenden Sie dieses Verfahren für eine schnelle und einfache Einrichtung des OpenSearch-Dienstes nur zu Test-/Demo-Zwecken. Wenn Sie On-Premises-Elasticsearch für Suchintegrationsservices verwenden, lesen Sie den Abschnitt Suchintegrations-Services für On-Premises-Elasticsearch.

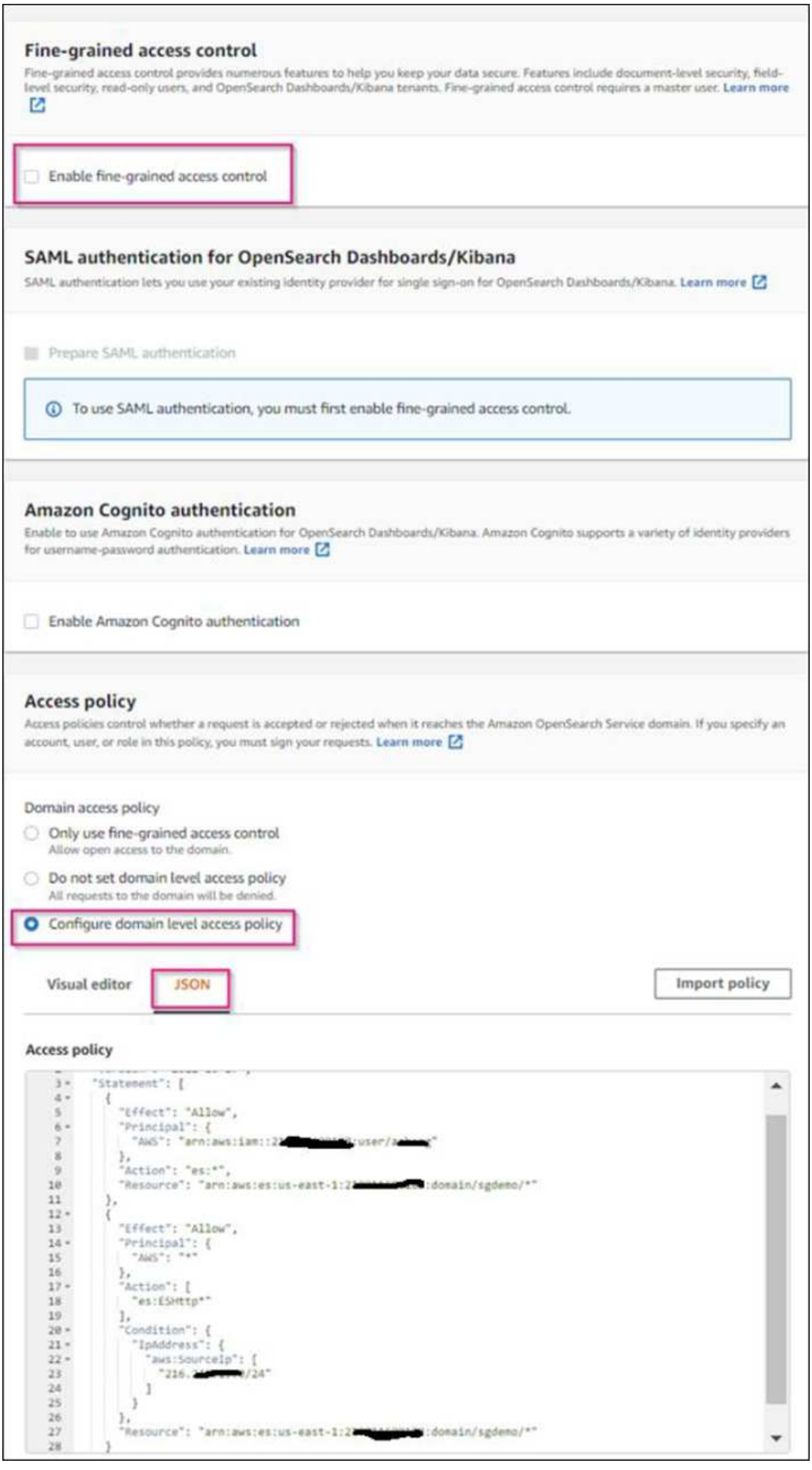
> Sie müssen über eine gültige Anmeldung für die AWS-Konsole, einen Zugriffsschlüssel, einen geheimen Zugriffsschlüssel und die Berechtigung zum Abonnieren des OpenSearch-Dienstes verfügen.

1. Erstellen Sie mithilfe der Anweisungen von eine neue Domäne "AWS OpenSearch Service – erste Schritte", Mit Ausnahme der folgenden:
   - Schritt 4: Domain-Name: Sgdemo
   - Schritt 10: Feinkörnige Zugriffssteuerung: Deaktivieren Sie die Option Enable Fine-genarainable Access Control.

◦ Schritt 12: Zugriffsrichtlinie: Wählen Sie Zugriffsrichtlinie auf Zugriffsebene konfigurieren, wählen Sie die Registerkarte JSON aus, um die Zugriffsrichtlinie anhand des folgenden Beispiels zu ändern:

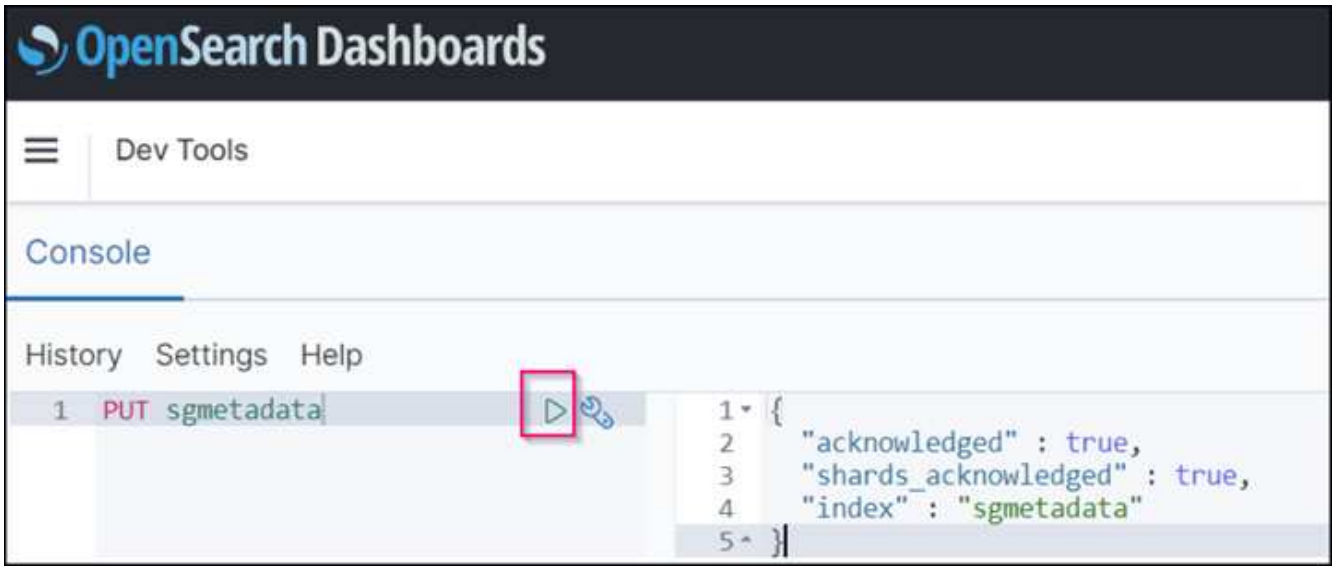
- Ersetzen Sie den hervorgehobenen Text durch Ihre eigene AWS IAM-ID (Identity and Access Management) und Ihren Benutzernamen.
- Ersetzen Sie den markierten Text (die IP-Adresse) durch die öffentliche IP-Adresse Ihres lokalen Computers, über den Sie auf die AWS-Konsole zugreifen.
- Öffnen Sie eine Browserregisterkarte für "https://checkip.amazonaws.com" Um Ihre öffentliche IP zu finden.

```
{

    "Version": "2012-10-17",
    "Statement": [
        {
        "Effect": "Allow",
        "Principal":
        {"AWS": "arn:aws:iam:: nnnnnn:user/xyzabc"},
        "Action": "es:*",
        "Resource": "arn:aws:es:us-east-1:nnnnnn: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:nnnnnn:domain/sgdemo/*"
        }
    ]
}
```

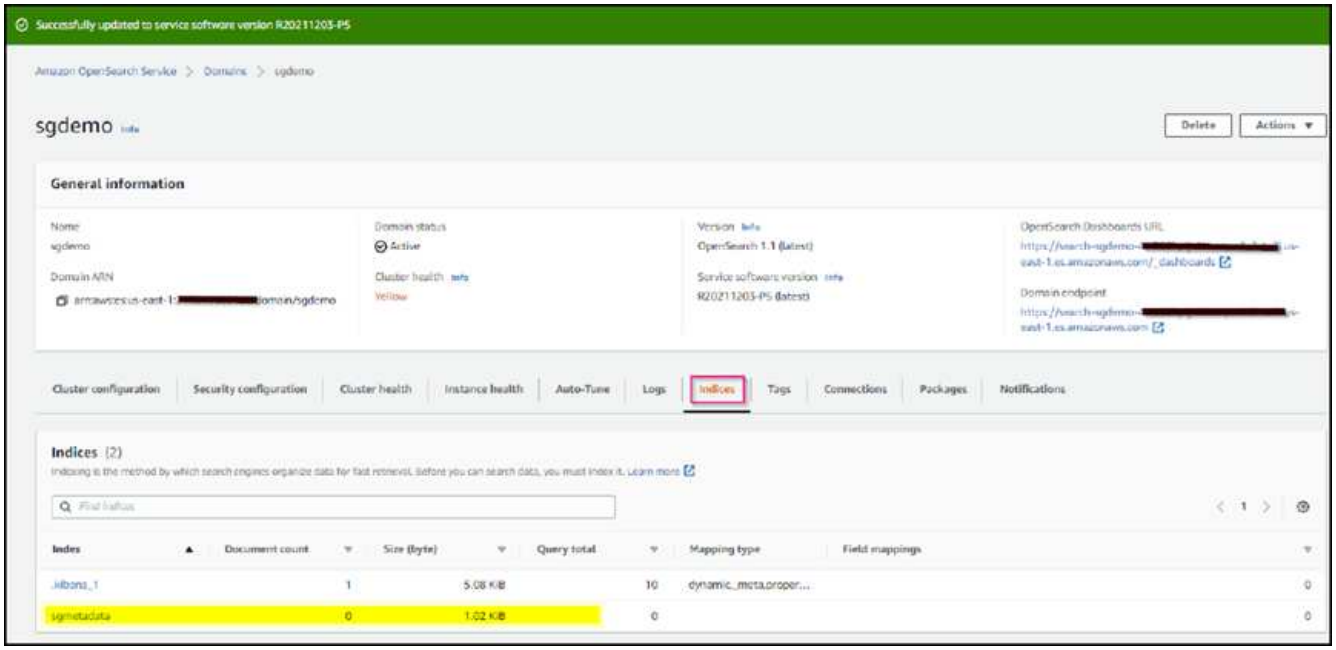

Fine-grained access control
Fine-grained access control provides numerous features to help you keep your data secure. Features include document-level security, field-level security, read-only users, and OpenSearch Dashboards/Kibana tenants. Fine-grained access control requires a master user. Learn more
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
"Statement": [
{
"Effect": "Allow",
"Principal": {
"AWS": "arn:aws:iam::2[redacted]:user/A[redacted]"
},
"Action": "es:*",
"Resource": "arn:aws:es:us-east-1:2[redacted]:domain/sgdemo/*"
},
{
"Effect": "Allow",
"Principal": {
"AWS": "*"
},
"Action": [
"es:ESHttp*"
],
"Condition": {
"IpAddress": {
"aws:SourceIp": [
"216.[redacted]/24"
]
}
},
"Resource": "arn:aws:es:us-east-1:2[redacted]:domain/sgdemo/*"
}

2. Warten Sie 15 bis 20 Minuten, bis die Domain aktiv ist.

3. Klicken Sie auf OpenSearch Dashboards URL, um die Domäne in einer neuen Registerkarte zu öffnen, um auf das Dashboard zuzugreifen. Wenn ein Fehler beim Zugriff verweigert wird, überprüfen Sie, ob die Quell-IP-Adresse der Zugriffsrichtlinie korrekt auf Ihre öffentliche IP-Adresse des Computers eingestellt ist, um den Zugriff auf das Domain-Dashboard zu ermöglichen.
4. Wählen Sie auf der Willkommensseite des Dashboards „Explore“ auf eigene Faust aus. Wählen Sie im Menü „Management → Entwicklungstools“
5. Geben Sie unter Dev Tools → Console ein `PUT <index>` Wo Sie den Index zum Speichern von StorageGRID-Objektmetadaten verwenden. Im folgenden Beispiel verwenden wir den Indexnamen 'sgmetadaten'. Klicken Sie auf das kleine Dreieck-Symbol, um den PUT-Befehl auszuführen. Das erwartete Ergebnis wird im rechten Bereich angezeigt, wie im folgenden Beispiel Screenshot dargestellt.



6. Überprüfen Sie, ob der Index über die Benutzeroberfläche von Amazon OpenSearch unter sgdomain > Indizes sichtbar ist.

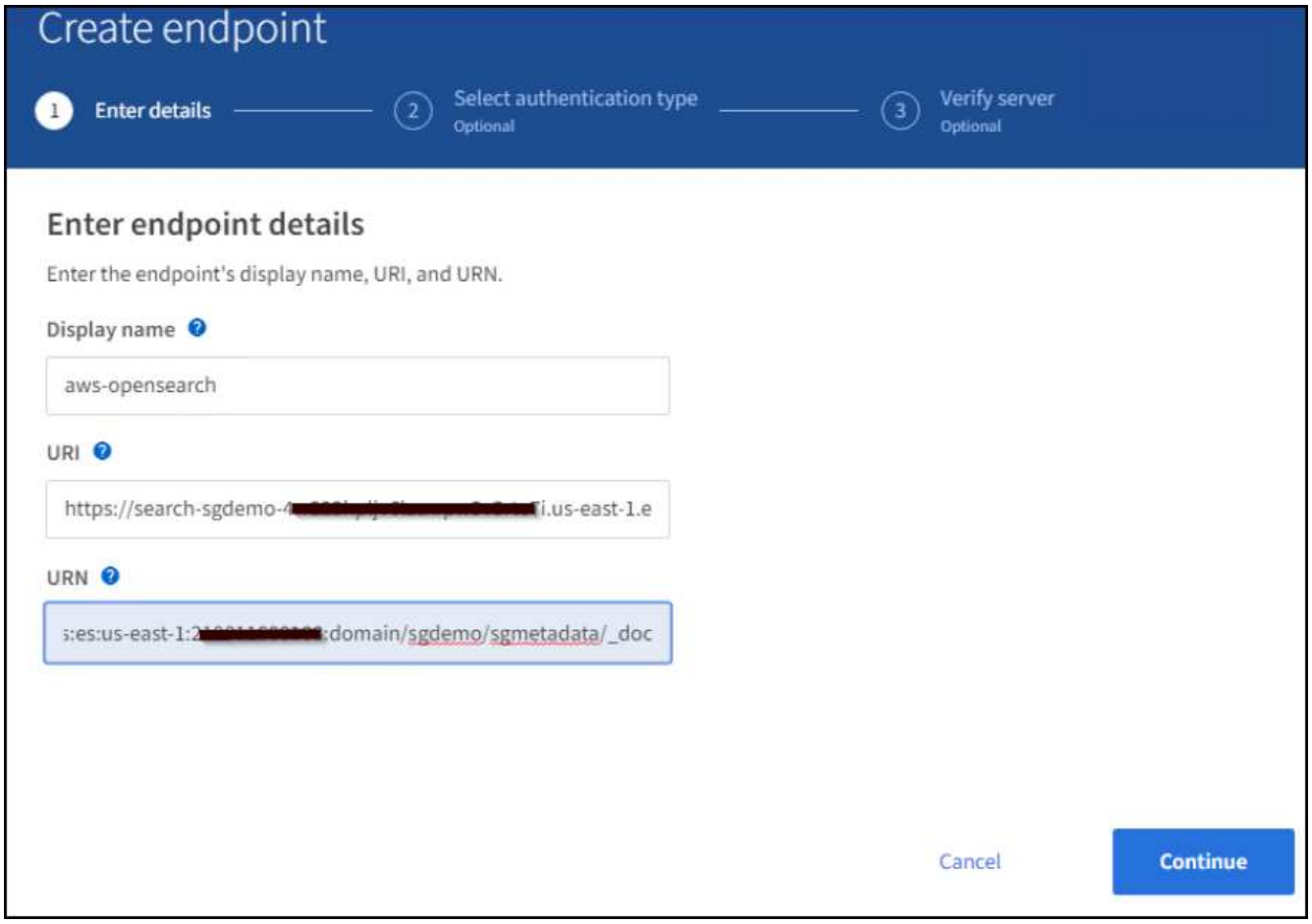


Endpoint-Konfiguration für Plattform-Services

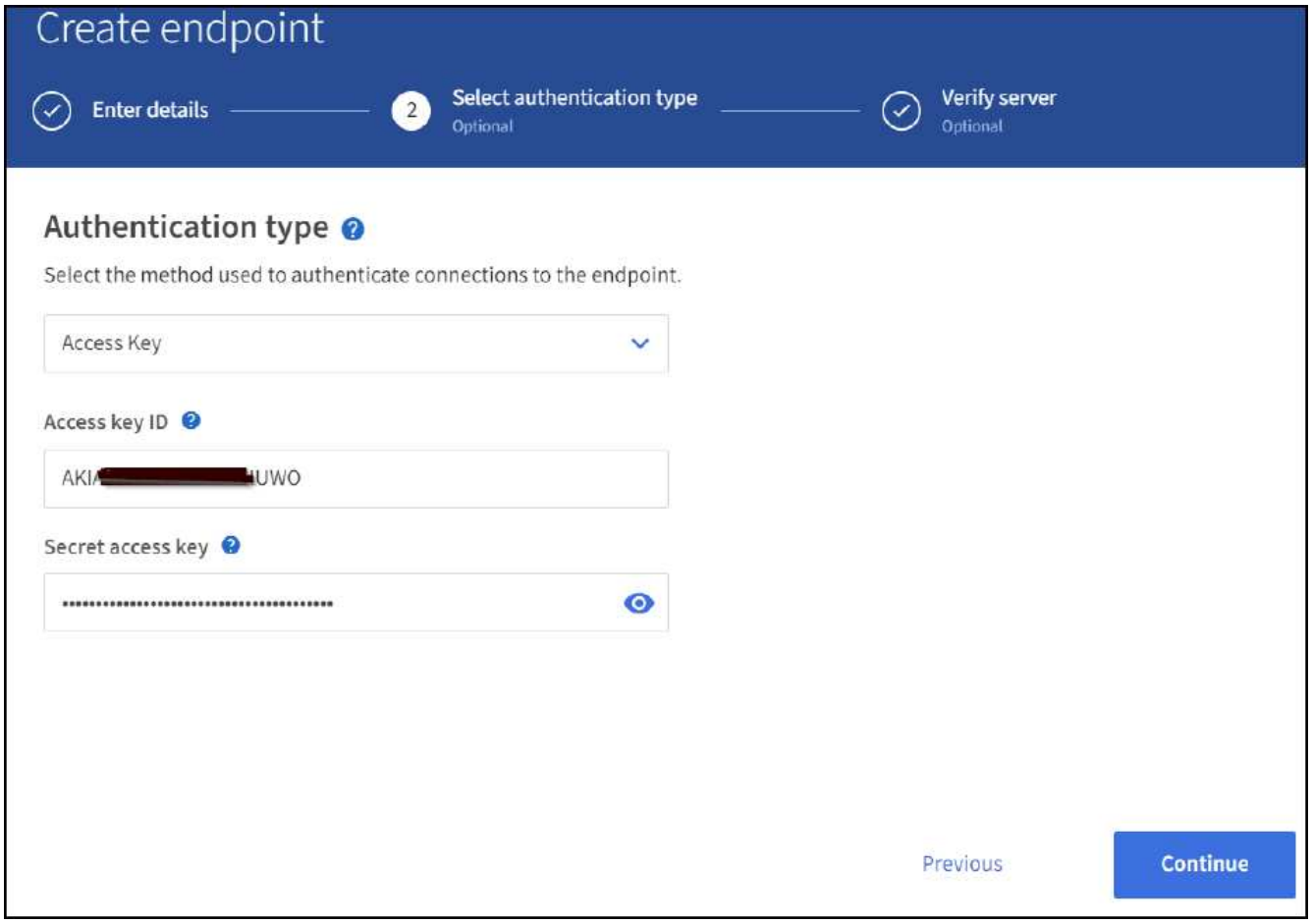
Gehen Sie wie folgt vor, um die Endpunkte der Plattformservices zu konfigurieren:

1. In Tenant Manager wechseln Sie zu STORAGE(S3) > Plattform-Services-Endpunkten.
2. Klicken Sie auf Endpunkt erstellen, geben Sie Folgendes ein und klicken Sie dann auf Weiter:
   - Beispiel für einen Anzeigenamen `aws-opensearch`
   - Der Domänenendpunkt im Beispiel-Screenshot unter Schritt 2 des vorhergehenden Verfahrens im URI-Feld.
   - Die Domäne ARN, die in Schritt 2 des vorhergehenden Verfahrens im Feld URN verwendet wurde und addieren `/<index>/_doc` Bis zum Ende von ARN.

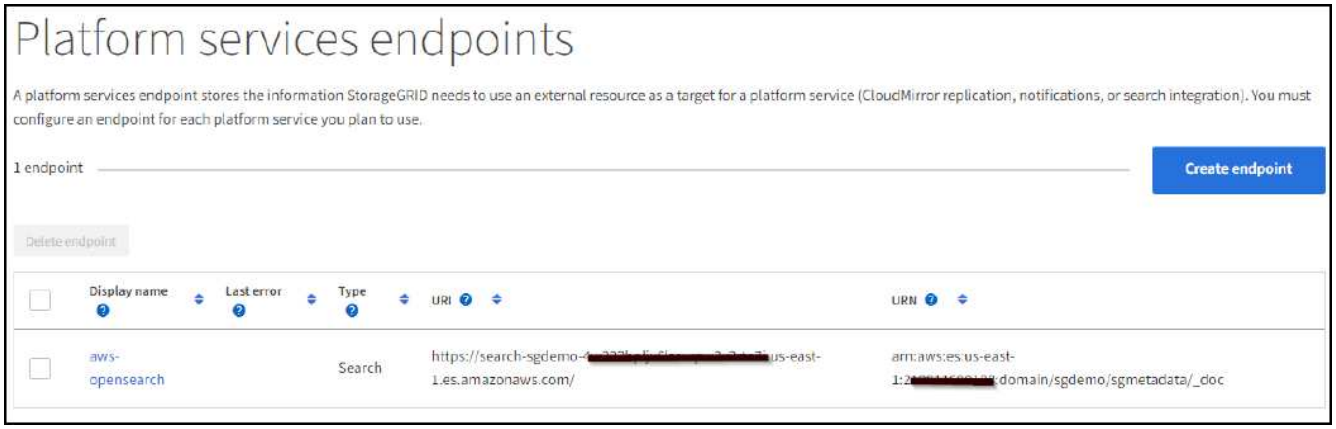
     In diesem Beispiel wird URN `arn:aws:es:us-east-1:211234567890:domain/sgdemo /sgmedata/_doc`.



3. Um auf die Amazon OpenSearch sgdomain zuzugreifen, wählen Sie als Authentifizierungstyp den Zugriffsschlüssel aus, und geben Sie dann den Amazon S3-Zugriffsschlüssel und den geheimen Schlüssel ein. Um zur nächsten Seite zu gelangen, klicken Sie auf Weiter.



4. Um den Endpunkt zu überprüfen, wählen Sie Operating System CA Certificate und Test and Create Endpoint aus. Wenn die Überprüfung erfolgreich ist, wird ein Endpunkt-Bildschirm angezeigt, der der folgenden Abbildung entspricht. Wenn die Überprüfung fehlschlägt, überprüfen Sie, ob der URN umfasst `/<index>/_doc` Am Ende des Pfads und der AWS Zugriffsschlüssel und der Geheimschlüssel sind korrekt.



Suchintegrations-Services für On-Premises-Elasticsearch

Elasticsearch-Einrichtung vor Ort

Dieses Verfahren dient der schnellen Einrichtung von vor-Ort-Elasticsearch und Kibana mit Docker nur zu Testzwecken. Wenn Elasticsearch und Kibana-Server bereits vorhanden sind, fahren Sie mit Schritt 5 fort.

1. Folgen Sie diesen Anweisungen "Docker-Installationsvorgang" So installieren Sie den Docker. Wir verwenden den "CentOS Docker Installationsverfahren" In diesem Setup.

```
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
```

   - Um den Docker nach dem Neustart zu starten, geben Sie Folgendes ein:

```
sudo systemctl enable docker
```

   - Stellen Sie die ein `vm.max_map_count` Wert für 262144:

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

   - Um die Einstellung nach dem Neustart zu behalten, geben Sie Folgendes ein:

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

2. Folgen Sie den "Elasticsearch Quick Start Guide" Selbstverwalteter Abschnitt zum Installieren und Ausführen der Elasticsearch- und Kibana-Docker. In diesem Beispiel wurde die Version 8.1 installiert.

Beachten Sie den Benutzernamen/das Kennwort und das Token, das Elasticsearch erstellt hat. Sie müssen diese zum Starten der Kibana UI und der StorageGRID-Plattform-Endpunktauthentifizierung verwenden.

Elasticsearch Service | Self-managed

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.

NOTE 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.

NOTE 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.

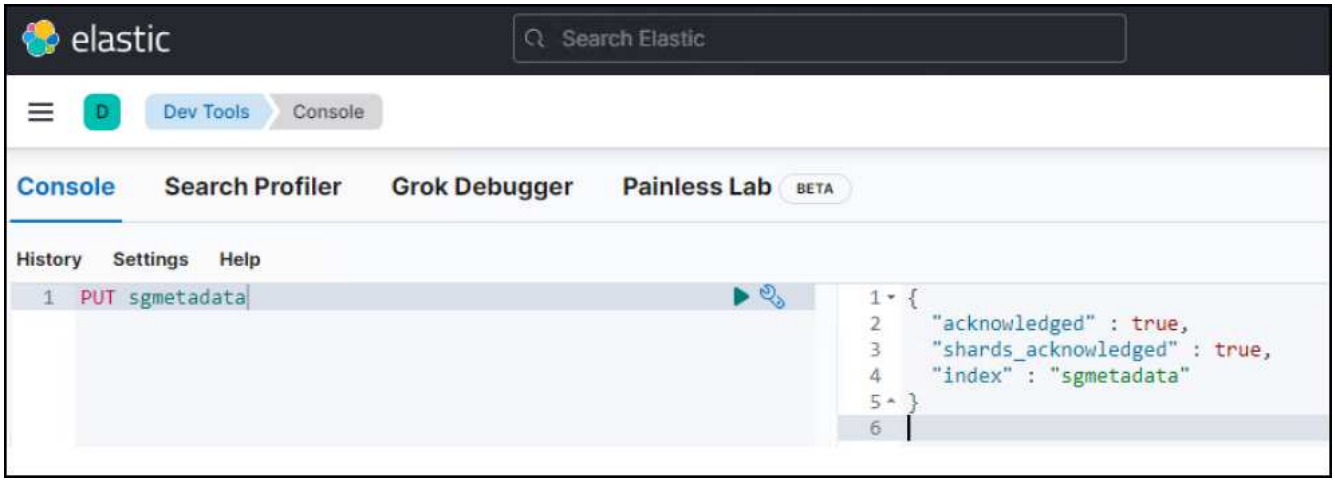
1. In a new terminal session, run:

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

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.

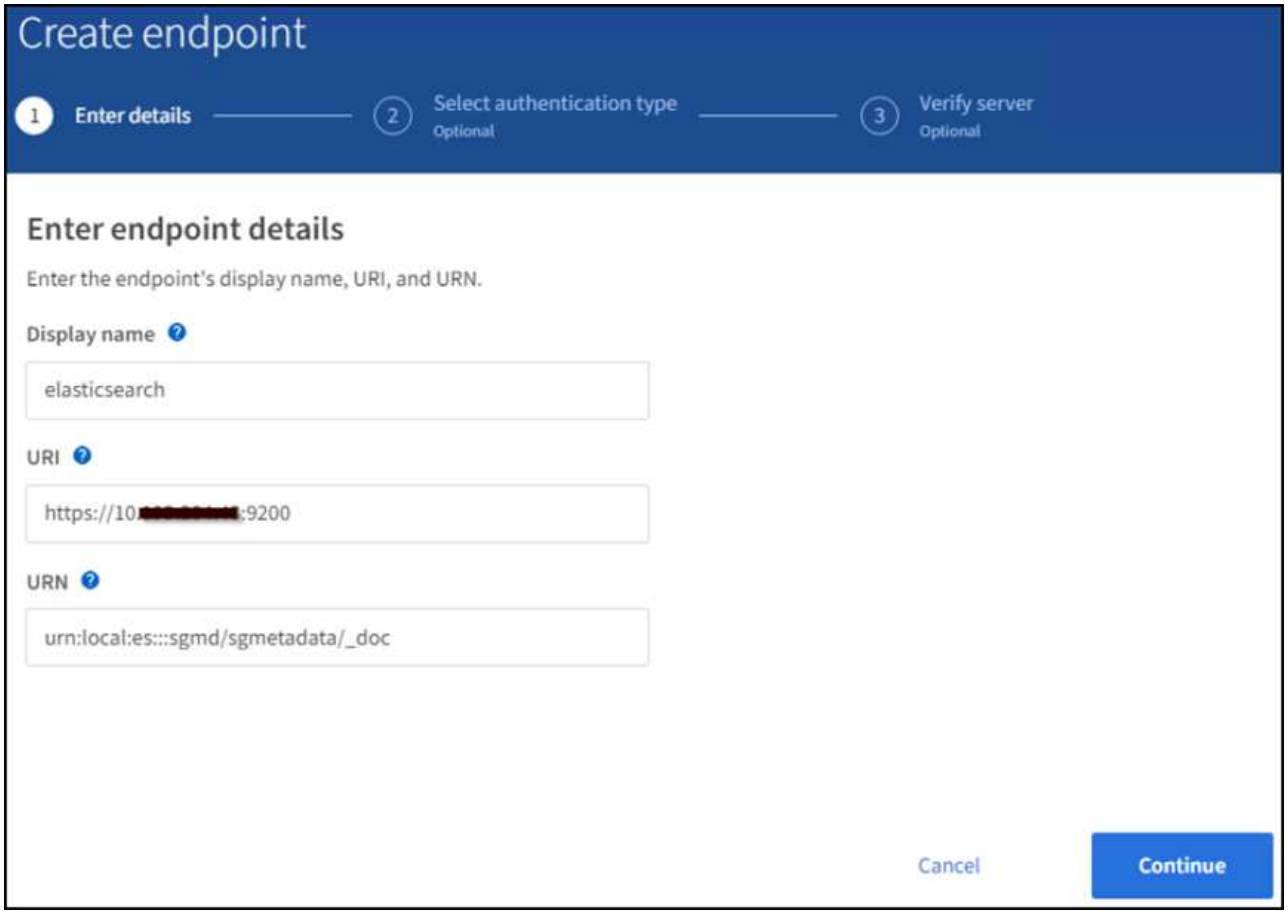
3. Nachdem der Kibana-Docker-Container gestartet wurde, wird der URL-Link aufgerufen `https://0.0.0.0:5601` Wird in der Konsole angezeigt. Ersetzen Sie 0.0.0.0 durch die Server-IP-Adresse in der URL.
4. Melden Sie sich mit dem Benutzernamen bei der Kibana-Benutzeroberfläche an `elastic` Und das Passwort, das im vorherigen Schritt von Elastic generiert wurde.
5. Wenn Sie sich zum ersten Mal anmelden möchten, wählen Sie auf der Begrüßungsseite „Explore“. Wählen Sie im Menü Verwaltung > Entwicklungstools.
6. Geben Sie auf dem Bildschirm Dev Tools Console die Eingabe ein `PUT <index>` Dort, wo Sie diesen Index zum Speichern von StorageGRID-Objektmetadaten verwenden. Wir verwenden den Indexnamen `sgmetadata` In diesem Beispiel. Klicken Sie auf das kleine Dreieck-Symbol, um den PUT-Befehl auszuführen. Das erwartete Ergebnis wird im rechten Bereich angezeigt, wie im folgenden Beispiel Screenshot dargestellt.



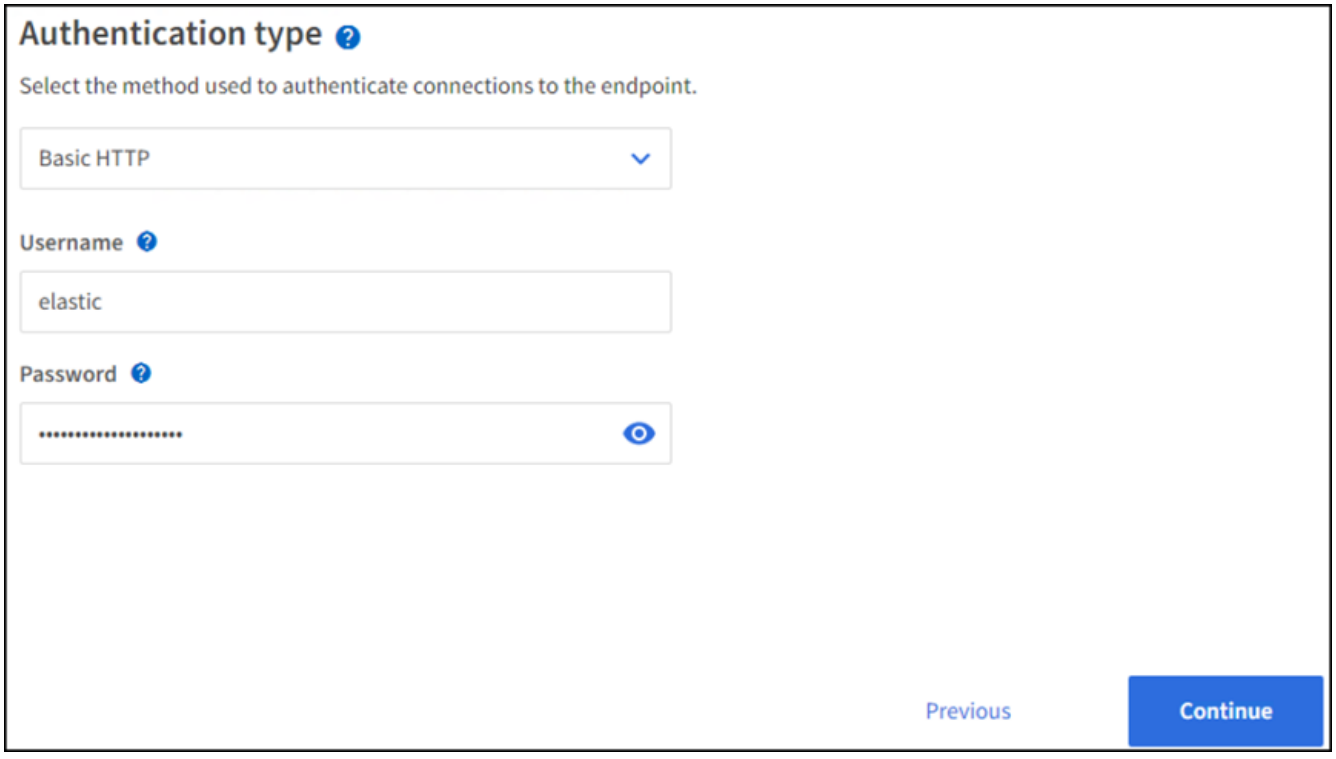
Endpoint-Konfiguration für Plattform-Services

Gehen Sie wie folgt vor, um Endpunkte für Plattformservices zu konfigurieren:

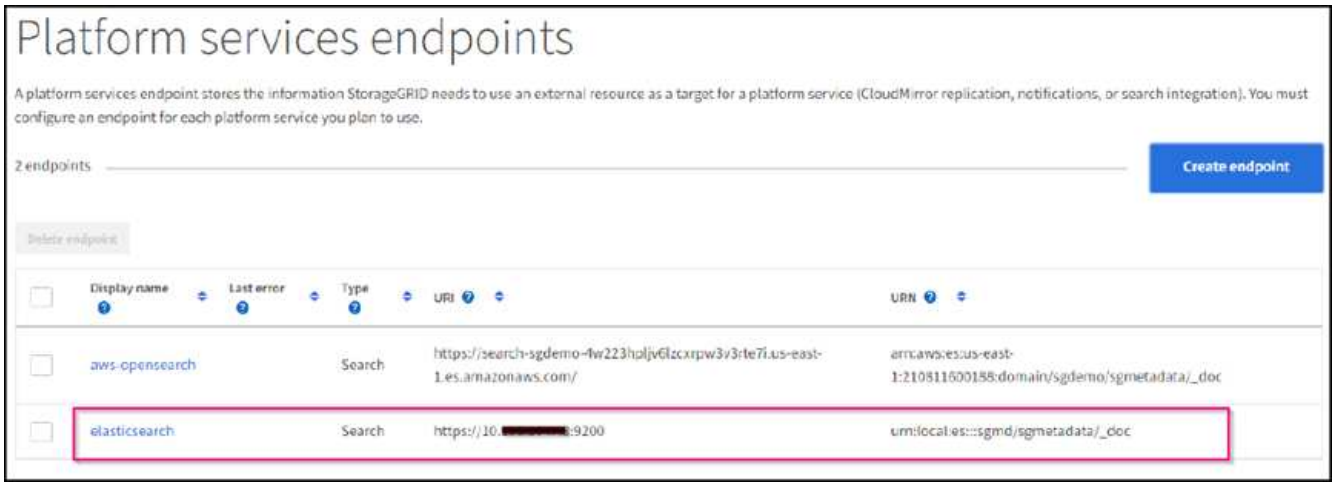
1. In Tenant Manager wechseln Sie zu STORAGE (S3) > Plattform-Services-Endpunkten
2. Klicken Sie auf Endpunkt erstellen, geben Sie Folgendes ein und klicken Sie dann auf Weiter:
   - Beispiel für Anzeigename: `elasticsearch`
   - URI: `https://<elasticsearch-server-ip or hostname>:9200`
   - URNE: `urn:<something>:es:::<some-unique-text>/<index-name>/_doc` Wobei der Indexname der Name ist, den Sie auf der Kibana-Konsole verwendet haben. Beispiel: `urn:local:es:::sgmd/sgmetadata/_doc`



3. Wählen Sie Basic HTTP als Authentifizierungstyp, geben Sie den Benutzernamen ein `elastic` Und das durch den Elasticsearch-Installationsprozess generierte Passwort. Um zur nächsten Seite zu gelangen, klicken Sie auf Weiter.



4. Wählen Sie Zertifikat nicht überprüfen und Endpunkt erstellen und testen, um den Endpunkt zu überprüfen. Wenn die Überprüfung erfolgreich ist, wird ein Endpunkt-Bildschirm angezeigt, der dem folgenden Screenshot ähnelt. Wenn die Überprüfung fehlschlägt, überprüfen Sie, ob die Einträge für URN, URI und Benutzername/Passwort korrekt sind.



Konfiguration des integrierten Service für die Bucket-Suche

Nachdem der Plattform-Service-Endpunkt erstellt wurde, besteht der nächste Schritt darin, diesen Service auf Bucket-Ebene zu konfigurieren, um Objektmetadaten an den definierten Endpunkt zu senden, sobald ein Objekt erstellt, gelöscht oder seine Metadaten oder Tags aktualisiert werden.

Sie können die Suchintegration mit Tenant Manager konfigurieren, um eine benutzerdefinierte StorageGRID-Konfigurations-XML auf einen Bucket anzuwenden wie folgt:

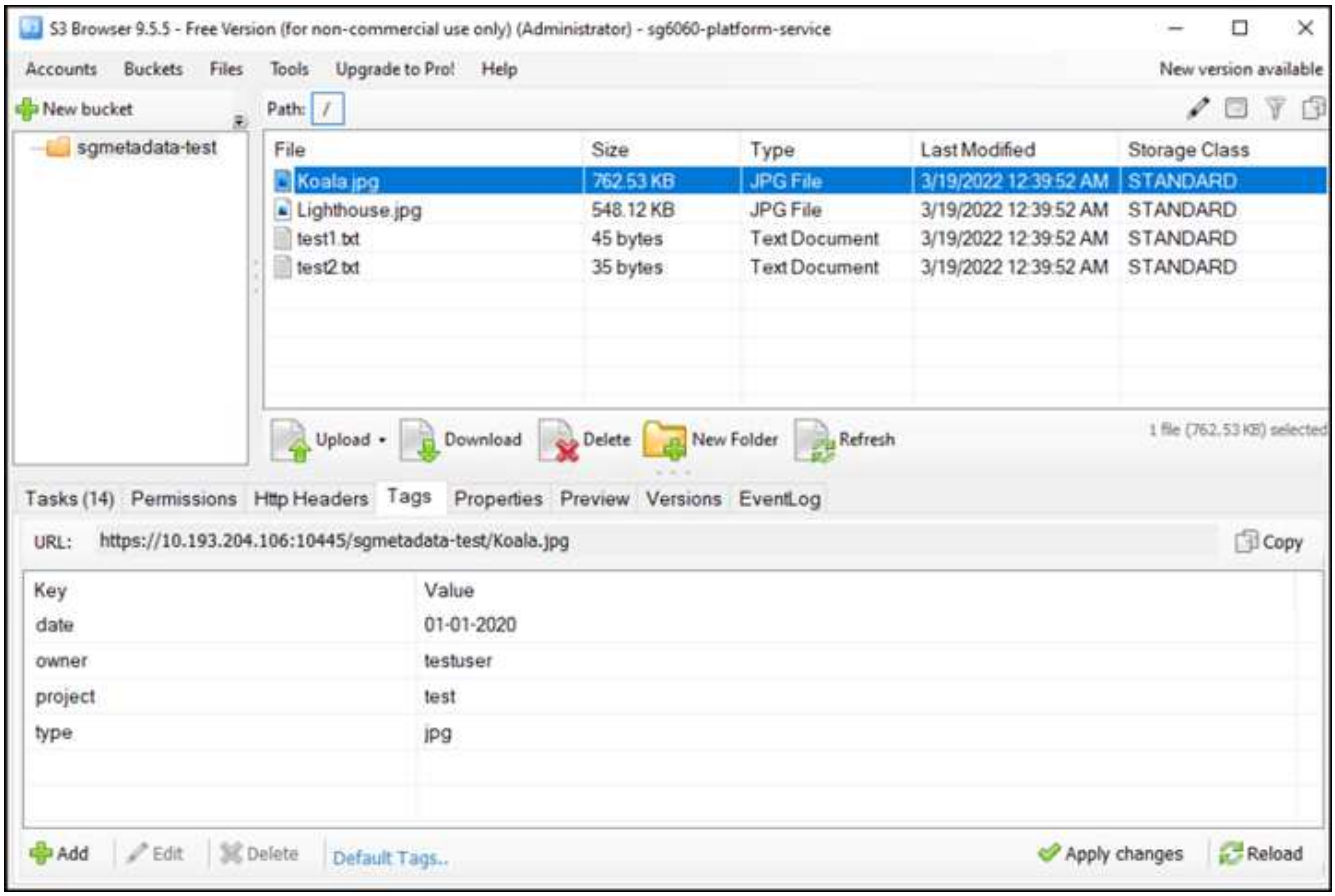
1. Wählen Sie in Tenant Manager „STORAGE(S3)“ > „Buckets“
2. Klicken Sie auf Create Bucket. Geben Sie den Bucket-Namen ein (z. B. `sgmetadata-test`) Und akzeptieren Sie die Standardeinstellung `us-east-1` Werden.
3. Klicken Sie Auf Weiter > Bucket Erstellen.
4. Um die Seite „Bucket-Übersicht“ aufzurufen, klicken Sie auf den Bucket-Namen und wählen Sie „Platform Services“ aus.
5. Wählen Sie das Dialogfeld Integration der Suche aktivieren aus. Geben Sie im angegebenen XML-Feld die Konfigurations-XML-XML-Datei unter Verwendung dieser Syntax ein.

   Der hervorgehobene URN muss mit dem von Ihnen definierten Endpunkt für Plattformservices übereinstimmen. Sie können eine weitere Browserregisterkarte öffnen, um auf den Mandantenmanager zuzugreifen und URN vom definierten Endpunkt der Plattformdienste zu kopieren.

   In diesem Beispiel haben wir kein Präfix verwendet, was bedeutet, dass die Metadaten für jedes Objekt in diesem Bucket an den zuvor definierten Elasticsearch-Endpunkt gesendet werden.

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

6. Verwenden Sie S3-Browser, um eine Verbindung zu StorageGRID mit dem Mandantenzugriff/geheimen Schlüssel herzustellen und Testobjekte in hochzuladen `sgmetadata-test` Bucket und fügen Sie Tags oder benutzerdefinierte Metadaten zu Objekten hinzu.



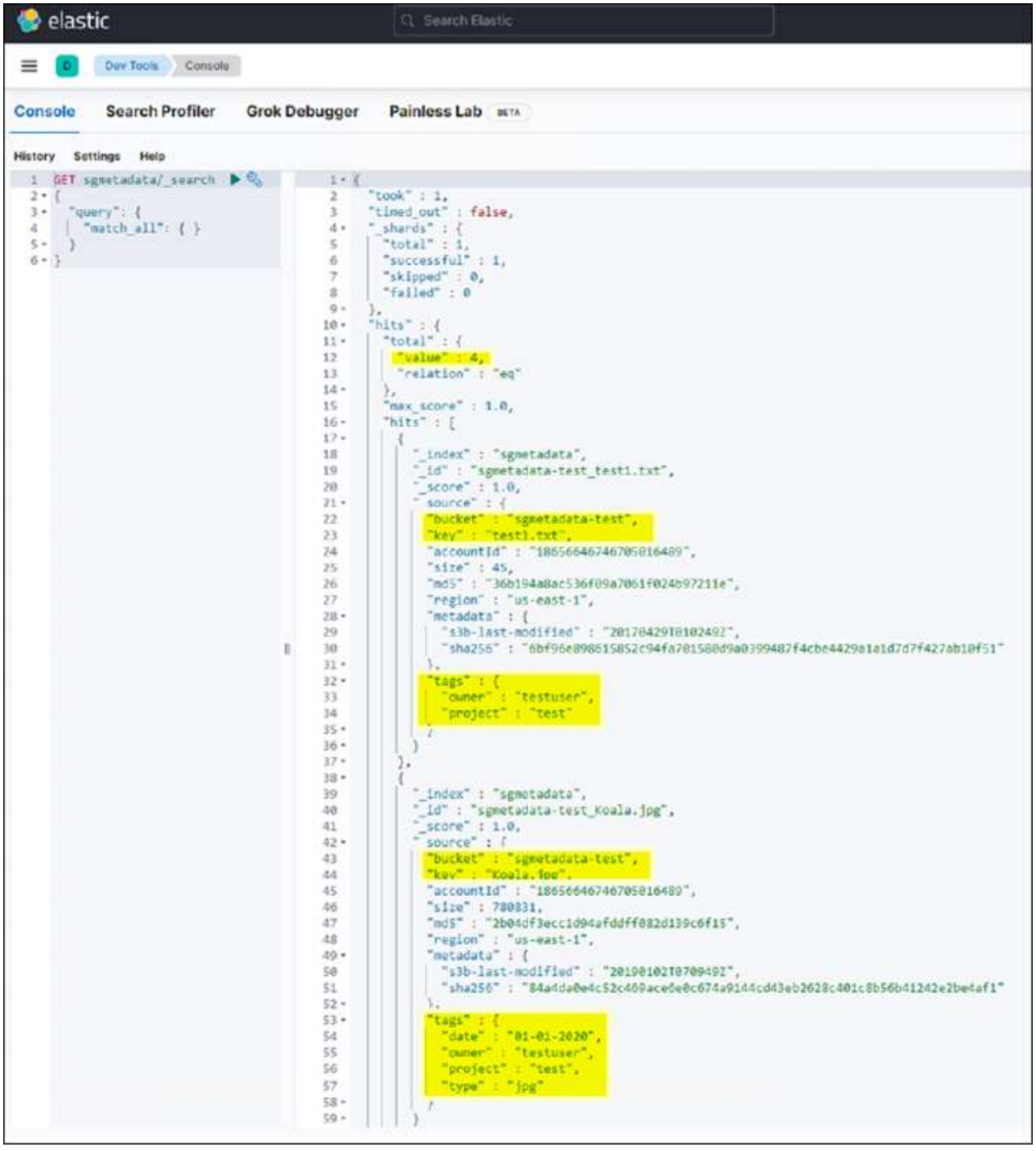
7. Verwenden Sie die Kibana UI, um zu überprüfen, ob die Objektmetadaten in den Index der sgmetadaten geladen wurden.

   a. Wählen Sie im Menü Verwaltung > Entwicklungstools.

   b. Fügen Sie die Beispielabfrage in das Konsolenfenster auf der linken Seite ein, und klicken Sie auf das Dreieckssymbol, um sie auszuführen.

      Das Beispielergebnis für die Abfrage 1 im folgenden Beispiel-Screenshot zeigt vier Datensätze. Dies entspricht der Anzahl der Objekte im Bucket.

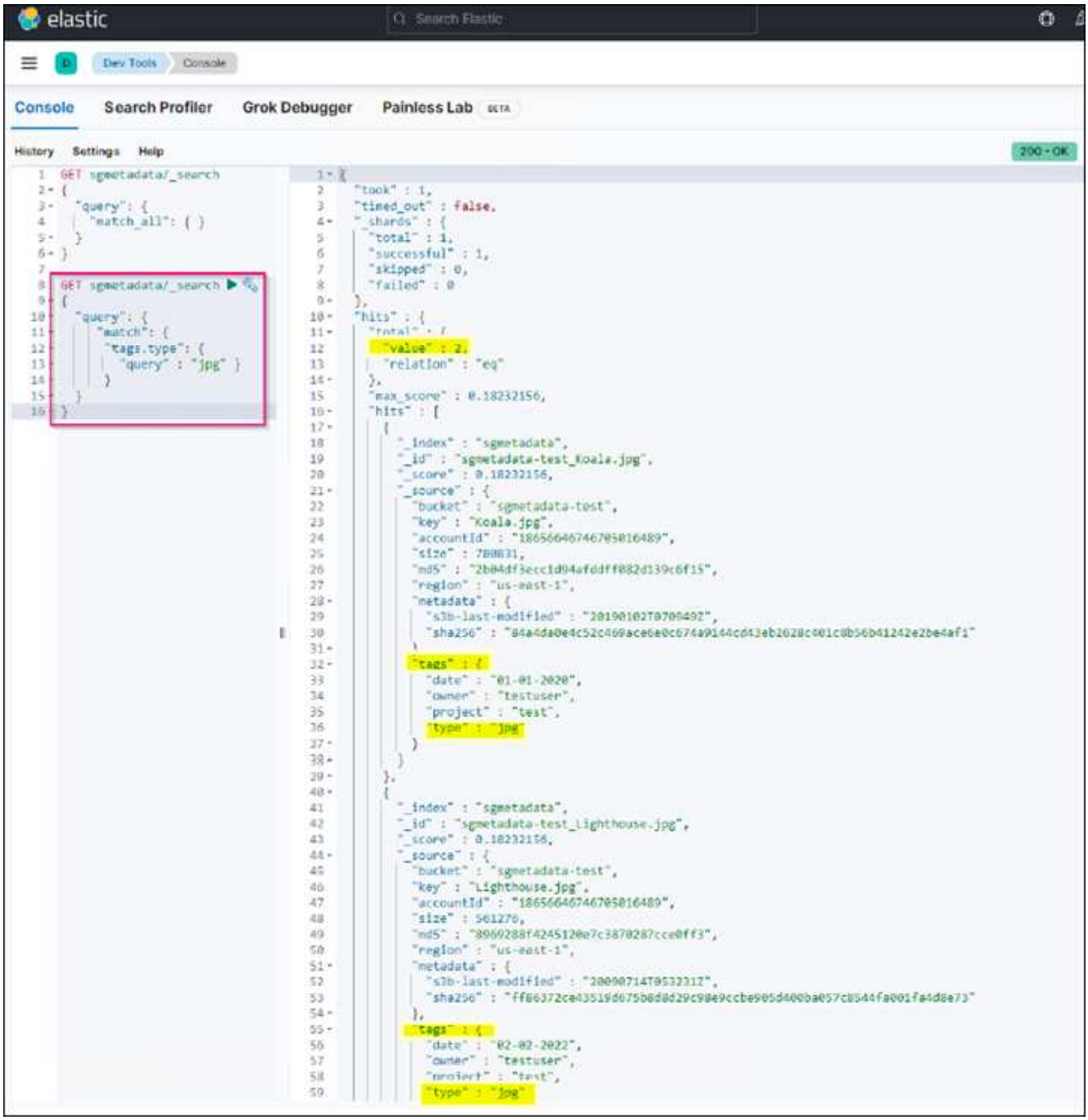
```
GET sgmetadata/_search
{
    "query": {
        "match_all": { }
}
}
```



Das Beispielergebnis für Abfrage 2 im folgenden Screenshot zeigt zwei Datensätze mit Tag-Typ jpg.

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

+



Wo Sie weitere Informationen finden

Sehen Sie sich die folgenden Dokumente und/oder Websites an, um mehr über die in diesem Dokument beschriebenen Informationen zu erfahren:

- "Was sind Plattform-Services"
- "StorageGRID 11.6-Dokumentation"

Von Angela Cheng

Node-Klonen

Überlegungen und Performance von Node-Klonen

Überlegungen zu Node-Klonen

Node-Klone können eine schnellere Methode zum Austausch vorhandener Appliance-Nodes für eine Technologieaktualisierung sein, die Kapazität erhöhen oder die Performance Ihres StorageGRID Systems steigern. Node-Klon kann auch für die Konvertierung in Node-Verschlüsselung mit einem KMS oder die Änderung eines Storage-Node von DDP8 zu DDP16 nützlich sein.

- Die genutzte Kapazität des Quell-Node ist nicht relevant für die Zeit, die für den Abschluss des Klonprozesses erforderlich ist. Node-Klon ist eine vollständige Kopie des Node, einschließlich freiem Speicherplatz im Node.
- Quell- und Ziel-Appliances müssen dieselbe PGE-Version aufweisen
- Der Zielknoten muss immer eine größere Kapazität als die Quelle haben
    - Stellen Sie sicher, dass die neue Ziel-Appliance eine größere Laufwerksgröße als die Quelle hat
    - Wenn das Zielgerät über Laufwerke gleicher Größe verfügt und für DDP8 konfiguriert ist, können Sie das Ziel für DDP16 konfigurieren. Wenn die Quelle bereits für DDP16 konfiguriert ist, ist ein Node-Klon nicht möglich.
    - Beachten Sie beim Wechsel von SG5660 oder SG5760 Appliances zu SG6060 Appliances, dass die SG5x60 60 Laufwerke mit Kapazität haben, bei denen die SG6060 nur 58 hat.
- Für den Klonprozess eines Node muss der Quell-Node für die Dauer des Klonens offline im Grid sein. Wenn ein zusätzlicher Knoten während dieser Zeit offline geht, sind möglicherweise die Client-Services betroffen.
- Ein Storage-Node kann nur 15 Tage lang offline sein. Wenn der Klonprozess fast 15 Tage beträgt oder 15 Tage überschreitet, können Sie das Erweiterungs- und Ausmusterung verwenden.
- Bei einem SG6060 mit Erweiterungs-Shelfs müssen Sie die Zeit für die richtige Shelf-Laufwerksgröße zur Zeit der Basis-Appliance hinzufügen, um die volle Klondauer zu erhalten.
- Die Anzahl der Volumes in einer Ziel-Storage-Appliance muss größer oder gleich der Anzahl der Volumes im Quell-Node sein. Sie können einen Quell-Node mit 16 Objektspeicher-Volumes (rangedb) nicht auf einer Ziel-Storage-Appliance mit 12 Objektspeicher-Volumes klonen, selbst wenn die Ziel-Appliance über eine größere Kapazität als der Quell-Node verfügt. Die meisten Storage Appliances verfügen über 16 Objektspeicher-Volumes, außer der SGF6112 Storage Appliance mit nur 12 Objektspeicher-Volumes. Sie können beispielsweise nicht von einem SG5760 in ein SGF6112 klonen.

Schätzungen der Performance von Node-Klonen

Die folgenden Tabellen enthalten berechnete Schätzungen für die Dauer von Node-Klonen. Die Bedingungen variieren, sodass Einträge in **BOLD** das 15-Tage-Limit für einen Knoten nach unten überschreiten können.

DDP8

SG5612 → beliebig

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	1 Tag	2 Tage	2.5 Tage	3 Tage	4 Tage	4.5 Tage
25 GB	1 Tag	2 Tage	2.5 Tage	3 Tage	4 Tage	4.5 Tage

SG5712 → beliebig

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	1 Tag	2 Tage	2.5 Tage	3 Tage	4 Tage	4.5 Tage
25 GB	1 Tag	2 Tage	2.5 Tage	3 Tage	4 Tage	4.5 Tage

SG5660 → SG5760

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	3 Tage	6 Tage	7 Tage	8.5 Tage	11.5 Tage	**13 Tage**
25 GB	3 Tage	6 Tage	7 Tage	8.5 Tage	11.5 Tage	**13 Tage**

SG5660 → SG6060

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	2.5 Tage	4.5 Tage	5.5 Tage	6.5 Tage	9 Tage	10 Tage
25 GB	2 Tage	4 Tage	5 Tage	6 Tage	8 Tage	9 Tage

SG5760 → SG5760

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	3 Tage	6 Tage	7 Tage	8.5 Tage	11.5 Tage	**13 Tage**
25 GB	3 Tage	6 Tage	7 Tage	8.5 Tage	11.5 Tage	**13 Tage**

SG5760 → SG6060

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	2.5 Tage	4.5 Tage	5.5 Tage	6.5 Tage	9 Tage	10 Tage
25 GB	1.5 Tage	3 Tage	3.5 Tage	4.5 Tage	6 Tage	6.5 Tage

SG6060 → SG6060

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	2.5 Tage	4.5 Tage	5.5 Tage	6.5 Tage	8.5 Tage	9.5 Tage
25 GB	1.5 Tage	3 Tage	3.5 Tage	4 Tage	5.5 Tage	6 Tage

DDP16

SG5760 → SG5760

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	3.5 Tage	6.5 Tage	8 Tage	9.5 Tage	12.5 Tage	**14 Tage**
25 GB	3.5 Tage	6.5 Tage	8 Tage	9.5 Tage	12.5 Tage	**14 Tage**

SG5760 → SG6060

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	2.5 Tage	5 Tage	6 Tage	7.5 Tage	10 Tage	11 Tage

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
25 GB	2 Tage	3.5 Tage	4 Tage	5 Tage	6.5 Tage	7 Tage

SG6060 → SG6060

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	3.5 Tage	5 Tage	6 Tage	7 Tage	9.5 Tage	10.5 Tage
25 GB	2 Tage	3 Tage	4 Tage	4.5 Tage	6 Tage	7 Tage

Erweiterungs-Shelf (oberhalb von SG6060 für jedes Shelf in der Quell-Appliance hinzufügen)

Geschwindigkeit der Netzwerkschnittstelle	Größe des 4-TB-Laufwerks	8-TB-Laufwerke	10-TB-Laufwerkgröße	12-TB-Laufwerksgröße	16-TB-Laufwerkgröße	18 TB Laufwerksgröße
10 GBIT	3.5 Tage	5 Tage	6 Tage	7 Tage	9.5 Tage	10.5 Tage
25 GB	2 Tage	3 Tage	4 Tage	4.5 Tage	6 Tage	7 Tage

Von Aron Klein

So verwenden Sie die Port-Neuzuordnung

Möglicherweise müssen Sie einen eingehenden oder ausgehenden Port aus mehreren Gründen neu zuordnen. Möglicherweise wechseln Sie vom älteren CLB Load Balancer-Dienst zum aktuellen nginx Service Load Balancer-Endpunkt und behalten denselben Port bei, um die Auswirkungen auf Clients zu reduzieren. Sie möchten Port 443 für Client S3 in einem Client-Netzwerk des Admin-Knotens oder für Firewall-Einschränkungen verwenden.

Migrieren Sie S3-Clients von CLB auf NGINX mit Port-Neuzuordnung

In Versionen vor StorageGRID 11.3 ist der Verbindungs-Lastausgleich (CLB) der enthaltene Load Balancer auf den Gateway-Nodes. Im StorageGRID 11.3 stellt NetApp den NGINX-Service als funktionsreiche integrierte Lösung für den Lastausgleich von HTTP(s) Traffic vor. Da der CLB-Dienst in der aktuellen Version von StorageGRID verfügbar bleibt, können Sie Port 8082 nicht in der neuen Endpunktkonfiguration des Load Balancer wiederverwenden. Um dies zu umgehen, wird der eingehende Port 8082 erneut 10443 zugeordnet. Dadurch werden alle HTTPS-Anfragen an Port 8082 des Gateways an Port 10443 umgeleitet, wobei der CLB-Dienst umgangen und stattdessen eine Verbindung zum NGINX-Dienst hergestellt wird. Obwohl die folgenden Anweisungen für VMware gelten, ist die FUNKTION PORT_REMAP für alle Installationsmethoden vorhanden, und Sie können einen ähnlichen Prozess für Bare-Metal-Bereitstellungen und -Appliances verwenden.

Implementierung von VMware Virtual Machine Gateway Node

Die folgenden Schritte gelten für eine StorageGRID-Bereitstellung, bei der der Gateway-Knoten oder -Knoten in VMware vSphere 7 als VMs mit dem StorageGRID Open Virtualization Format (OVF) bereitgestellt werden. Der Prozess beinhaltet das zerstörerische Entfernen der VM und die erneute Bereitstellung der VM mit demselben Namen und derselben Konfiguration. Bevor Sie die VM einschalten, ändern Sie die vApp-Eigenschaft, um den Port neu zuzuordnen, schalten Sie die VM ein und folgen Sie dem Wiederherstellungsprozess für den Knoten.

Voraussetzungen

- Sie verwenden StorageGRID 11.3 oder höher
- Sie haben heruntergeladen und haben Zugriff auf die installierten VMware-Installationsdateien der StorageGRID-Version.
- Sie haben ein vCenter Konto mit Berechtigungen zum ein-/Ausschalten von VMs, Ändern der Einstellungen der VMs und vApps, Entfernen von VMs aus vCenter und Bereitstellen von VMs durch OVF.
- Sie haben einen Load Balancer-Endpunkt erstellt
  - Der Port ist für den gewünschten Umleitungsport konfiguriert
  - Das Endpunkt-SSL-Zertifikat ist dasselbe wie für den CLB-Dienst im Serverzertifikat für Konfiguration/Serverzertifikate/Objekt-Storage-API-Dienst installiert, oder der Client kann eine Änderung des Zertifikats akzeptieren.

> ⓘ
>
> ```
> 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.
> ```

Zerstören Sie den ersten Gateway-Node

Gehen Sie wie folgt vor, um den ersten Gateway-Node zu zerstören:

1. Wählen Sie den Gateway Node aus, mit dem Sie beginnen möchten, wenn das Grid mehr als einen enthält.
2. Entfernen Sie die Node-IPs von allen DNS-Round-Robin-Entitäten oder Load-Balancer-Pools, falls zutreffend.
3. Warten Sie, bis Time-to-Live (TTL) und geöffnete Sitzungen ablaufen.
4. Schalten Sie den VM-Knoten aus.
5. Entfernen Sie den VM-Node von der Festplatte.

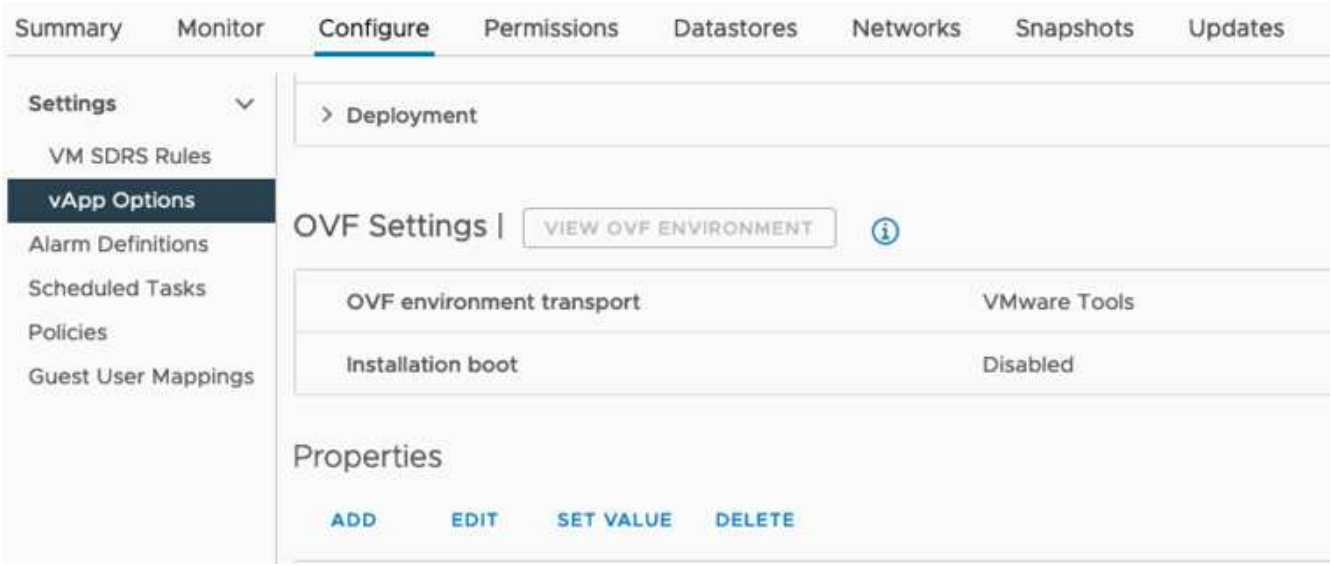
Implementieren Sie den Ersatz-Gateway-Node

Gehen Sie wie folgt vor, um den Ersatz-Gateway-Node bereitzustellen:

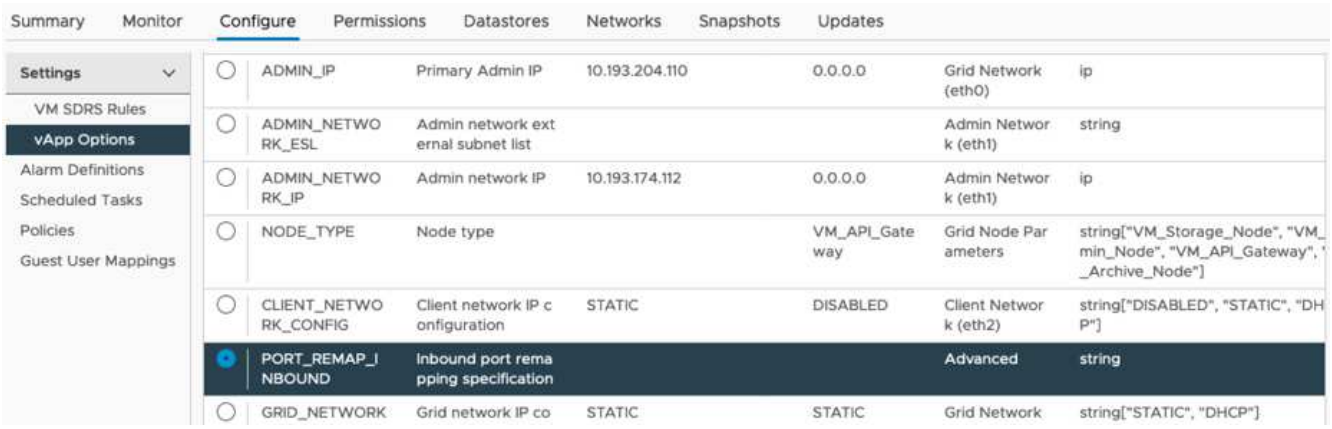
1. Implementieren Sie die neue VM über OVF, und wählen Sie die Dateien .ovf, .MF und .vmdk aus dem von der Support-Website heruntergeladenen Installationspaket aus:
   - vsphere-Gateway.MF
   - vsphere-Gateway.ovf

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

2. Nachdem die VM bereitgestellt wurde, wählen Sie sie aus der Liste der VMs aus und wählen Sie die Registerkarte Konfigurieren vApp Options aus.



3. Blättern Sie nach unten zum Abschnitt Eigenschaften, und wählen Sie die Eigenschaft PORT_REMAP_INBOUND aus

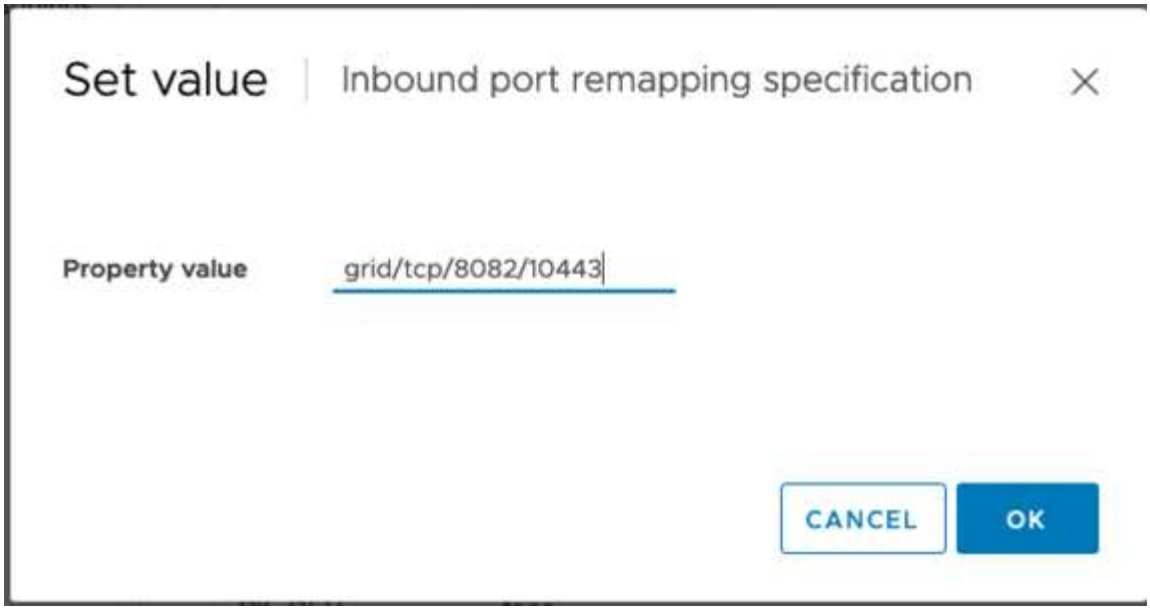


4. Blättern Sie nach oben in der Liste Eigenschaften, und klicken Sie auf Bearbeiten

5. Wählen Sie die Registerkarte Typ aus, bestätigen Sie, dass das Kontrollkästchen Benutzerkonfigurierbar aktiviert ist, und klicken Sie dann auf Speichern.

6. Klicken Sie oben in der Liste Eigenschaften, wenn die Eigenschaft „PORT_REMAP_INBOUND“ noch ausgewählt ist, auf Wert festlegen.

7. Geben Sie im Feld Eigenschaftswert das Netzwerk (Grid, admin oder Client), TCP, den ursprünglichen Port (8082) und den neuen Port (10443) mit „/“ zwischen den einzelnen Werten ein, wie nachfolgend dargestellt.

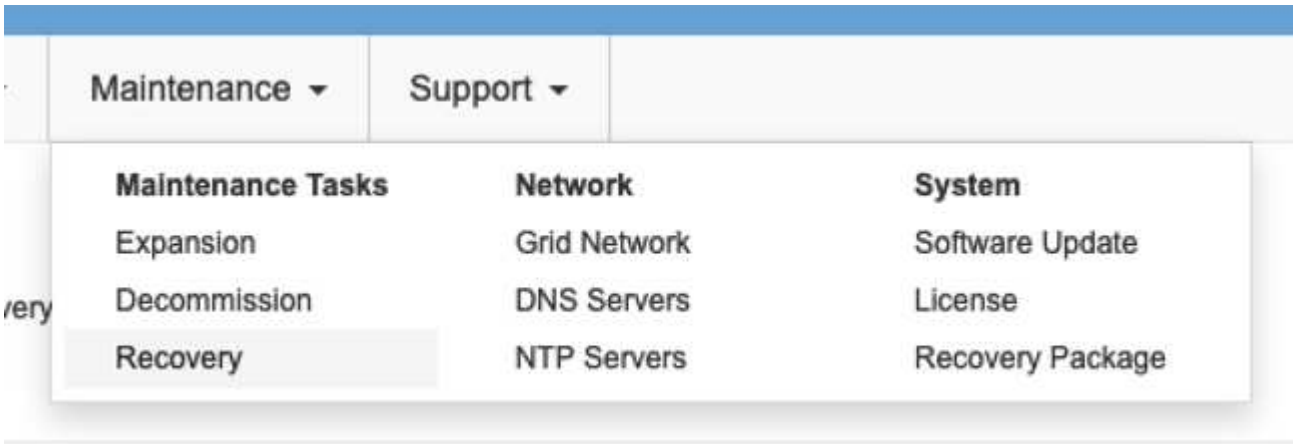


8. Wenn Sie mehrere Netzwerke verwenden, trennen Sie die Netzwerkzeichenfolgen mit einem Komma (,), z. B. GRID/tcp/8082/10443,admin/tcp/8082/10443,Client/tcp/8082/10443

Stellen Sie den Gateway-Node wieder her

Gehen Sie wie folgt vor, um den Gateway-Node wiederherzustellen:

1. Navigieren Sie zum Abschnitt Wartung/Wiederherstellung der Grid Management-Benutzeroberfläche.



2. Schalten Sie den VM-Knoten ein, und warten Sie, bis der Knoten im Abschnitt Wartung/Wiederherstellung ausstehender Knoten der Grid Management-Benutzeroberfläche angezeigt wird.

> 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]

3. Nachdem der Node wiederhergestellt wurde, kann die IP in alle DNS-Round-Robin-Einheiten oder, falls zutreffend, in Load-Balancer-Pools enthalten sein.

   Jetzt gehen alle HTTPS-Sitzungen auf Port 8082 zu Port 10443

Port 443 für den Client-S3-Zugriff auf einen Admin-Node neu zuordnen

Die Standardkonfiguration im StorageGRID-System für einen Admin-Node oder eine HA-Gruppe mit einem Admin-Node lautet, dass Port 443 und 80 für die Management- und Mandantenmanager-UI reserviert werden und nicht für Load Balancer-Endpunkte verwendet werden können. Die Lösung hierfür besteht darin, die Funktion zur Portzuordnung zu verwenden und den eingehenden Port 443 an einen neuen Port weiterzuleiten, der als Load Balancer-Endpunkt konfiguriert wird. Sobald der Client-S3-Datenverkehr abgeschlossen ist, kann Port 443 verwendet werden, die Grid-Management-UI ist nur über Port 8443 zugänglich, und die Mandantenmanagement-UI ist nur über Port 9443 zugänglich. Die Neuzuordnungsfunktion kann nur zum Installationszeitpunkt des Node konfiguriert werden. Um eine Port-Neuzuordnung eines aktiven Node im Grid zu implementieren, muss dieser auf den vorinstallierten Status zurückgesetzt werden. Dies ist ein destruktives Verfahren, das nach Durchführung der Konfigurationsänderung eine Recovery des Node einschließt.

Backup-Protokolle und Datenbanken

Administrator-Nodes enthalten Audit-Protokolle, prometheus-Kennzahlen sowie Verlaufsinformationen zu Attributen, Alarmen und Alarmen. Bei mehreren Administrator-Nodes haben Sie mehrere Kopien dieser Daten. Wenn sich in dem Grid nicht mehrere Administrator-Nodes befinden, sollten Sie diese Daten zur Wiederherstellung beibehalten, nachdem der Node nach Abschluss dieses Prozesses wiederhergestellt wurde. Wenn sich in Ihrem Grid ein anderer Administrator-Node befindet, können Sie die Daten von diesem Node während des Recovery-Prozesses kopieren. Wenn sich kein weiterer Admin-Node im Raster befindet, können Sie die Daten vor dem Zerstören des Node anhand der folgenden Anweisungen kopieren.

Prüfprotokolle kopieren

1. Melden Sie sich beim Admin-Knoten an:
   a. Geben Sie den folgenden Befehl ein: `ssh admin@grid_node_IP`
   b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`

d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`

f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

2. Erstellen Sie das Verzeichnis, um alle Audit-Log-Dateien an einen temporären Speicherort auf einem separaten Grid-Knoten zu kopieren. Verwenden Sie *Storage_Node_01*:

a. `ssh admin@storage_node_01_IP`

b. `mkdir -p /var/local/tmp/saved-audit-logs`

3. Beenden Sie den AMS-Dienst wieder auf dem Admin-Knoten, um zu verhindern, dass er eine neue Protokolldatei erstellt: `service ams stop`

4. Benennen Sie die Datei audit.log um, damit sie die vorhandene Datei nicht überschreiben kann, wenn Sie sie in den wiederhergestellten Admin-Node kopieren.

a. Benennen Sie audit.log in einen eindeutigen nummerierten Dateinamen um, z. B. yyyy-mm-dd.txt.1. Sie können beispielsweise die Audit-Log-Datei in 2015-10-25.txt.1 umbenennen

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

5. AMS-Dienst neu starten: `service ams start`

6. Alle Audit-Log-Dateien kopieren: `scp * admin@storage_node_01_IP:/var/local/tmp/saved-audit-logs`

Kopieren Sie Prometheus Daten

Das Kopieren der Prometheus-Datenbank dauert möglicherweise ein Stunde oder länger. Einige Grid Manager-Funktionen sind nicht verfügbar, während Dienste auf dem Admin-Knoten angehalten werden.

1. Erstellen Sie das Verzeichnis, um die prometheus-Daten an einen temporären Speicherort auf einem separaten Grid-Knoten zu kopieren, auch hier wird *Storage_Node_01* verwendet:

a. Melden Sie sich beim Speicher-Node an:

i. Geben Sie den folgenden Befehl ein: `ssh admin@storage_node_01_IP`

ii. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

iii. Mkdir -p /var/local/tmp/prometheus`

2. Melden Sie sich beim Admin-Knoten an:

a. Geben Sie den folgenden Befehl ein: `ssh admin@admin_node_IP`

b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`

d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`

f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

3. Halten Sie vom Admin-Knoten den Prometheus-Service an: `service prometheus stop`

   a. Prometheus-Datenbank vom Quell-Admin-Node auf den Speicher-Node-Backup-Speicherort kopieren Knoten: `/rsync -azh --stats "/var/local/mysql_ibdata/prometheus/data" "`*`storage_node_01_IP`*`:/var/local/tmp/prometheus/"`

4. Starten Sie den Prometheus-Service auf dem Quell-Admin-Node neu.`service prometheus start`

Sichern Sie Verlaufsinformationen

Die historischen Informationen werden in einer mysql-Datenbank gespeichert. Um eine Kopie der Datenbank abzuladen, benötigen Sie den Benutzer und das Passwort von NetApp. Wenn sich in der Tabelle ein weiterer Admin-Node befindet, ist dieser Schritt nicht erforderlich. Die Datenbank kann während der Recovery von einem verbleibenden Admin-Node geklont werden.

1. Melden Sie sich beim Admin-Knoten an:

   a. Geben Sie den folgenden Befehl ein: `ssh admin@`*`admin_node_IP`*

   b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

   c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`

   d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

   e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`

   f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

2. Stoppen Sie StorageGRID-Dienste auf Admin-Knoten und starten sie ntp und mysql

   a. Beenden Sie alle Dienste: `service servermanager stop`

   b. Starten Sie den ntp-Service neu: `service ntp start`..Neustart mysql-Dienst: `service mysql start`

3. Dump mi-Datenbank in /var/local/tmp

   a. Geben Sie den folgenden Befehl ein: `mysqldump –u `*`username`*` –p `*`password`*` mi > /var/local/tmp/mysql-mi.sql`

4. Kopieren Sie die mysql Dump-Datei auf einen alternativen Knoten, wir verwenden *Storage_Node_01: `scp /var/local/tmp/mysql-mi.sql _storage_node_01_IP`:/var/local/tmp/mysql-mi.sql*

a. Wenn Sie keinen passwortlosen Zugriff auf andere Server mehr benötigen, entfernen Sie den privaten Schlüssel vom SSH-Agent. Geben Sie Ein: `ssh-add -D`

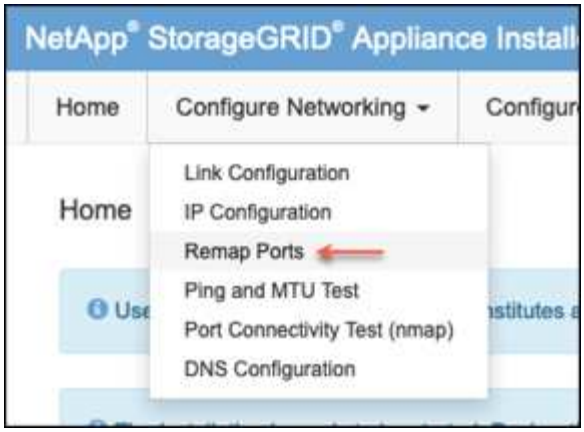
Erstellen Sie den Admin-Knoten neu

Nachdem Sie nun über eine Backup-Kopie aller gewünschten Daten und Protokolle verfügen, die sich entweder auf einem anderen Admin-Node im Grid oder an einem temporären Speicherort befinden, ist es an der Zeit, die Appliance zurückzusetzen, damit die Port-Neuzuordnung konfiguriert werden kann.

1. Wenn Sie eine Appliance zurücksetzen, wird sie in den vorinstallierten Zustand zurückversetzt, wobei nur der Hostname, die IP-Adressen und die Netzwerkkonfigurationen beibehalten werden. Alle Daten gehen verloren, weshalb wir dafür gesorgt haben, dass alle wichtigen Informationen gesichert sind.
   a. Geben Sie den folgenden Befehl ein: `sgareinstall`

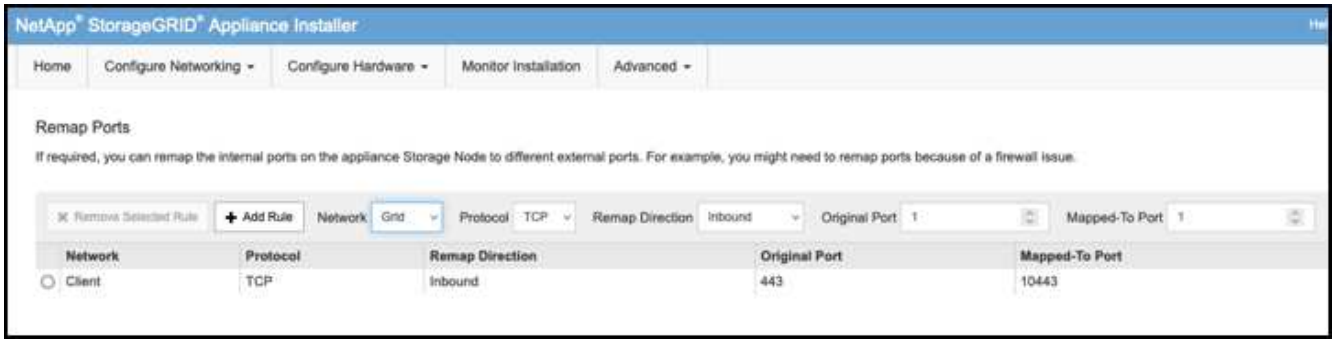
```
root@sg100-01:~ # sgareinstall
WARNING: All StorageGRID Webscale services on this node will be shut
down.
WARNING: Data stored on this node may be lost.
WARNING: You will have to reinstall StorageGRID Webscale to this
node.

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. Nach einiger Zeit wird die Appliance neu gestartet, und Sie können auf die Knoten-PGE-Benutzeroberfläche zugreifen.
3. Navigieren Sie zum Fenster Netzwerk konfigurieren



4. Wählen Sie das gewünschte Netzwerk, Protokoll, Richtung und Ports aus, und klicken Sie dann auf die Schaltfläche Regel hinzufügen.

> Die Neuzuordnung von eingehendem Port 443 auf dem GRID-Netzwerk bricht die Installation und die Erweiterungsverfahren ab. Es wird nicht empfohlen, Port 443 im NETZNETZWERK neu zuzuordnen.



5. Eine der gewünschten Port-Neuzuordnungen wurde hinzugefügt. Sie können zur Registerkarte „Home“ zurückkehren und auf die Schaltfläche „Installation starten“ klicken.

Sie können nun die Wiederherstellungsverfahren für den Admin-Knoten in befolgen "Produktdokumentation"

Wiederherstellung von Datenbanken und Protokollen

Nach der Wiederherstellung des Admin-Node können Sie nun die Metriken, Protokolle und Verlaufsinformationen wiederherstellen. Wenn sich ein anderer Administrator-Node im Raster befindet, folgen Sie den Anweisungen "Produktdokumentation" Verwenden der Skripte *prometheus-Clone-db.sh* und *mi-Clone-db.sh*. Wenn dies der einzige Admin-Node ist und Sie diese Daten sichern möchten, können Sie die folgenden Schritte ausführen, um die Informationen wiederherzustellen.

Kopieren Sie die Prüfprotokolle zurück

1. Melden Sie sich beim Admin-Knoten an:
   a. Geben Sie den folgenden Befehl ein: `ssh admin@grid_node_IP`
   b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:
   c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`
   d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`

f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

2. Kopieren Sie die erhaltenen Audit-Log-Dateien auf den wiederhergestellten Admin-Knoten: `scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY* .`
3. Löschen Sie aus Sicherheitsgründen die Prüfprotokolle vom fehlgeschlagenen Grid-Knoten, nachdem Sie überprüft haben, ob sie erfolgreich auf den wiederhergestellten Admin-Node kopiert wurden.
4. Aktualisieren Sie die Benutzer- und Gruppeneinstellungen der Audit-Log-Dateien auf dem wiederhergestellten Admin-Knoten: `chown ams-user:bycast *`

Sie müssen auch alle bereits vorhandenen Clientzugriffe auf die Revisionsfreigabe wiederherstellen. Weitere Informationen finden Sie in den Anweisungen zum Verwalten von StorageGRID.

Restore von Prometheus Kennzahlen

> Das Kopieren der Prometheus-Datenbank dauert möglicherweise ein Stunde oder länger. Einige Grid Manager-Funktionen sind nicht verfügbar, während Dienste auf dem Admin-Knoten angehalten werden.

1. Melden Sie sich beim Admin-Knoten an:
   a. Geben Sie den folgenden Befehl ein: `ssh admin@grid_node_IP`
   b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:
   c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`
   d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:
   e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`
   f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

2. Halten Sie vom Admin-Knoten den Prometheus-Service an: `service prometheus stop`
   a. Kopieren Sie die Prometheus Datenbank vom Speicherort für temporäre Backups auf den Admin-Node: `/rsync -azh --stats "backup_node:/var/local/tmp/prometheus/" "/var/local/mysql_ibdata/prometheus/"`
   b. Überprüfen Sie, ob sich die Daten im richtigen Pfad befinden und vollständig sind `ls /var/local/mysql_ibdata/prometheus/data/`
3. Starten Sie den Prometheus-Service auf dem Quell-Admin-Node neu.`service prometheus start`

Historische Informationen wiederherstellen

1. Melden Sie sich beim Admin-Knoten an:

a. Geben Sie den folgenden Befehl ein: `ssh admin@grid_node_IP`

b. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

c. Geben Sie den folgenden Befehl ein, um zum Root zu wechseln: `su -`

d. Geben Sie das im aufgeführte Passwort ein `Passwords.txt` Datei:

e. Fügen Sie den SSH-privaten Schlüssel zum SSH-Agenten hinzu. Geben Sie Ein: `ssh-add`

f. Geben Sie das SSH-Zugriffspasswort ein, das im aufgeführt ist `Passwords.txt` Datei:

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

2. Kopieren Sie die mysql-Dump-Datei vom alternativen Knoten: `scp grid_node_IP_:/var/local/tmp/mysql-mi.sql /var/local/tmp/mysql-mi.sql`
3. Stoppen Sie StorageGRID-Dienste auf Admin-Knoten und starten sie ntp und mysql
   a. Beenden Sie alle Dienste: `service servermanager stop`
   b. Starten Sie den ntp-Service neu: `service ntp start`..Neustart mysql-Dienst: `service mysql start`
4. Legen Sie die mi-Datenbank ab und erstellen Sie eine neue leere Datenbank: `mysql -u username -p password -A mi -e "drop database mi; create database mi;"`
5. Stellen Sie die mysql-Datenbank aus dem Datenbank-Dump wieder her: `mysql -u username -p password -A mi < /var/local/tmp/mysql-mi.sql`
6. Starten Sie alle anderen Dienste neu `service servermanager start`

Von Aron Klein

Standortverlagerung von Grid-Standorten und standortweites Netzwerkänderungsverfahren

Dieser Leitfaden beschreibt die Vorbereitung und das Verfahren für den Standortwechsel in einem Grid mit mehreren Standorten von StorageGRID. Sie sollten über ein vollständiges Verständnis dieser Vorgehensweise verfügen und im Voraus planen, um einen reibungslosen Prozess zu gewährleisten und Unterbrechungen für Kunden zu minimieren.

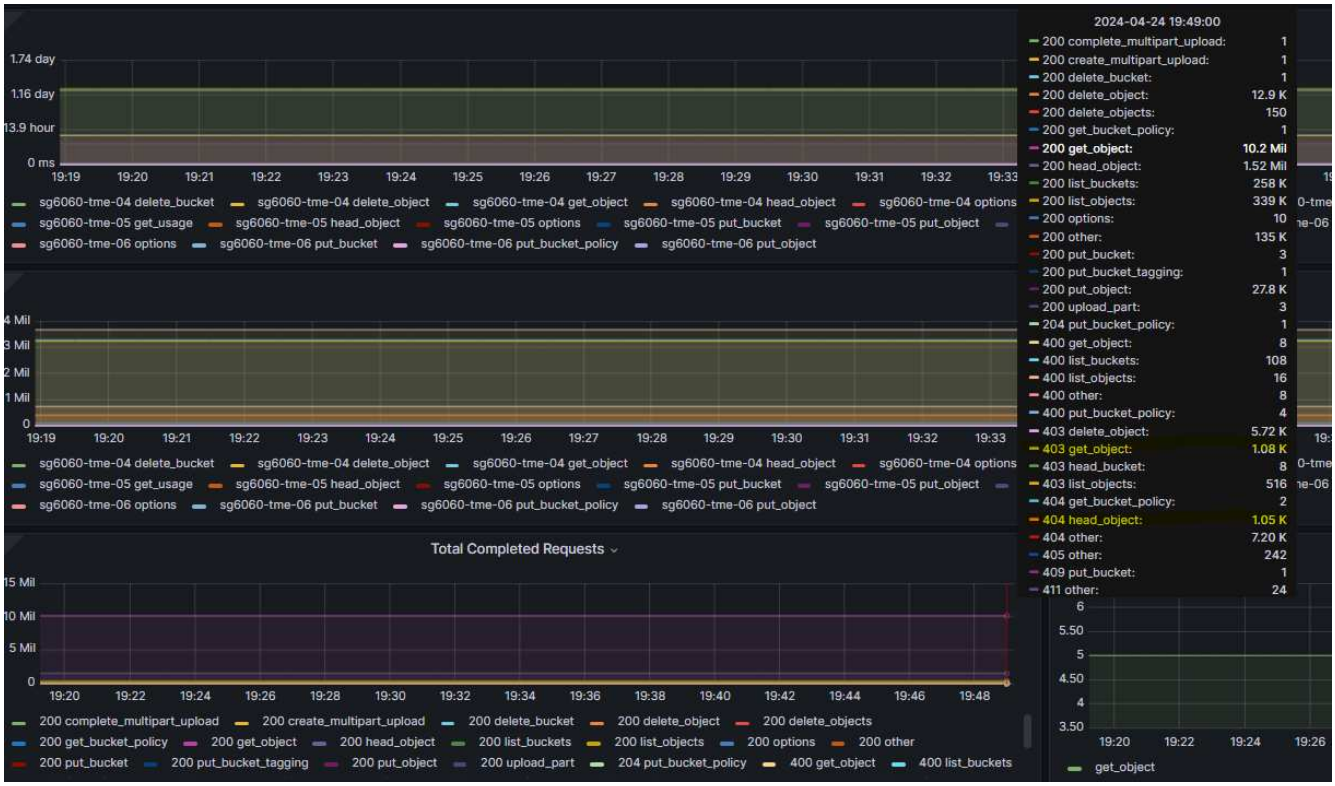
Informationen zum Ändern des Grid-Netzwerks des gesamten Grid finden Sie unter "Ändern Sie die IP-Adressen für alle Nodes im Grid".

Überlegungen vor Standortverlagerung

- Die Standortverschiebungen sollten abgeschlossen sein und alle Nodes innerhalb von 15 Tagen online sein, um eine Wiederherstellung der Cassandra-Datenbank zu vermeiden.
  "Stellen Sie Storage Node länger als 15 Tage wieder her"
- Wenn eine ILM-Regel in der aktiven Richtlinie striktes Aufnahmeverhalten verwendet, sollten Sie sie in Erwägung ziehen, um einen Ausgleich oder eine doppelte Provisionierung zu erreichen, wenn der Kunde

weiterhin Objekte im Grid bei der Standortverlagerung ABLEGEN möchte.

- Bei Storage Appliances mit 60 oder mehr Laufwerken: Verschieben Sie das Shelf niemals bei installierten Festplatten. Beschriften Sie die einzelnen Laufwerke, und entfernen Sie sie vor dem Verpacken/Verschieben aus dem Speichergehäuse.
- Ändern der StorageGRID-Appliance Grid-Netzwerk-VLAN kann Remote über das Admin-Netzwerk oder das Client-Netzwerk durchgeführt werden. Oder planen Sie, vor Ort zu sein, um die Änderung vor oder nach dem Umzug durchzuführen.
- Prüfen Sie, ob die Kundenanwendung vor dem PUT ein Objekt vom TYP HEAD oder GET Nonexistent verwendet. Wenn ja, ändern Sie die Bucket-Konsistenz in strong-site, um HTTP 500-Fehler zu vermeiden. Wenn Sie sich nicht sicher sind, überprüfen Sie die S3-Übersicht Grafana-Diagramme **Grid-Manager > Support > Metriken**, bewegen Sie die Maus über das Diagramm 'gesamte abgeschlossene Anfrage'. Wenn eine sehr hohe Anzahl von 404 get Object oder 404 Head Objects vorhanden ist, verwenden wahrscheinlich eine oder mehrere Anwendungen Head oder Get Nonexistence Objects. Die Zählung wird akkumuliert, Maus über verschiedene Zeitachse, um den Unterschied zu sehen.



Verfahren zum Ändern der Grid-IP-Adresse vor Standortverlagerung

Schritte

1. Wenn das neue Netznetz-Subnetz am neuen Standort verwendet wird, "Fügen Sie das Subnetz der Subnetzliste des Netznetzes hinzu"
2. Melden Sie sich beim primären Admin-Knoten an, verwenden Sie Change-IP, um Grid IP-Änderungen vorzunehmen, müssen Sie die Änderung * inszenieren*, bevor Sie den Knoten für die Verlagerung herunterfahren.
   a. Wählen Sie 2 und dann 1 für Grid IP-Änderung

```
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. Wählen Sie 5, um die Änderungen anzuzeigen

```
================================================================================
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
```

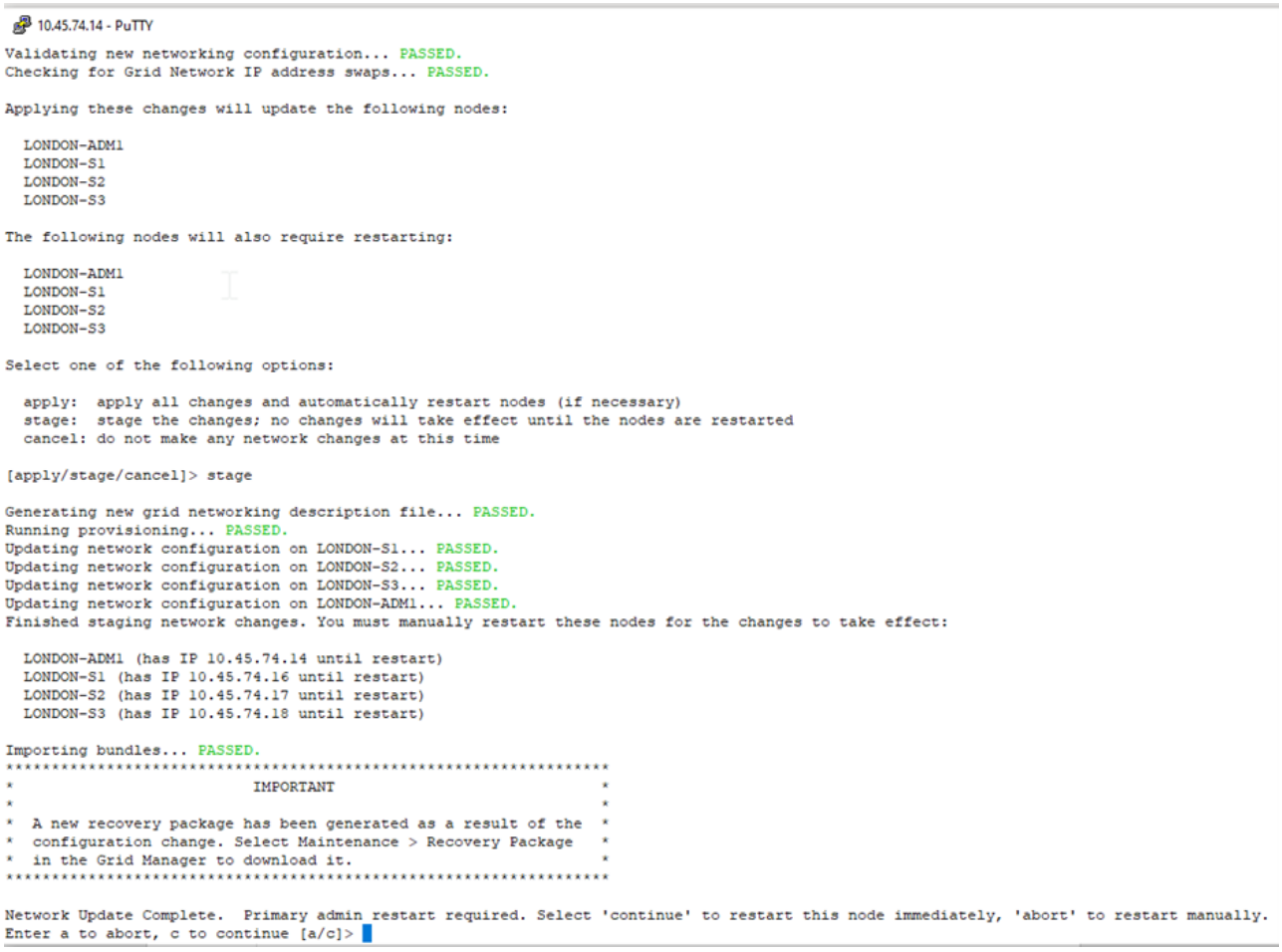
c. Wählen Sie 10, um die Änderung zu validieren und anzuwenden.

```
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
```

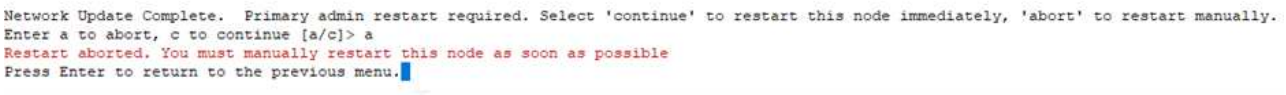
d. In diesem Schritt muss **Stufe** ausgewählt werden.

```
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. Wenn der primäre Admin-Knoten in der obigen Änderung enthalten ist, geben Sie **'a' ein, um den primären Admin-Knoten manuell neu zu starten**



f. Drücken Sie ENTER, um zum vorherigen Menü zurückzukehren und die Change-ip-Schnittstelle zu verlassen.



3. Laden Sie das neue Wiederherstellungspaket vom Grid Manager herunter. **Grid-Manager** > **Wartung** > **Recovery-Paket**
4. Wenn eine VLAN-Änderung auf der StorageGRID-Appliance erforderlich ist, lesen Sie den Abschnitt VLAN-Änderung der Appliance.
5. Fahren Sie alle Knoten und/oder Geräte am Standort herunter, kennzeichnen/entfernen Sie ggf. Festplattenlaufwerke, und entfernen Sie sie aus dem Rack, packen Sie sie aus, und verschieben Sie sie.
6. Wenn Sie die ip-Adresse des Admin-Netzwerks und/oder des Client-VLAN und der ip-Adresse ändern möchten, können Sie die Änderung nach der Verlagerung vornehmen.

VLAN-Änderung der Appliance

Bei der folgenden Vorgehensweise wird davon ausgegangen, dass Sie Remote-Zugriff auf das Admin- oder Client-Netzwerk der StorageGRID Appliance haben, um die Änderung Remote durchzuführen.

Schritte

1. Vor dem Herunterfahren des Geräts "Stellen Sie das Gerät in den Wartungsmodus".

2. Verwenden eines Browsers für den Zugriff auf die StorageGRID-Appliance-Installer-GUI mit https://<admin-or-client-network-ip>:8443. Grid IP kann nicht verwendet werden, da die neue Grid-IP bereits vorhanden ist, sobald die Appliance im Wartungsmodus gestartet wird.
3. Ändern Sie das VLAN für das Grid-Netzwerk. Wenn Sie über das Client-Netzwerk auf die Appliance zugreifen, können Sie das Client-VLAN derzeit nicht ändern. Sie können es nach dem Umzug ändern.
4. ssh zur Appliance und Herunterfahren des Node mit 'shutdoown -h now'
5. Sobald die Appliances an einem neuen Standort bereit sind, können Sie über die Benutzeroberfläche des StorageGRID-Appliance-Installationsprogramms auf zugreifen https://<grid-network-ip>:8443. Überprüfen Sie mithilfe der Ping/nmap-Tools in der GUI, ob sich der Speicher im optimalen Zustand und der Netzwerkverbindung zu anderen Grid-Nodes befindet.
6. Wenn Sie planen, die Client-Netzwerk-IP zu ändern, können Sie das Client-VLAN zu diesem Zeitpunkt ändern. Das Client-Netzwerk ist erst bereit, wenn Sie die Client-Netzwerk-ip-Adresse mit dem Change-ip-Tool in einem späteren Schritt aktualisieren.
7. Beenden Sie den Wartungsmodus. Wählen Sie im Installationsprogramm der StorageGRID-Appliance die Option **Erweitert** > **Controller neu starten** aus, und wählen Sie dann **Neustart in StorageGRID** aus.
8. Wenn alle Nodes eingeschaltet sind und Grid kein Verbindungsproblem zeigt, aktualisieren Sie ggf. das Admin-Netzwerk und das Client-Netzwerk der Appliance mithilfe von Change-ip.

Tool- und Anwendungsleitfäden

Nutzen Sie Hadoop S3A-Connector von Cloudera mit StorageGRID

Hadoop ist bereits seit einiger Zeit ein beliebtes Datenwissenschaftlerteam. Hadoop ermöglicht die verteilte Verarbeitung großer Datensätze über Computer-Cluster mithilfe von einfachen ProgrammierFrameworks. Hadoop wurde entwickelt, um von einzelnen Servern auf Tausende von Machines zu skalieren, wobei jede Maschine über lokale Computing- und Storage-Ressourcen verfügt.

Vorteile von S3A für Hadoop Workflows

Mit der Zeit hat das Datenvolumen zugenommen, aber die Nutzung der IT-Infrastruktur mit eigenen Computing- und Storage-Ressourcen ist ineffizient. Lineare Skalierung führt zu Herausforderungen bei der effizienten Nutzung von Ressourcen und dem Management der Infrastruktur.

Zur Bewältigung dieser Herausforderungen bietet der Hadoop S3A-Client hochperformante I/O-Vorgänge im Vergleich zu S3-Objekt-Storage. Durch die Implementierung eines Hadoop Workflows mit S3A können Sie Objekt-Storage als Daten-Repository nutzen und Computing- und Storage-Ressourcen separat voneinander skalieren. Dadurch wiederum können Sie Computing- und Storage-Ressourcen unabhängig voneinander skalieren. Die Abkopplung von Computing und Storage eröffnet Ihnen auch die Möglichkeit, die passende Menge an Ressourcen für Ihre Rechneraufgaben zu beanspruchen und Kapazitäten basierend auf der Größe Ihres Datensatzes zu bereitstellen. Somit lassen sich die Gesamtbetriebskosten für Hadoop Workflows verringern.

S3A-Anschluss für die Verwendung von StorageGRID konfigurieren

Voraussetzungen

- Einen StorageGRID S3-Endpunkt-URL, einen S3-Zugriffsschlüssel für Mandanten und einen geheimen Schlüssel für Hadoop S3A-Verbindungstests.
- Ein Cloudera Cluster und Root- oder sudo-Berechtigung für jeden Host im Cluster, um das Java-Paket zu installieren.

Seit April 2022 wurde Java 11.0.14 mit Cloudera 7.1.7 gegen StorageGRID 11.5 und 11.6 getestet. Die Java-Versionsnummer kann jedoch bei einer neuen Installation unterschiedlich sein.

Installieren Sie das Java-Paket

1. Prüfen Sie die "Cloudera Support-Matrix" Für die unterstützte JDK-Version.
2. Laden Sie die herunter "Java 11.x-Paket" Das dem Cloudera Cluster-Betriebssystem entspricht. Kopieren Sie dieses Paket auf jeden Host im Cluster. In diesem Beispiel wird das rpm-Paket für CentOS verwendet.
3. Melden Sie sich bei jedem Host als Root an oder verwenden Sie ein Konto mit sudo-Berechtigung. Führen Sie für jeden Host folgende Schritte durch:
   a. Installieren Sie das Paket:

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

b. Überprüfen Sie, wo Java installiert ist. Wenn mehrere Versionen installiert sind, legen Sie die neu installierte Version als Standard fest:

```
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. Fügen Sie diese Zeile am Ende von hinzu `/etc/profile`. Der Pfad sollte dem Pfad der obigen Auswahl entsprechen:

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

d. Führen Sie den folgenden Befehl aus, damit das Profil wirksam wird:

```
source /etc/profile
```

Konfiguration von Cloudera HDFS S3A

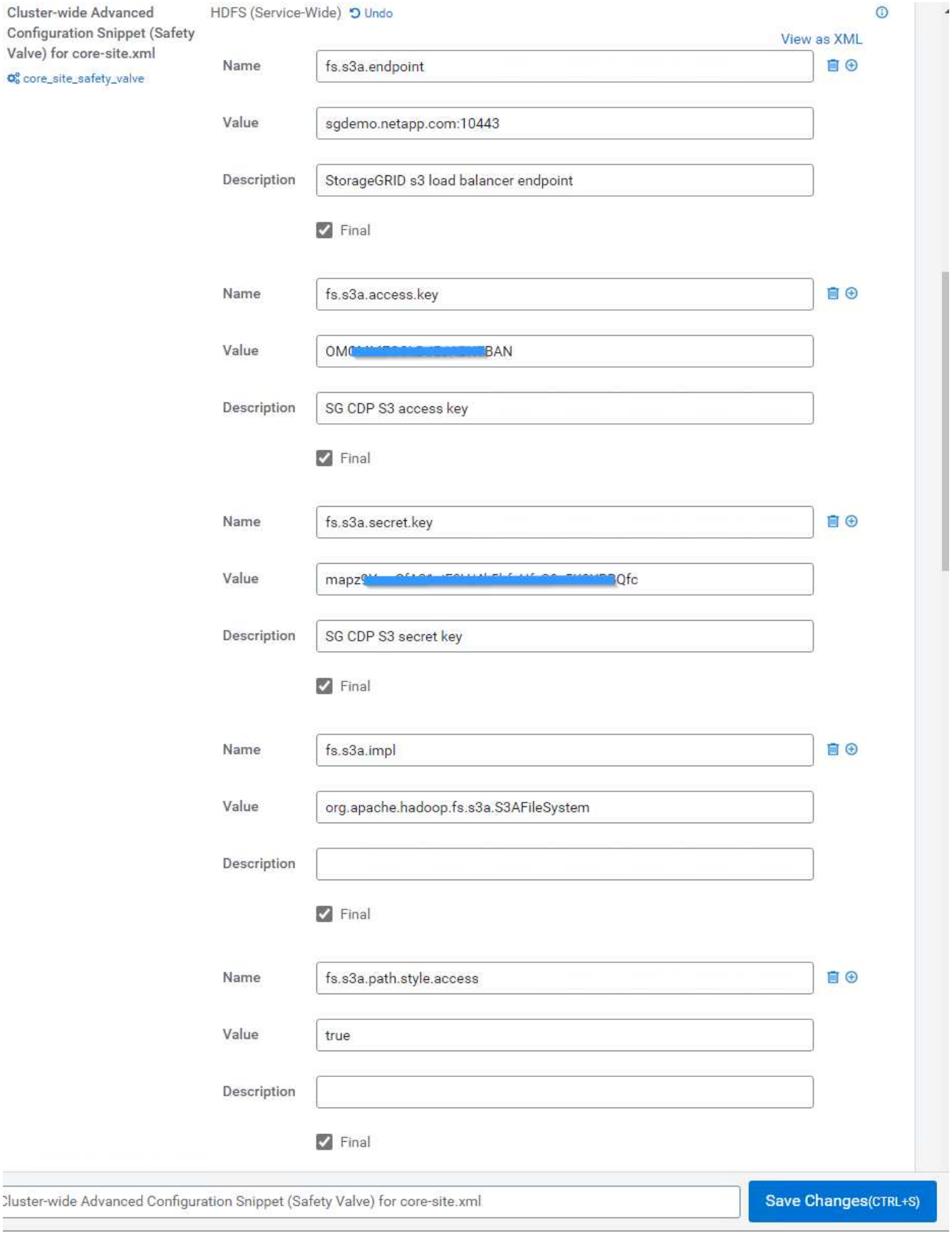
Schritte

1. Wählen Sie in der Cloudera Manager GUI Cluster > HDFS aus, und wählen Sie Konfiguration aus.
2. Wählen Sie unter KATEGORIE die Option Erweitert aus, und blättern Sie nach unten, um zu suchen `Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml`.
3. Klicken Sie auf das (+)-Zeichen und fügen Sie folgende Wertpaare hinzu.

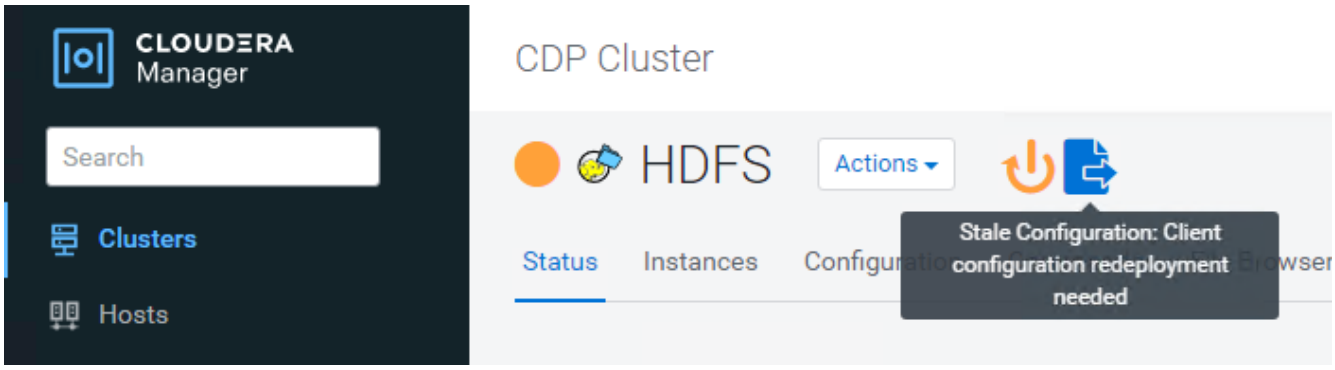
Name	Wert
fs.s3a.access.key	*<S3-Zugriffsschlüssel des Mandanten von StorageGRID>*
fs.s3a.secret.key	*<Mandant s3 geheimen Schlüssel von StorageGRID>*
fs.s3a.connection.ssl.enabled	[Wahr oder falsch] (Standardeinstellung: Https, wenn dieser Eintrag fehlt)

Name	Wert
fs.s3a.Endpunkt	*<StorageGRID S3 Endpunkt:Port>*
fs.s3a.mpl	Org.apache.hadoop.fs.s3a.S3AFileSystem
fs.s3a.path.style.Access	[True oder false] (Standard ist der Stil des virtuellen Hosts, wenn dieser Eintrag fehlt)

Beispiel Screenshot



4. Klicken Sie auf die Schaltfläche Änderungen speichern. Wählen Sie in der HDFS-Menüleiste das Symbol „veraltete Konfiguration“ aus, wählen Sie auf der nächsten Seite „veraltete Dienste neu starten“ und anschließend „Jetzt neu starten“ aus.



S3A-Verbindung mit StorageGRID testen

Führen Sie einen grundlegenden Verbindungstest durch

Melden Sie sich bei einem der Hosts im Cloudera Cluster an, und geben Sie ein `hadoop fs -ls s3a://<bucket-name>/`.

Im folgenden Beispiel wird Pfadsyle mit einem vorhandenen hdfs-Test-Bucket und einem Testobjekt verwendet.

```
[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.
```

Fehlerbehebung

Szenario 1

Verwenden Sie eine HTTPS-Verbindung zu StorageGRID, und holen Sie ein `handshake_failure` Fehler nach einem Timeout von 15 Minuten.

Grund: alte JRE/JDK-Version mit veralteter oder nicht unterstützter TLS-Chiffre-Suite für die Verbindung zu

StorageGRID.

Beispiel-Fehlermeldung

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

Auflösung: stellen Sie sicher, dass JDK 11.x oder höher installiert ist und auf die Java-Bibliothek eingestellt ist. Siehe Installieren Sie das Java-Paket Weitere Informationen finden Sie in.

Szenario 2:

Fehler beim Herstellen der Verbindung zum StorageGRID mit Fehlermeldung `Unable to find valid certification path to requested target`.

Grund: StorageGRID S3-Endpoint-Server-Zertifikat wird nicht von Java-Programm vertrauenswürdig.

Beispielfehlermeldung:

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

Auflösung: NetApp empfiehlt die Verwendung eines Serverzertifikats, das von einer bekannten öffentlichen Zertifizierungsstelle ausgestellt wurde, um die Sicherheit der Authentifizierung sicherzustellen. Alternativ können Sie dem Java Trust Store ein benutzerdefiniertes CA- oder Serverzertifikat hinzufügen.

Führen Sie die folgenden Schritte aus, um eine benutzerdefinierte StorageGRID-Zertifizierungsstelle oder ein Serverzertifikat zum Java-Treuhandspeicher hinzuzufügen.

1. Sichern Sie die vorhandene Standard-Java-Cacaerts-Datei.

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

2. Importieren Sie das StorageGRID S3-Endpunktcert in den Java-Treuhandspeicher.

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

Tipps zur Fehlerbehebung

1. Erhöhen sie den hadoop Protokolllevel zum DEBUGGEN.

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

2. Führen Sie den Befehl aus und leiten Sie die Protokollmeldungen an ERROR.log.

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

Von Angela Cheng

Verwenden Sie S3cmd, um den S3-Zugriff auf StorageGRID zu testen und zu demonstrieren

S3cmd ist ein kostenloses Befehlszeilen-Tool und Client für S3-Vorgänge. Sie können s3cmd verwenden, um den s3-Zugriff auf StorageGRID zu testen und zu demonstrieren.

S3cmd installieren und konfigurieren

Um S3cmd auf einer Workstation oder einem Server zu installieren, laden Sie ihn von herunter "Kommandozeile S3-Client". S3cmd ist vorinstalliert auf jedem StorageGRID-Knoten als Tool zur Unterstützung der Fehlerbehebung.

Erste Konfigurationsschritte

1. S3cmd --configure
2. Geben Sie nur Access_Key und Secret_Key an, für den Rest behalten Sie die Standardeinstellungen.
3. Zugriff mit den angegebenen Zugangsdaten testen? [J/n]: n (den Test umgehen, da er fehlschlägt)
4. Einstellungen speichern? [J/N] J
   a. Konfiguration in '/root/.s3cfg' gespeichert
5. In .s3cfg make fields Host_base and Host_bucket leerer nach dem "=" Zeichen :
   a. Host_Base =
   b. Host_Bucket =

Wenn Sie in Schritt 4 Host_Base und Host_Bucket angeben, müssen Sie in der CLI keinen Endpunkt mit --Host angeben. Beispiel:

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

Beispiele für grundlegende Befehle

- **Erstellen Sie einen Eimer:**

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

- **Alle Buckets auflisten:**

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

- **Alle Eimer und deren Inhalt auflisten:**

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

- **Objekte in einem bestimmten Bucket auflisten:**

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

- **Ein Eimer löschen:**

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

- **Legen Sie ein Objekt:**

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

- **Holen Sie sich ein Objekt:**

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

- **Ein Objekt löschen:**

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

Von Aron Klein

Vertica Eon-Modus-Datenbank mit NetApp StorageGRID als gemeinschaftliche Storage-Lösung

In diesem Leitfaden werden die Verfahren zum Erstellen einer Vertica Eon-Modus-Datenbank mit gemeinsamem Speicher auf NetApp StorageGRID beschrieben.

Einführung

Vertica ist eine Software für das Analyse-Datenbankmanagement. Es handelt sich um eine spaltenbasierte Storage-Plattform zur Verarbeitung großer Datenvolumen. Damit ermöglicht sie eine sehr schnelle Abfrage-Performance in einem klassisch intensiven Szenario. Eine Vertica-Datenbank läuft in einem der beiden Modi: Eon oder Enterprise. Beide Modi können sowohl lokal als auch in der Cloud implementiert werden.

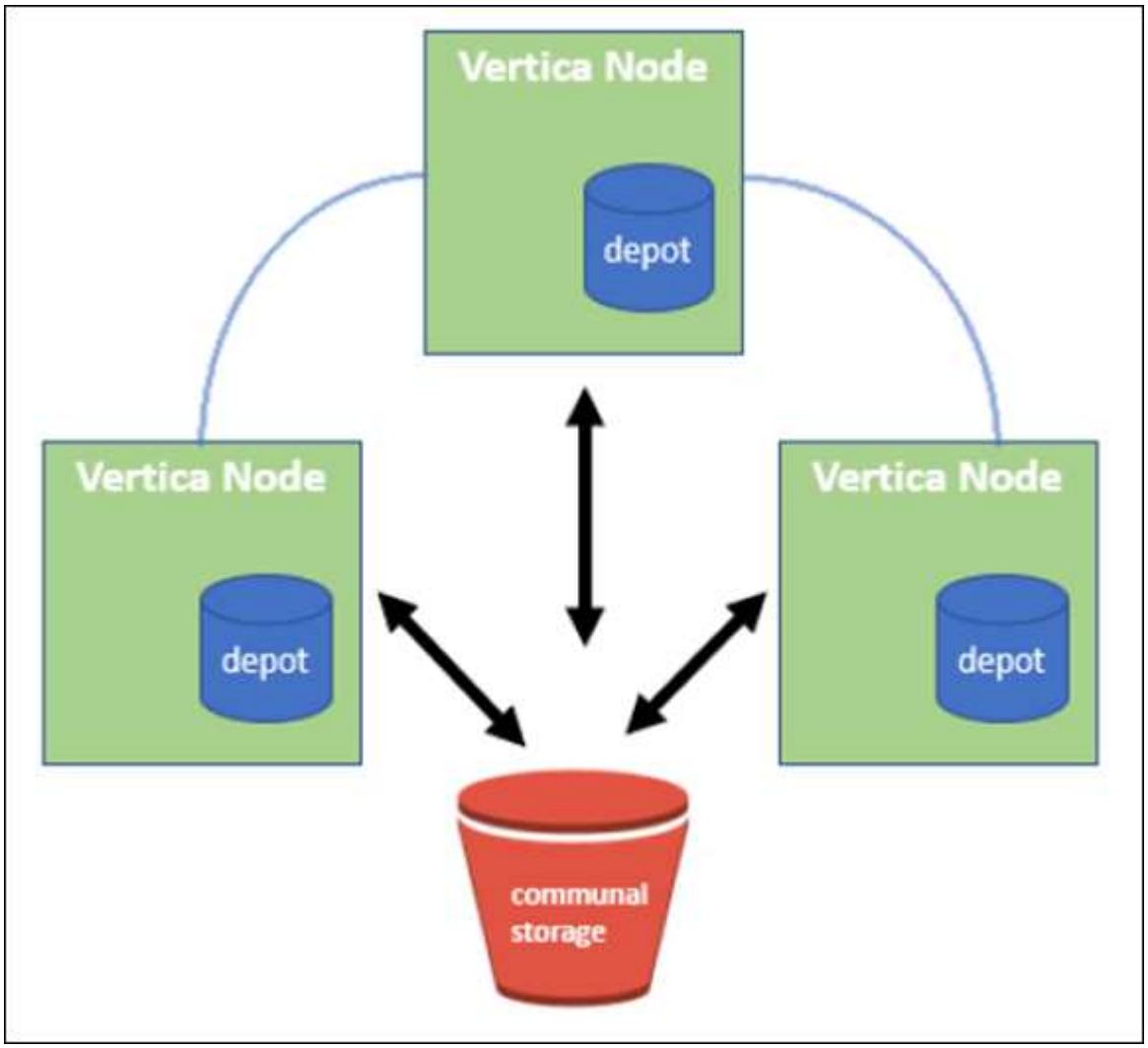
Eon- und Enterprise-Modi unterscheiden sich hauptsächlich darin, wo sie Daten speichern:

- Die Eon-Modus-Datenbanken verwenden einen gemeinsamen Speicher für ihre Daten. Dies wird von Vertica empfohlen.
- Die Enterprise-Mode-Datenbanken speichern die Daten lokal im Dateisystem der Knoten, aus denen die Datenbank besteht.

Eon-Mode-Architektur

Eon-Modus trennt die Computing-Ressourcen von der gemeinsamen Storage-Schicht der Datenbank, wodurch Computing- und Storage-Ressourcen getrennt skaliert werden können. Vertica im Eon-Modus ist für variable Workloads optimiert und kann durch die Nutzung separater Computing- und Storage-Ressourcen voneinander isoliert werden.

Eon-Modus speichert Daten in einem gemeinsam genutzten Objektspeicher, dem sogenannten Communal Storage – einem S3-Bucket, der entweder lokal oder in Amazon S3 gehostet wird.



Gemeinschaftsspeicher

Statt Daten lokal zu speichern, verwendet der Eon-Modus einen einzigen gemeinsamen Speicherort für alle Daten und den Katalog (Metadaten). Der zentrale Speicherort der Datenbank, der von den Datenbank-Nodes gemeinsam genutzt wird, ist die gemeinsame Speicherung.

Kommunale Lagerung hat folgende Eigenschaften:

- Kommunaler Storage in der Cloud oder On-Premises-Objekt-Storage ist ausfallsicherer und anfällig für Datenverluste aufgrund von Storage-Ausfällen als Storage auf Festplatte an individuellen Maschinen.
- Alle Daten können von jedem Node aus denselben Pfad gelesen werden.
- Die Kapazität ist nicht durch den Festplattenspeicher auf Nodes begrenzt.
- Da Daten communal gespeichert werden, können Sie den Cluster flexibel skalieren, um den sich ändernden Anforderungen gerecht zu werden. Falls die Daten lokal auf den Nodes gespeichert wurden, müssten beim Hinzufügen oder Entfernen von Nodes umfangreiche Datenmengen zwischen Nodes verschoben werden, um sie entweder von entfernten Nodes oder auf neu erstellte Nodes zu verschieben.

Das Depot

Ein Nachteil der kommunalen Lagerung ist seine Geschwindigkeit. Der Zugriff auf Daten von einem gemeinsam genutzten Cloud-Speicherort ist langsamer als das Lesen von einer lokalen Festplatte. Darüber hinaus kann die Verbindung zu kommunalem Storage zu einem Engpass werden, wenn viele Nodes die Daten gleichzeitig lesen. Zur Verbesserung der Zugriffsgeschwindigkeit für Daten führen die Knoten in einer Eon-Modus-Datenbank einen lokalen Festplatten-Cache mit Daten, die als Depot bezeichnet werden. Bei der Ausführung einer Abfrage prüfen die Knoten zunächst, ob sich die erforderlichen Daten im Depot befinden. Ist dies der Fall, wird die Abfrage beendet, indem die lokale Kopie der Daten verwendet wird. Wenn sich die Daten nicht im Depot befindet, ruft der Knoten die Daten aus dem gemeinsamen Lager ab und speichert eine Kopie im Depot.

NetApp StorageGRID-Empfehlungen

Vertica speichert Datenbankdaten im Objekt-Storage als Tausende (oder Millionen) komprimierter Objekte (die beobachtete Größe beträgt 200 bis 500 MB pro Objekt. Wenn ein Benutzer Datenbankabfragen ausführt, ruft Vertica den ausgewählten Datenbereich aus diesen komprimierten Objekten parallel mit dem GET-Aufruf des Byte-Bereichs ab. Jeder Byte-Bereich HAT ca. 8 KB.

Während des 10-TB-Datenbankdepots zur Prüfung von Benutzeranfragen wurden 4,000 bis 10,000 GET-Anforderungen (Byte-Bereich GET) pro Sekunde an das Grid gesendet. Bei diesem Test werden SG6060 Appliances eingesetzt, obwohl die CPU-Auslastung des % pro Appliance-Node niedrig ist (etwa 20 bis 30 %), warten 2/3 der CPU-Zeit auf I/O Bei SGF6024 wird ein sehr kleiner Prozentsatz (0 % bis 0.5 %) des I/O-Wartens beobachtet.

Aufgrund der hohen Anforderungen an kleine IOPS mit sehr niedrigen Latenzanforderungen (der Durchschnitt liegt bei weniger als 0.01 Sekunden) empfiehlt NetApp die Verwendung von SFG6024 für Objekt-Storage-Services. Falls das SG6060 für sehr große Datenbanken benötigt wird, sollte der Kunde mit dem Vertica Account-Team zusammenarbeiten, um den aktiv abgefragten Datensatz zu unterstützen.

Für den Admin-Node und den API-Gateway-Node kann der Kunde das SG100 oder SG1000 verwenden. Die Wahl hängt von der Anzahl der Abfragen der Benutzer in paralleler und Datenbank-Größe ab. Wenn der Kunde einen Drittanbieter-Load-Balancer einsetzen möchte, empfiehlt NetApp einen dedizierten Load Balancer für Workloads mit hohen Performance-Anforderungen. Informationen zur StorageGRID Dimensionierung erhalten Sie vom NetApp Account Team.

Weitere Empfehlungen für die StorageGRID-Konfiguration:

- **Grid-Topologie**. Kombinieren Sie SGF6024 nicht mit anderen Storage Appliance-Modellen am selben Grid-Standort. Wenn Sie den SG6060 für den langfristigen Archivierungsschutz verwenden möchten, sollten Sie den SGF6024 mit einem dedizierten Grid Load Balancer am eigenen Grid-Standort (entweder am physischen oder logischen Standort) für eine aktive Datenbank aufbewahren, um die Performance zu steigern. Die Kombination verschiedener Appliance-Modelle am selben Standort verringert die Gesamtleistung am Standort.

- **Datenschutz**. Verwenden Sie Replizieren-Kopien für die Sicherheit. Verwenden Sie kein Erasure Coding für eine aktive Datenbank. Der Kunde kann das Verfahren zur Einhaltung von Datenkonsistenz (Erasure Coding) verwenden, um inaktive Datenbanken langfristig zu schützen.
- **Gitterkompression nicht aktivieren**. Vertica komprimiert Objekte, bevor sie in Objekt-Storage gespeichert werden. Durch die Aktivierung der Grid-Komprimierung wird die Storage-Auslastung nicht weiter gesenkt und DIE GET-Performance im Byte-Bereich wird deutlich verringert.
- **HTTP im Vergleich zu HTTPS S3-Endpunktverbindung**. Während des Benchmark-Tests konnten wir bei der Verwendung einer HTTP S3-Verbindung vom Vertica Cluster zum StorageGRID Load Balancer-Endpunkt eine Performance-Steigerung von ca. 5 % feststellen. Diese Auswahl sollte auf den Sicherheitsanforderungen des Kunden basieren.

Empfehlungen für eine Vertica Konfiguration sind:

- **Die Standardeinstellungen des Vertica Datenbank-Depots sind aktiviert (Wert = 1) für Lese- und Schreibvorgänge**. NetApp empfiehlt dringend, diese Depoteinstellungen für die Performance-Steigerung zu aktivieren.
- **Streaming-Einschränkungen deaktivieren**. Weitere Informationen zur Konfiguration finden Sie im Abschnitt Deaktivieren von Streaming-Einschränkungen.

Installation von Eon-Modus vor Ort mit kommunalem Speicher auf StorageGRID

In den folgenden Abschnitten wird das Verfahren beschrieben, um den Eon-Modus vor Ort mit kommunalem Speicher auf StorageGRID zu installieren. Das Verfahren zur Konfiguration von S3-kompatiblem Objektspeicher (Simple Storage Service) ähnelt dem Verfahren im Vertica-Leitfaden. "Installation einer On-Premises-Eon-Mode-Datenbank".

Für den Funktionstest wurde folgendes Setup verwendet:

- StorageGRID 11.4.0.4
- Vertica 10.1.0
- Drei Virtual Machines (VMs) mit CentOS 7.x OS für Vertica Nodes zu einem Cluster bilden. Dieses Setup gilt nur für den Funktionstest, nicht für das Produktions-Datenbank-Cluster der Vertica.

Diese drei Nodes sind mit einem SSH-Schlüssel (Secure Shell) eingerichtet, um SSH ohne Passwort zwischen den Nodes innerhalb des Clusters zuzulassen.

Erforderliche Informationen von NetApp StorageGRID

Um den Eon-Modus vor Ort mit kommunalem Speicher auf StorageGRID zu installieren, müssen Sie die folgenden Vorbedingung-Informationen haben.

- IP-Adresse oder vollständig qualifizierter Domain-Name (FQDN) und Portnummer des StorageGRID S3-Endpunkts. Wenn Sie HTTPS verwenden, verwenden Sie eine CA (Custom Certificate Authority) oder ein selbstsigniertes SSL-Zertifikat, das am StorageGRID S3-Endpunkt implementiert wurde.
- Bucket-Name Er muss vorexistieren und leer sein.
- Schlüssel-ID und geheimer Zugriffsschlüssel mit Lese- und Schreibzugriff auf den Bucket

Erstellen einer Autorisierungsdatei für den Zugriff auf den S3-Endpunkt

Beim Erstellen einer Autorisierungsdatei für den Zugriff auf den S3-Endpunkt gelten die folgenden Voraussetzungen:

- Vertica ist installiert.
- Ein Cluster ist für die Datenbankerstellung eingerichtet, konfiguriert und bereit.

So erstellen Sie eine Autorisierungsdatei für den Zugriff auf den S3-Endpunkt:

1. Melden Sie sich beim Vertica-Knoten an, auf dem Sie ausgeführt werden `admintools` So erstellen Sie die Eon-Modus-Datenbank.

   Der Standardbenutzer ist `dbadmin`, Erstellt während der Vertica Cluster Installation.

2. Verwenden Sie einen Texteditor, um eine Datei unter dem zu erstellen `/home/dbadmin` Verzeichnis. Der Dateiname kann alles sein, was Sie wollen, z. B. `sg_auth.conf`.
3. Wenn der S3-Endpunkt einen Standard-HTTP-Port 80 oder HTTPS-Port 443 verwendet, überspringen Sie die Portnummer. Um HTTPS zu verwenden, legen Sie die folgenden Werte fest:

   ◦ `awsenablehttps = 1`, Sonst setzen Sie den Wert auf `0`.

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

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

   Um eine benutzerdefinierte CA oder ein selbstsigniertes SSL-Zertifikat für die HTTPS-Verbindung des StorageGRID S3-Endpunkts zu verwenden, geben Sie den vollständigen Dateipfad und den Dateinamen des Zertifikats an. Diese Datei muss sich am selben Speicherort auf jedem Vertica-Knoten befinden und über Leseberechtigung für alle Benutzer verfügen. Überspringen Sie diesen Schritt, wenn das StorageGRID S3 Endpoint SSL-Zertifikat von einer öffentlich bekannten CA signiert wurde.

   - `awscafile = <filepath/filename>`

   Informationen hierzu finden Sie beispielsweise in der folgenden Beispieldatei:

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

+

> In einer Produktionsumgebung muss der Kunde ein Serverzertifikat implementieren, das von einer öffentlich bekannten CA auf einem StorageGRID S3 Load Balancer-Endpunkt unterzeichnet wurde.

Auswählen eines Depotpfads auf allen Vertica-Knoten

Wählen Sie auf jedem Knoten ein Verzeichnis für den Depot-Speicherpfad aus oder erstellen Sie ein Verzeichnis. Das Verzeichnis, das Sie für den Parameter Depot-Speicherpfad bereitstellen, muss Folgendes haben:

- Derselbe Pfad auf allen Nodes im Cluster (z. B. `/home/dbadmin/depot`)
- Vom dbadmin-Benutzer lesbar und beschreibbar sein

- Ausreichende Lagerung

  Standardmäßig verwendet Vertica 60 % des Dateisystemspeichers, der das Verzeichnis für die Depotspeicherung enthält. Sie können die Größe des Depots mithilfe der begrenzen `--depot-size` Argument in `create_db` Befehl. Siehe "Dimensionierung des Vertica Clusters für eine Eon-Mode-Datenbank" Artikel für allgemeine Vertica Größenrichtlinien oder wenden Sie sich an Ihren Vertica Account Manager.

  Der `admintools create_db` Das Tool versucht, den Depotpfad für Sie zu erstellen, wenn dieser nicht vorhanden ist.

Erstellen der On-Premises-Datenbank von Eon

So erstellen Sie die On-Premises-Datenbank von Eon:

1. Verwenden Sie zum Erstellen der Datenbank die `admintools create_db` Werkzeug.

   Die folgende Liste enthält eine kurze Erläuterung der Argumente, die in diesem Beispiel verwendet werden. Eine detaillierte Erläuterung aller erforderlichen und optionalen Argumente finden Sie im Dokument Vertica.

   - -X <Pfad/Dateiname der in erstellten Autorisierungsdatei „Erstellen einer Autorisierungsdatei für den Zugriff auf den S3-Endpunkt“ >.

     Die Autorisierungsdetails werden nach erfolgreicher Erstellung in der Datenbank gespeichert. Sie können diese Datei entfernen, um zu vermeiden, dass der S3-Geheimschlüssel offengelegt wird.

   - --communal-Storage-location <s3://storagegrid buchname>
   - -S <kommagetrennte Liste der Vertica-Knoten, die für diese Datenbank verwendet werden sollen>
   - -D <Name der zu erstellenden Datenbank>
   - -P <Kennwort für diese neue Datenbank> festlegen. Den folgenden Beispielbefehl können Sie z. B. einsehen:

     ```
     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>'
     ```

     Das Erstellen einer neuen Datenbank dauert abhängig von der Anzahl der Nodes für die Datenbank mehrere Minuten. Wenn Sie die Datenbank zum ersten Mal erstellen, werden Sie aufgefordert, die Lizenzvereinbarung zu akzeptieren.

Informationen hierzu finden Sie z. B. in der folgenden Beispielautorisierungsdatei und `create db` Befehl:

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

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

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

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

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
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

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
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

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
33	s3://vertica/acd/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29_0_0.dfs
133	s3://vertica/b98/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a4d_0_0.dfs
38	s3://vertica/db3/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a49_0_0.dfs
38	s3://vertica/eba/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a59_0_0.dfs
21521920	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000215e2.tar
6865408	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021602/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021602.tar
204217344	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021610/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021610.tar
16109056	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000217e0.tar
12853248	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000021800/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021800.tar

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
8937984	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a.tar
56260608	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000218b2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2.tar
53947904	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000219ba/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba.tar
44932608	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000000219de/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de.tar
256306688	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar
8062464	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar
20024832	s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.tar
10444	s3://vertica/metadata/VMart/cluster_config.json
823266	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/chkpt_1.cat.gz

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
254	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/completed
2958	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/chkpt_1.cat.gz
231	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c2_2/completed
822521	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/chkpt_1.cat.gz
231	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c4_4/completed
746513	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g14.cat
2596	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_3_g3.cat.gz
821065	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_4_g4.cat.gz
6440	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_5_g5.cat
8518	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_8_g8.cat

Objektgröße (Byte)	Bucket/Objektschlüssel vollständiger Pfad
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat
822922	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz
232	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed
822930	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz
755033	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat
0	s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat

Deaktivieren von Streaming-Einschränkungen

Dieses Verfahren basiert auf dem Vertica-Leitfaden für andere On-Premises-Objektspeicher und sollte für StorageGRID angewendet werden.

1. Deaktivieren Sie nach dem Erstellen der Datenbank das `AWSStreamingConnectionPercentage` Konfigurationsparameter durch Festlegen auf `0`. Diese Einstellung ist für eine On-Premises-Installation im Eon-Modus mit kommunalem Speicher nicht erforderlich. Dieser Konfigurationsparameter steuert die Anzahl der Verbindungen zu dem Objektspeicher, den Vertica für das Streaming von Lesevorgängen verwendet. In einer Cloud-Umgebung verhindert diese Einstellung, dass aus dem Objektspeicher Daten gestreamt werden, alle verfügbaren Datei-Handles nutzen. Einige Datei-Handles stehen für andere Objektspeichervorgänge zur Verfügung. Aufgrund der niedrigen Latenz von On-Premises-Objektspeichern ist diese Option nicht erforderlich.
2. Verwenden Sie A `vsql` Anweisung zum Aktualisieren des Parameterwerts. Das Passwort ist das Datenbank-Passwort, das Sie unter „Erstellen der On-Premises-Datenbank von Eon“ festgelegt haben. Informationen hierzu finden Sie z. B. in der folgenden Beispielausgabe:

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

Depot-Einstellungen werden überprüft

Standarddepot-Einstellungen der Vertica-Datenbank sind aktiviert (Wert = 1) für Lese- und Schreibvorgänge. NetApp empfiehlt dringend, diese Depoteinstellungen für die Performance-Steigerung zu aktivieren.

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

Laden von Probendaten (optional)

Wenn diese Datenbank zu Testzwecken bereit ist und entfernt werden wird, können Sie Beispieldaten zu Testzwecken in diese Datenbank laden. Vertica kommt mit Probendatensatz, VMart, gefunden unter `/opt/vertica/examples/VMart_Schema/` Auf jedem Vertica-Knoten. Weitere Informationen zu diesem Beispieldatensatz finden Sie hier "Hier".

Führen Sie die folgenden Schritte aus, um die Probendaten zu laden:

1. Melden Sie sich als dbadmin an einem der Vertica-Knoten an: cd /opt/vertica/examples/VMart_Schema/
2. Laden Sie Beispieldaten in die Datenbank, und geben Sie das Datenbank-Passwort ein, wenn Sie in den Unterschritten c und d aufgefordert werden:

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

b. `./vmart_gen`

c. `vsql < vmart_define_schema.sql`

d. `vsql < vmart_load_data.sql`

3. Es gibt mehrere vordefinierte SQL-Abfragen. Sie können einige davon ausführen, um zu bestätigen, dass die Testdaten erfolgreich in die Datenbank geladen wurden. Beispiel: `vsql < vmart_queries1.sql`

Wo Sie weitere Informationen finden

Sehen Sie sich die folgenden Dokumente und/oder Websites an, um mehr über die in diesem Dokument beschriebenen Informationen zu erfahren:

- "NetApp StorageGRID 11.7 Produktdokumentation"
- "Datenblatt zu StorageGRID"
- "Vertica 10.1 Produktdokumentation"

Versionsverlauf

Version	Datum	Versionsverlauf des Dokuments
Version 1.0	September 2021	Erste Version.

Von Angela Cheng

StorageGRID-Protokollanalyse mit ELK-Stack

Mit der Funktion StorageGRID 11.6 syslog Forward können Sie einen externen Syslog-Server konfigurieren, um StorageGRID-Protokollmeldungen zu erfassen und zu analysieren. ELK (Elasticsearch, Logstash, Kibana) hat sich zu einer der beliebtesten Log-Analytics-Lösungen entwickelt. Sehen Sie sich an "StorageGRID-Protokollanalyse mit ELK-Video" Zeigt eine ELK-Beispielkonfiguration an und wie sie verwendet werden kann, um fehlerhafte S3-Anforderungen zu identifizieren und zu beheben. Dieser Artikel enthält Beispieldateien der Logstash-Konfiguration, Kibana-Abfragen, Diagramme und Dashboard, die Ihnen einen schnellen Einstieg in die StorageGRID-Protokollverwaltung und -Analyse ermöglichen.

Anforderungen

- StorageGRID 11.6.0.2 oder höher
- ELK (Elasticsearch, Logstash und Kibana) 7.1x oder höher installiert und in Betrieb

Beispieldateien

- "Laden Sie das Paket Logstash 7.x Beispieldateien herunter" + **md5 Prüfsumme** 148c23d0021d9a4bb4a6c0287464deab + **sha256 Prüfsumme** f51ec9e2e3f842d5a786156b167a561beb4373038b4e7bb3c8be3d522adf2d6

- "Laden Sie das Paket Logstash 8.x Beispieldateien herunter" + **md5 Prüfsumme** e11bae3a662f87c310ef363d0fe06835 + **sha256 Prüfsumme** 5c670755742cfd5aa723a596ba087e0153a65bcae3934afddddddddd68d

Annahme

Leser kennen die Terminologie und den Betrieb von StorageGRID und ELK.

Anweisung

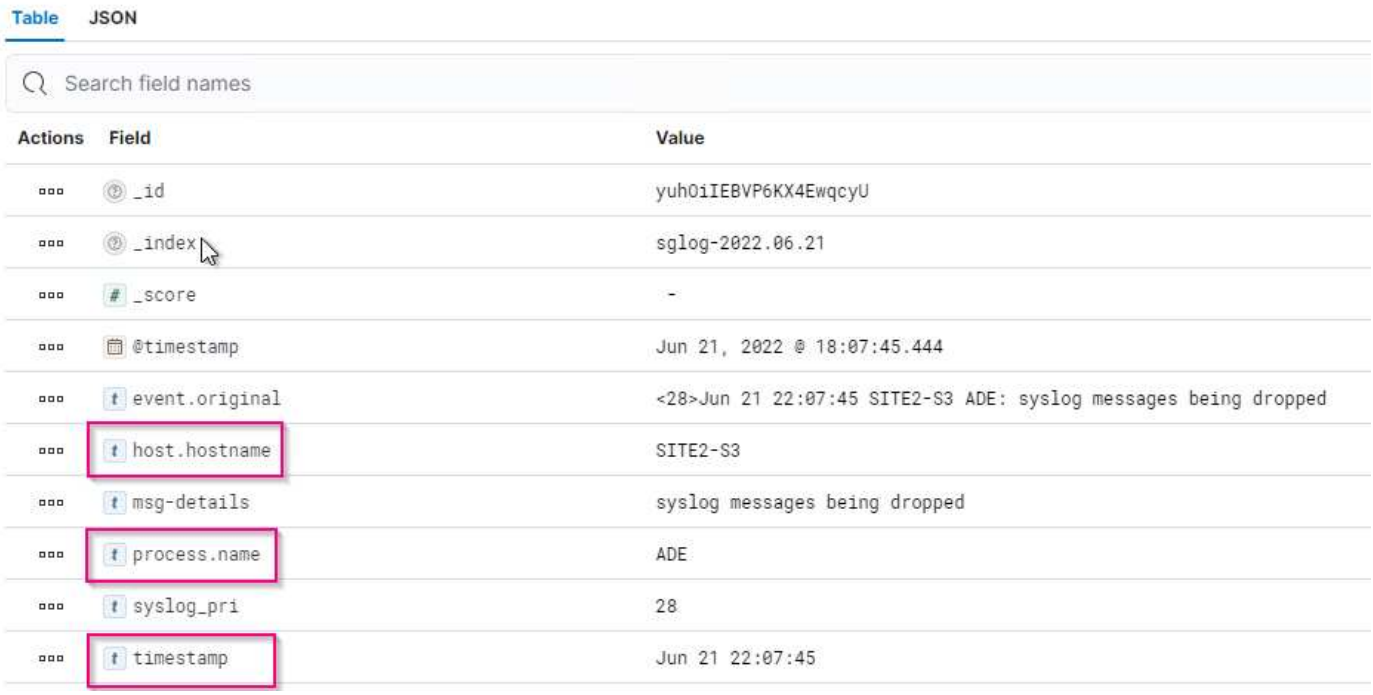
Zwei Beispielversionen werden aufgrund von Unterschieden in Namen bereitgestellt, die durch grok-Muster definiert wurden. + zum Beispiel definiert das SYSLOGBASE-grok-Muster in der Logstash config-Datei Feldnamen je nach installierter Logstash-Version unterschiedlich.

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

Logstash 7.17 Beispiel

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

Logstash 8.23 Beispiel



Schritte

1. Entpacken Sie das angegebene Muster anhand Ihrer installierten ELK-Version. + der Beispielordner enthält zwei Logstash-Konfigurationsbeispiele: + **sglog-2-file.conf:** Diese Konfigurationsdatei gibt StorageGRID-Protokollnachrichten ohne Datentransformation in eine Datei auf Logstash aus. Sie können auf diese Weise bestätigen, dass Logstash StorageGRID Nachrichten empfangen oder StorageGRID-Protokollmuster verstehen. + **sglog-2-es.conf:** Diese Konfigurationsdatei wandelt StorageGRID-Protokollmeldungen mithilfe verschiedener Muster und Filter um. Dazu gehören beispielsweise Drop-Statements, die Meldungen basierend auf Mustern oder Filtern ablegen. Die Ausgabe wird zur Indizierung an Elasticsearch gesendet. + Passen Sie die ausgewählte Konfigurationsdatei entsprechend der Anweisung in der Datei an.
2. Testen Sie die benutzerdefinierte Konfigurationsdatei:

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

   Wenn die letzte zurückgegebene Zeile der unten angegebenen Zeile ähnelt, weist die Konfigurationsdatei keine Syntaxfehler auf:

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

3. Benutzerdefinierte conf-Datei auf den Logstash-Server kopieren.config: /Etc/logstash/conf.d + Wenn Sie config.reload.automatic in /etc/logstash/logstash.yml nicht aktiviert haben, starten Sie den Logstash-Dienst neu. Andernfalls warten Sie, bis das Neueintervall der Konfiguration abgelaufen ist.

```
 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
```

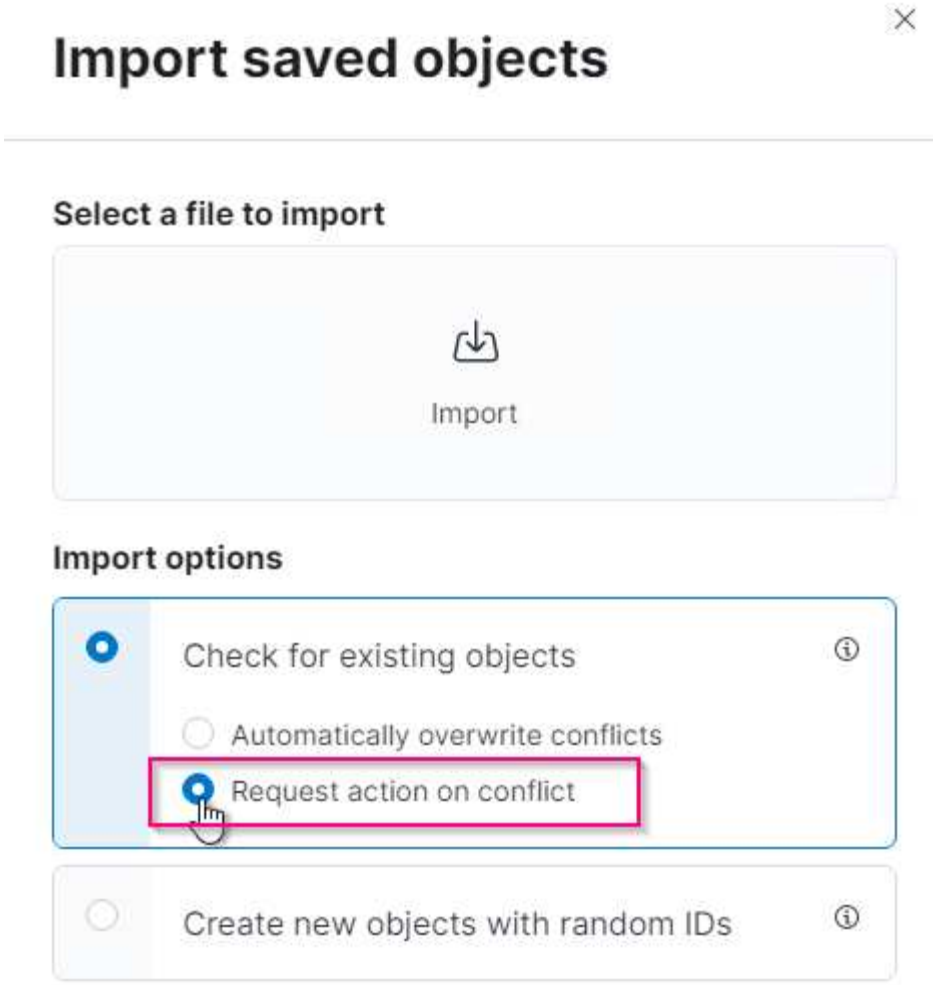
4. Prüfen Sie /var/log/logstash/logstash-plain.log und vergewissern Sie sich, dass beim Starten von Logstash mit der neuen Konfigurationsdatei keine Fehler auftreten.
5. Bestätigen Sie, dass der TCP-Port gestartet wurde und Sie zuhören. + in diesem Beispiel wird der TCP-Port 5000 verwendet.

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

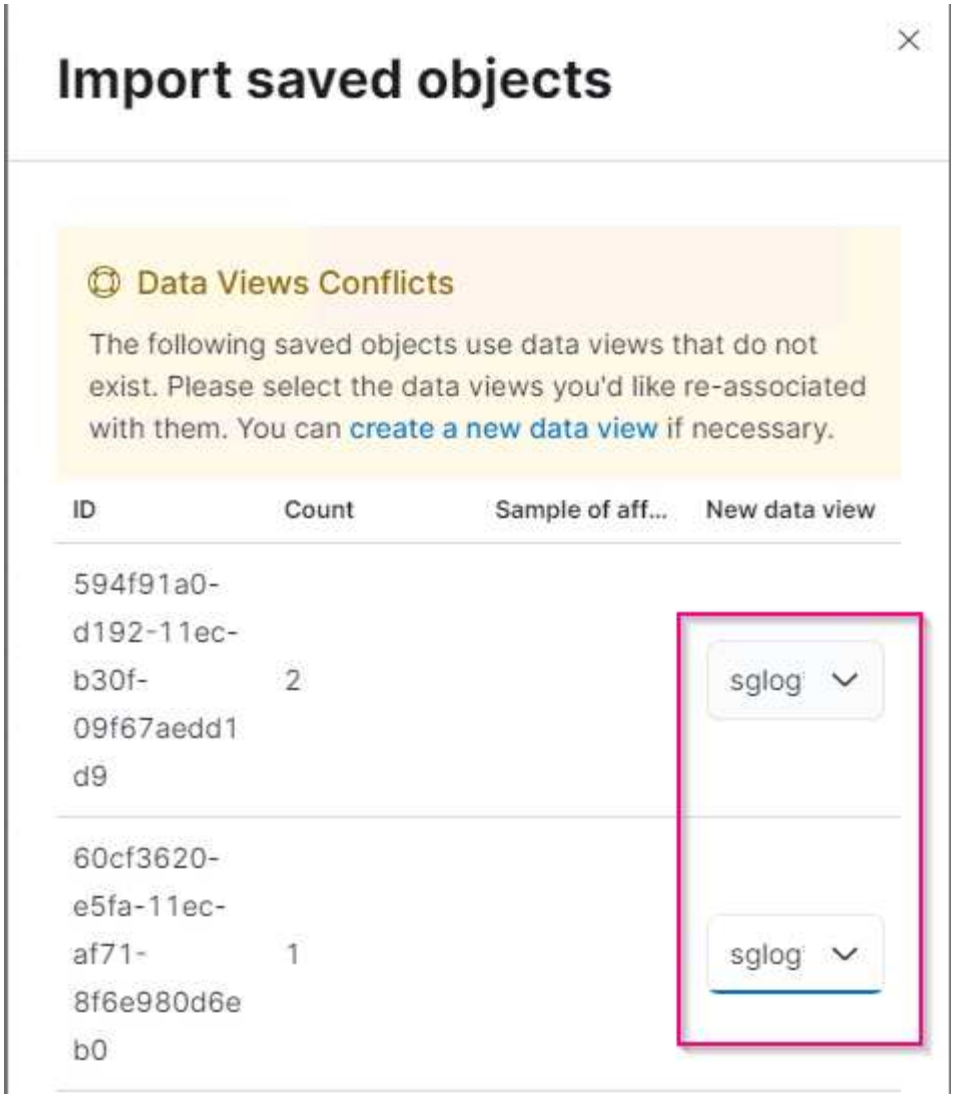
6. Konfigurieren Sie über die StorageGRID Manager-GUI einen externen Syslog-Server, um Protokollmeldungen an Logstash zu senden. Siehe "Demovideo" Entsprechende Details.
7. Sie müssen die Firewall auf dem Logstash-Server konfigurieren oder deaktivieren, damit StorageGRID-Knoten eine Verbindung zum definierten TCP-Port herstellen können.
8. Wählen Sie in der Kibana GUI die Option Management → Dev Tools. Führen Sie auf der Konsolenseite diesen BEFEHL GET aus, um zu bestätigen, dass neue Indizes auf Elasticsearch erstellt werden.

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

9. Erstellen Sie in der Kibana GUI Indexmuster (ELK 7.x) oder Datenansicht (ELK 8.x).
10. Geben Sie in der Kibana GUI in das Suchfeld, das sich in der oberen Mitte befindet, „abgetackte Objekte“ ein. + Wählen Sie auf der Seite gespeicherte Objekte die Option Importieren. Wählen Sie unter „Importoptionen“ die Option „Aktion für Konflikt anfordern“ aus.



Importieren Sie elk<Version>-query-Chart-sample.ndjson. + Wählen Sie bei Aufforderung zur Lösung des Konflikts das in Schritt 8 erstellte Indexmuster oder die Datenansicht aus.



Die folgenden Kibana-Objekte werden importiert: + **Query** + * Audit-msg-s3rq-orlm + * bycast Log s3 bezogene Nachrichten + * Loglevel Warnung oder höher + * fehlgeschlagenes Sicherheitsereignis + **Diagramm** + * s3 Anfragen zählen auf Basis von bycast.log + * HTTP-Statuscode + * Audit msg Aufschlüsselung nach Typ + * durchschnittliche s3-Antwort Time + **Dashboard** + * S3-Request-Dashboard mit den oben genannten Diagrammen.

Sie können nun die StorageGRID-Protokollanalyse mit Kibana durchführen.

Weitere Ressourcen

- "Syslog101"
- "Was ist der ELK-Stack"
- "Liste der Tülenmuster"
- "Ein Anfängerführer zum Logstash: Grok"
- "Eine praktische Anleitung zum Logstash: Syslog Deep Dive"
- "Kibana Guide – Erkunden Sie das Dokument"
- "Referenz für StorageGRID-Prüfprotokolle"

Von Angela Cheng

Mit Prometheus und Grafana können Sie die Aufbewahrung Ihrer Kennzahlen erweitern

Dieser technische Bericht enthält detaillierte Anweisungen zur Konfiguration von NetApp StorageGRID 11.6 mit externen Services Prometheus und Grafana.

Einführung

StorageGRID speichert Kennzahlen mithilfe von Prometheus und visualisiert diese Kennzahlen über integrierte Grafana Dashboards. Die Kennzahlen von Prometheus können über StorageGRID sicher abgerufen werden, indem Client-Zugriffszertifikate konfiguriert und prometheus-Zugriff für den angegebenen Client ermöglicht wird. Derzeit wird die Aufbewahrung dieser metrischen Daten durch die Storage-Kapazität des Administrations-Nodes begrenzt. Um eine längere Dauer zu erreichen und individuelle Visualisierungen dieser Kennzahlen zu erstellen, werden wir einen neuen Prometheus- und Grafana-Server einsetzen, unseren neuen Server für die Scrape der Kennzahlen aus der StorageGRIDs-Instanz konfigurieren und ein Dashboard mit den für uns wichtigen Kennzahlen erstellen. Weitere Informationen zu den in der erfassten Prometheus-Kennzahlen finden Sie unter "StorageGRID-Dokumentation".

Föderate Prometheus

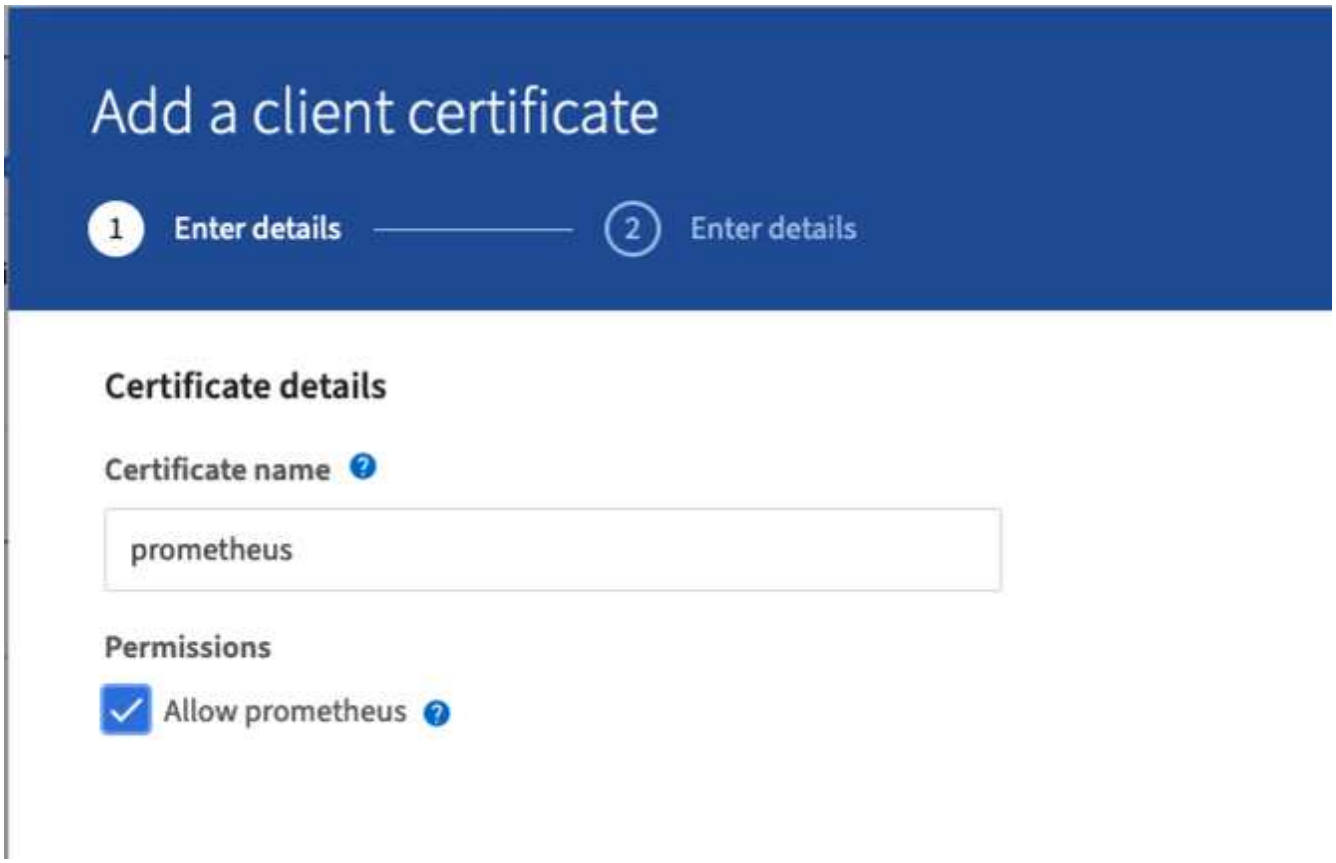
Labordetails

Für die Zwecke dieses Beispiels werde ich alle virtuellen Maschinen für StorageGRID 11.6 Knoten und einen Debian 11-Server verwenden. Die StorageGRID-Managementoberfläche ist mit einem öffentlich vertrauenswürdigen CA-Zertifikat konfiguriert. Dieses Beispiel wird die Installation und Konfiguration des StorageGRID-Systems oder der Debian linux-Installation nicht durchlaufen. Sie können jeden gewünschten Linux-Geschmack verwenden, der von Prometheus und Grafana unterstützt wird. Sowohl Prometheus als auch Grafana können als Docker-Container installiert, aus der Quelle erstellt oder vorkompilierte Binärdateien erstellt werden. In diesem Beispiel werde ich sowohl Prometheus- als auch Grafana-Binärdateien direkt auf dem gleichen Debian-Server installieren. Laden Sie sich die grundlegenden Installationsanweisungen von herunter https://prometheus.io Und https://grafana.com/grafana/ Jeweils.

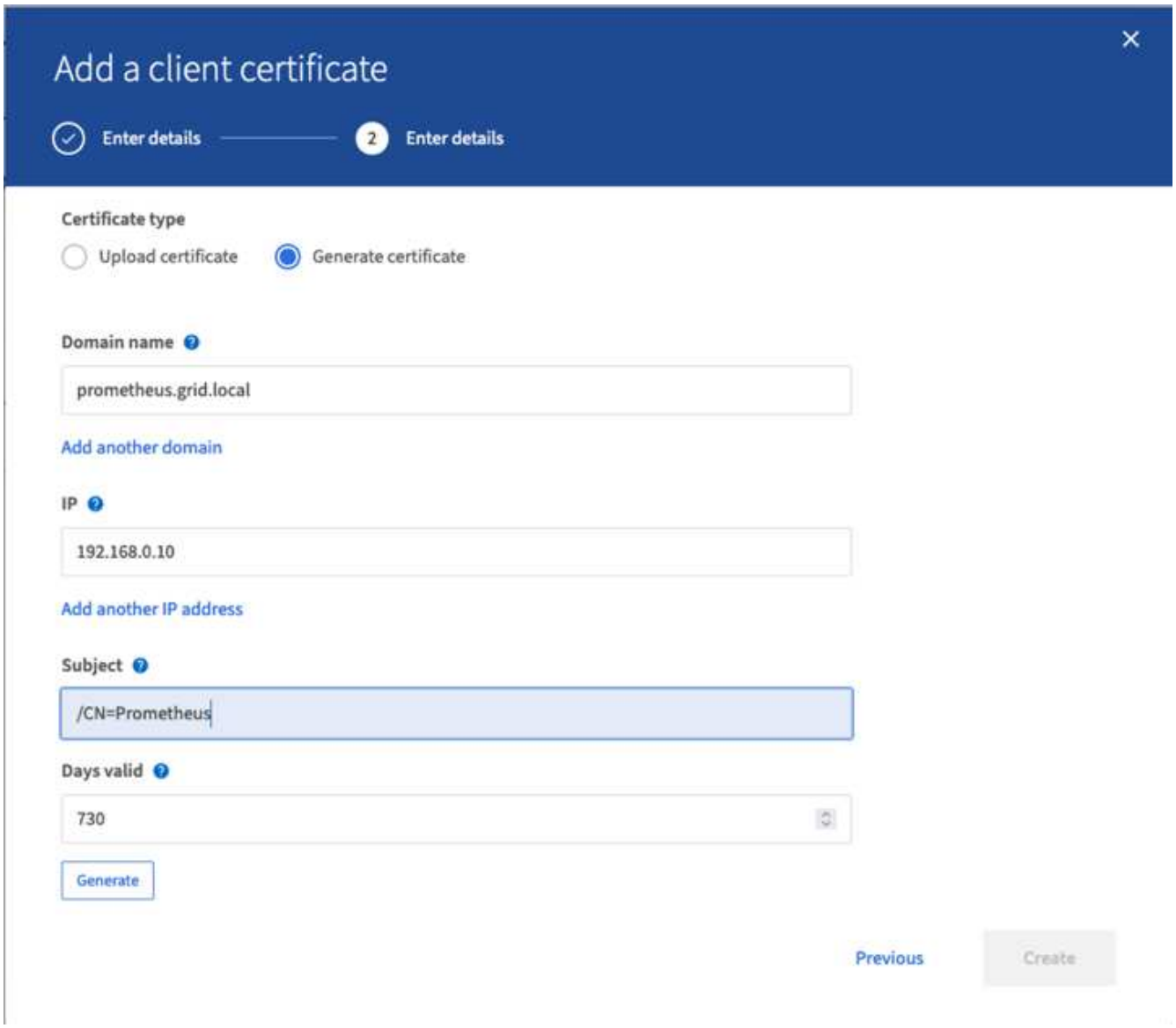
Konfigurieren Sie StorageGRID für Prometheus Client-Zugriff

Um Zugriff auf gespeicherte prometheus-Kennzahlen zu StorageGRIDs zu erhalten, müssen Sie ein Clientzertifikat mit privatem Schlüssel generieren oder hochladen und die Berechtigung für den Client aktivieren. Die StorageGRID-Schnittstelle muss ein SSL-Zertifikat haben. Dieses Zertifikat muss vom prometheus-Server entweder von einer vertrauenswürdigen CA oder manuell vertrauenswürdig sein, wenn es selbst signiert ist. Weitere Informationen finden Sie auf der "StorageGRID-Dokumentation".

1. Wählen Sie in der StorageGRID-Managementoberfläche unten links die Option „KONFIGURATION“ und klicken Sie in der zweiten Spalte unter „Sicherheit“ auf Zertifikate.
2. Wählen Sie auf der Seite Zertifikate die Registerkarte „Client“ aus und klicken Sie auf die Schaltfläche „Add“.
3. Geben Sie einen Namen für den Client an, dem Zugriff gewährt wird, und verwenden Sie dieses Zertifikat. Klicken Sie auf das Feld unter „Berechtigungen“ vor „Prometheus zulassen“ und klicken Sie auf die Schaltfläche „Weiter“.

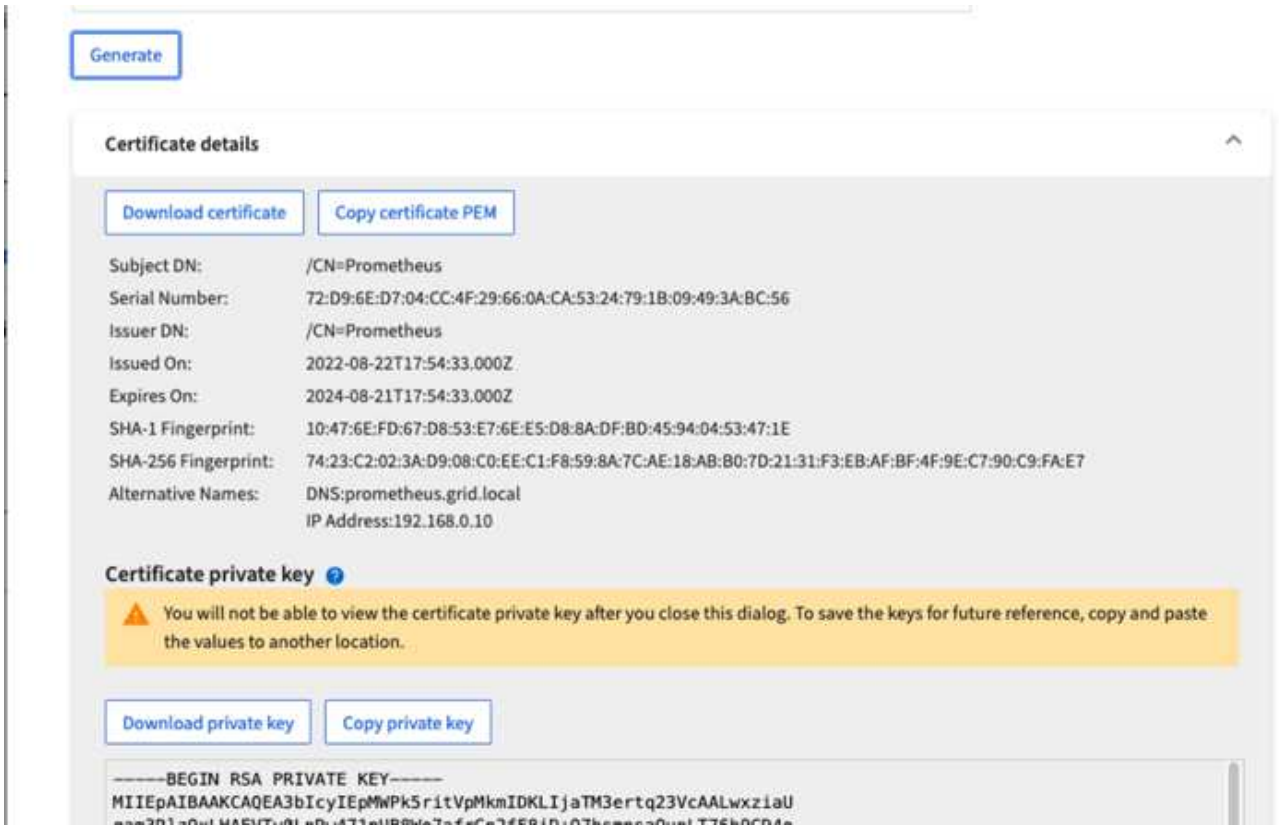


4. Wenn Sie ein CA-signiertes Zertifikat haben, können Sie das Optionsfeld für "Zertifikat hochladen" wählen, aber in unserem Fall werden wir StorageGRID das Client-Zertifikat generieren lassen, indem Sie das Optionsfeld für "Zertifikat generieren". Die Pflichtfelder werden angezeigt, in die ausgefüllt werden soll. Geben Sie den FQDN für den Client-Server, die IP des Servers, den Betreff und die gültigen Tage ein. Dann klicken Sie auf die Schaltfläche „Erzeugen“.



> 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. Laden Sie die Pem-Datei des Zertifikats und die Pem-Datei des privaten Schlüssels herunter.



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

Bereiten Sie den Linux-Server für die Prometheus-Installation vor

Vor der Installation von Prometheus möchte ich meine Umgebung mit einem Prometheus-Benutzer, der Verzeichnisstruktur vorbereiten und die Kapazität für den Speicherort der Kennzahlen konfigurieren.

1. Erstellen Sie den Prometheus-Benutzer.

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

2. Erstellen Sie die Verzeichnisse für Prometheus, Clientzertifikat und Kennzahlendaten.

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

3. Ich formatierte die Festplatte, die ich für die Aufbewahrung der Kennzahlen mit einem ext4 Dateisystem verwende.

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

4. Ich montierte dann das Dateisystem in das Prometheus-Kennzahlenverzeichnis.

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

5. Holen Sie die UUID der Festplatte, die Sie für Ihre Kennzahlendaten verwenden.

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

6. Hinzufügen eines Eintrags in /etc/fstab/ das Hinzufügen des Mount bei Neustarts mit der UUID von /dev/sdb.

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

Installation und Konfiguration von Prometheus

Nachdem der Server nun bereit ist, kann ich die Prometheus-Installation starten und den Service konfigurieren.

1. Extrahieren Sie das Prometheus Installationspaket

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

2. Kopieren Sie die Binärdateien in /usr/local/bin, und ändern Sie das Eigentumsrecht in den zuvor erstellten prometheus-Benutzer

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

3. Kopieren Sie die Konsolen und Bibliotheken auf /etc/prometheus

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

4. Kopieren Sie das Clientzertifikat und die pem-Dateien mit privaten Schlüsseln, die zuvor von StorageGRID heruntergeladen wurden, in /etc/prometheus/certs
5. Erstellen Sie die yaml-Konfigurationsdatei für prometheus

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

6. Geben Sie die folgende Konfiguration ein. Der Jobname kann alles sein, was Sie wünschen. Ändern Sie die „-Targets: ['']“ in den FQDN des Admin-Knotens. Wenn Sie die Namen des Zertifikats und der Dateinamen des privaten Schlüssels geändert haben, aktualisieren Sie bitte den abschnitt tls_config, um mit dem Eintrag übereinstimmen. Speichern Sie anschließend die Datei. Wenn Ihre Grid-Management-Schnittstelle ein selbstsigniertes Zertifikat verwendet, laden Sie das Zertifikat herunter und legen Sie es mit dem Clientzertifikat mit einem eindeutigen Namen ab, und fügen Sie im Abschnitt tls_config Ca_file: /Etc/prometheus/cert/UICert.pem hinzu

   a. In diesem Beispiel sammle ich alle Kennzahlen, die mit alertmanager, cassandra, Node und StorageGRID beginnen. Weitere Informationen zu den Prometheus-Kennzahlen finden Sie im "StorageGRID-Dokumentation".

```
# 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']
```

Wenn Ihre Grid-Managementoberfläche ein selbstsigniertes Zertifikat verwendet, laden Sie das Zertifikat herunter, und legen Sie es mit dem Clientzertifikat mit einem eindeutigen Namen ab. Fügen Sie im Abschnitt tls_config das Zertifikat über dem Clientzertifikat und den privaten Schlüsselzeilen hinzu

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

1. Ändern Sie das Eigentum aller Dateien und Verzeichnisse in /etc/prometheus und /var/lib/prometheus in den prometheus-Benutzer

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

2. Erstellen Sie eine prometheus-Servicedatei in /etc/systemd/System

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

3. Fügen Sie die folgenden Zeilen ein, beachten Sie die --Storage.tsdb.Retention.time=1y, welche die Aufbewahrung der metrischen Daten auf 1 Jahr festlegt. Alternativ können Sie zur Basis-Aufbewahrung auf Storage-Beschränkungen --Storage.tsdb.Retention.size=300gib verwenden. Dies ist der einzige Speicherort, der die Aufbewahrung von Kennzahlen vornimmt.

```
[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. Laden Sie den systemd-Dienst erneut, um den neuen prometheus-Service zu registrieren. Dann starten und aktivieren sie den prometheus Service.

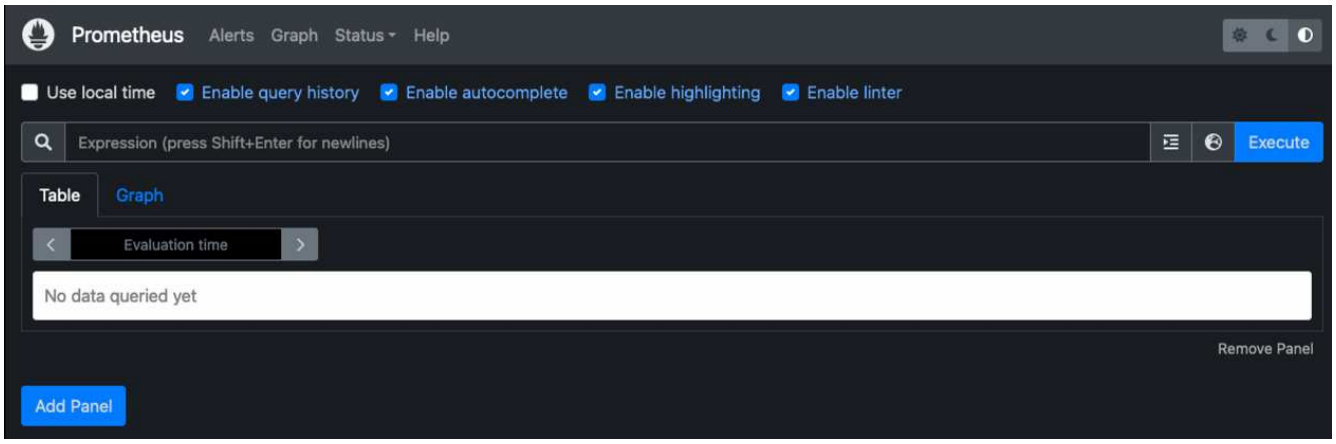
```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Überprüfen Sie, ob der Service ordnungsgemäß läuft

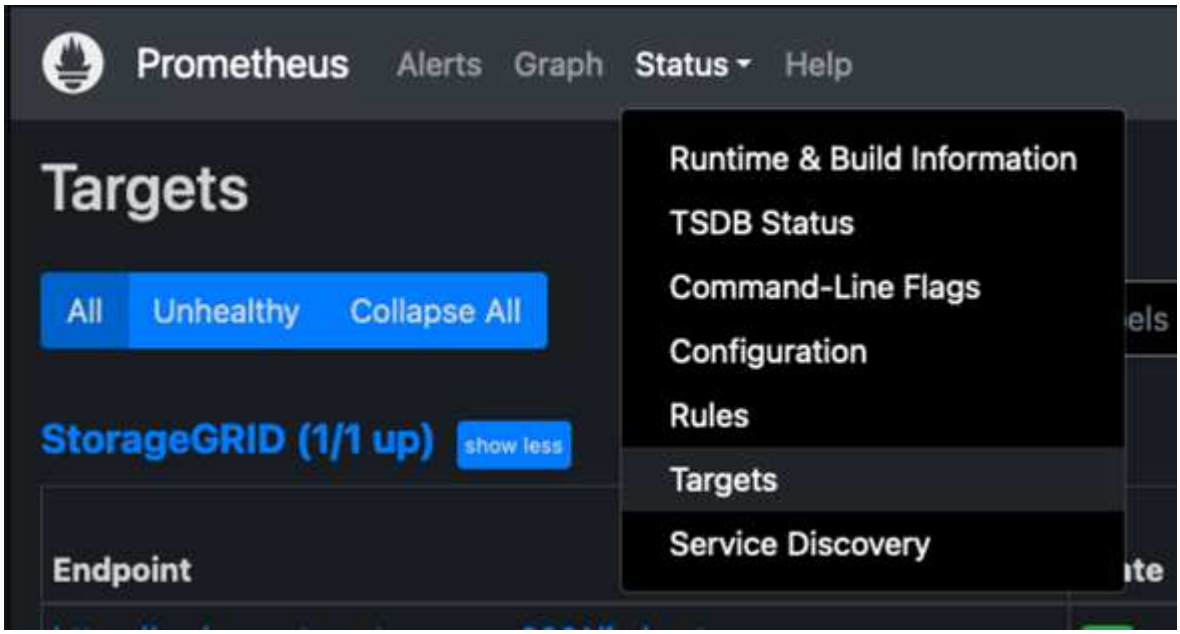
```
sudo systemctl status prometheus
```

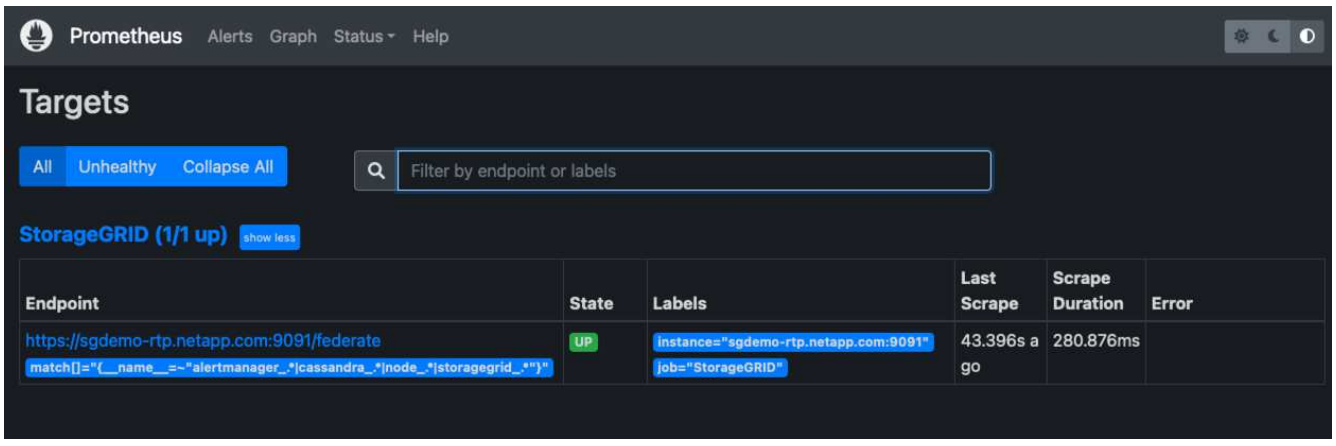
```
● prometheus.service - Prometheus Time Series Collection and Processing
Server
     Loaded: loaded (/etc/systemd/system/prometheus.service; enabled;
vendor preset: enabled)
     Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago
   Main PID: 6498 (prometheus)
      Tasks: 13 (limit: 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. Sie sollten nun in der Lage sein, auf die Benutzeroberfläche Ihres prometheus-Servers zu navigieren http://Prometheus-server:9090 Und siehe UI

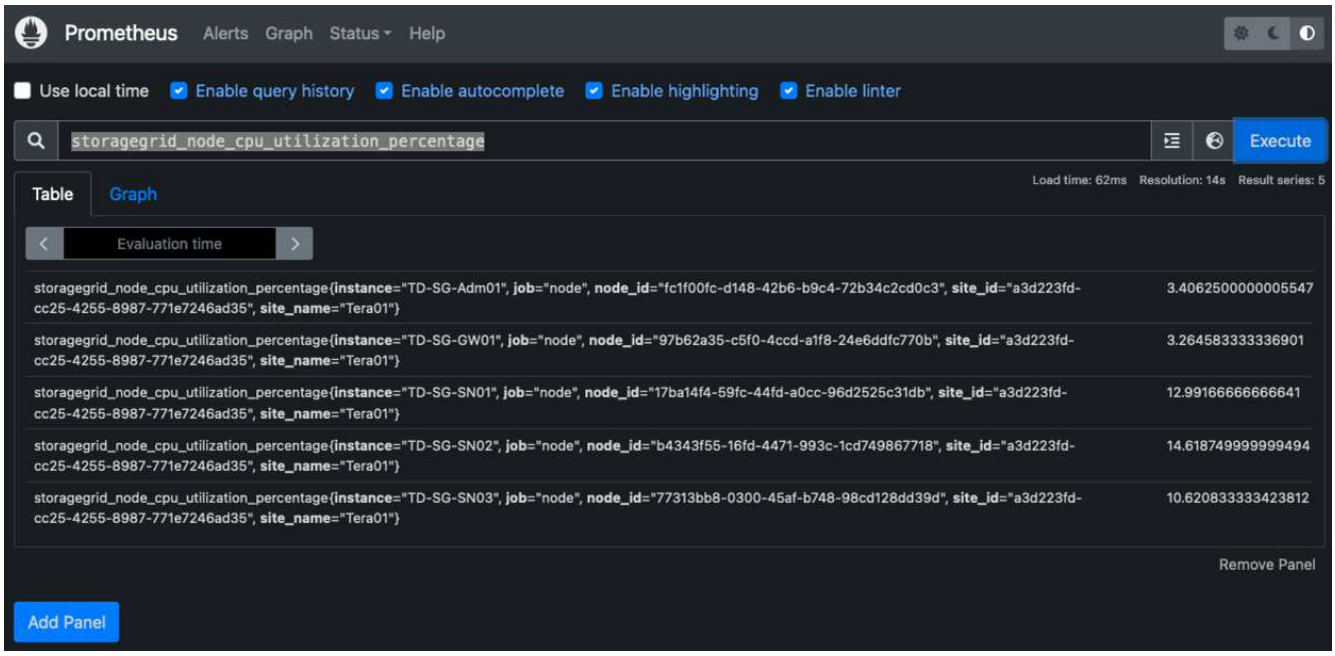


7. Unter "Status" Targets sehen Sie den Status des StorageGRID Endpunkts, den wir in prometheus.yml konfiguriert haben





8. Auf der Seite Diagramm können Sie eine Testabfrage ausführen und überprüfen, ob die Daten erfolgreich abgefangen wurden. Geben Sie beispielsweise „storagegrid_Node_cpu_Utiltiy_percenty“ in die Abfrageleiste ein und klicken Sie auf die Schaltfläche Ausführen.



Installation und Konfiguration von Grafana

Nach der Installation und dem Betrieb von prometheus können wir nun zur Installation von Grafana und zur Konfiguration eines Dashboards wechseln

Grafana-Instalation

1. Installieren Sie die neueste Enterprise Edition von 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. Dieses Repository für stabile Versionen hinzufügen:

```
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. Nachdem Sie das Repository hinzugefügt haben.

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

4. Laden Sie den systemd-Dienst neu, um den neuen grafana-Dienst zu registrieren. Starten und aktivieren Sie dann den Grafana-Service.

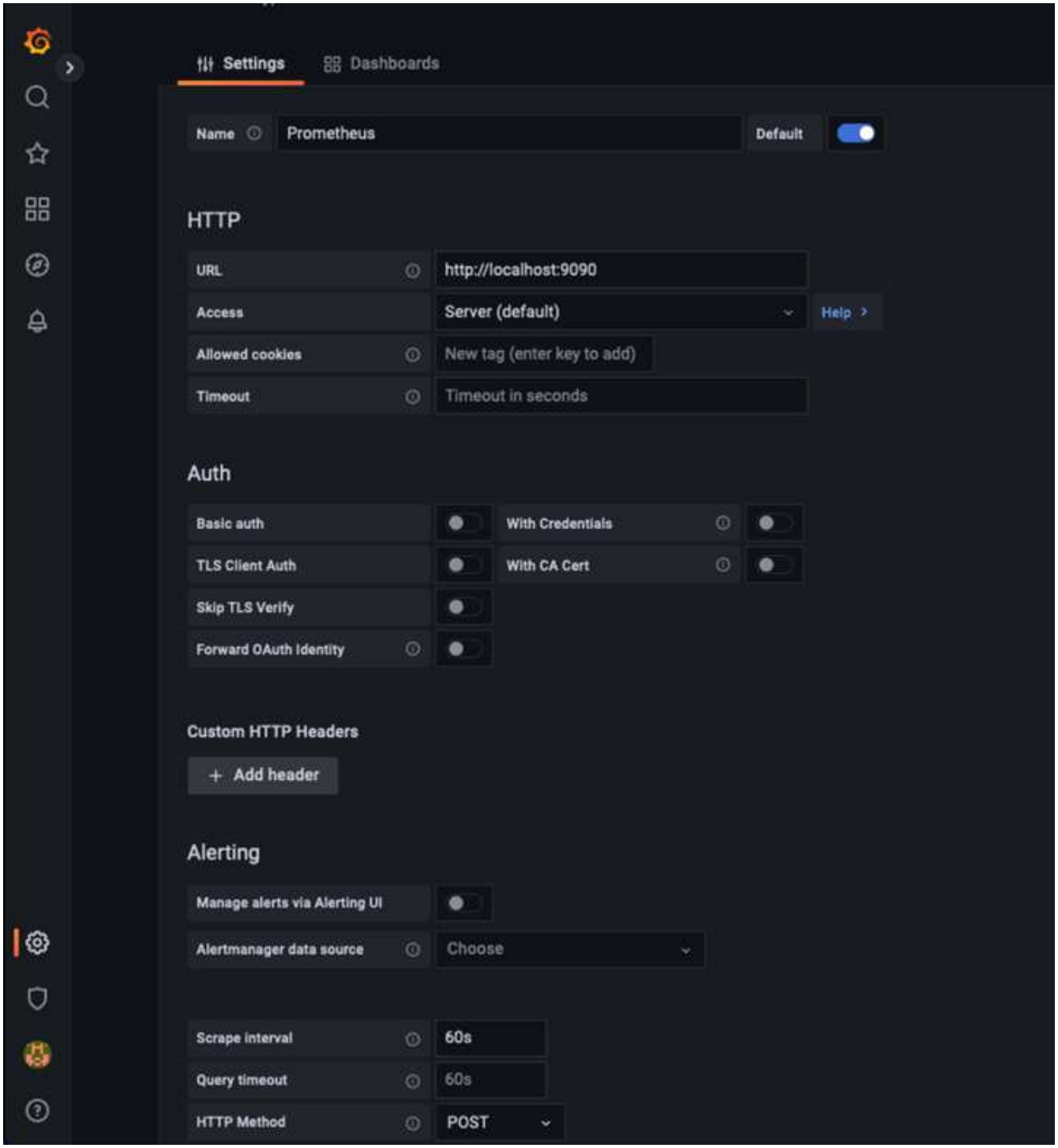
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. Grafana wird jetzt installiert und ausgeführt. Wenn Sie einen Browser zu HTTP://Prometheus-Server:3000 öffnen, werden Sie mit der Grafana-Anmeldeseite begrüßt.
6. Die Standard-Anmeldeinformationen sind admin/admin. Sie sollten ein neues Passwort festlegen, wenn Sie dazu aufgefordert werden.

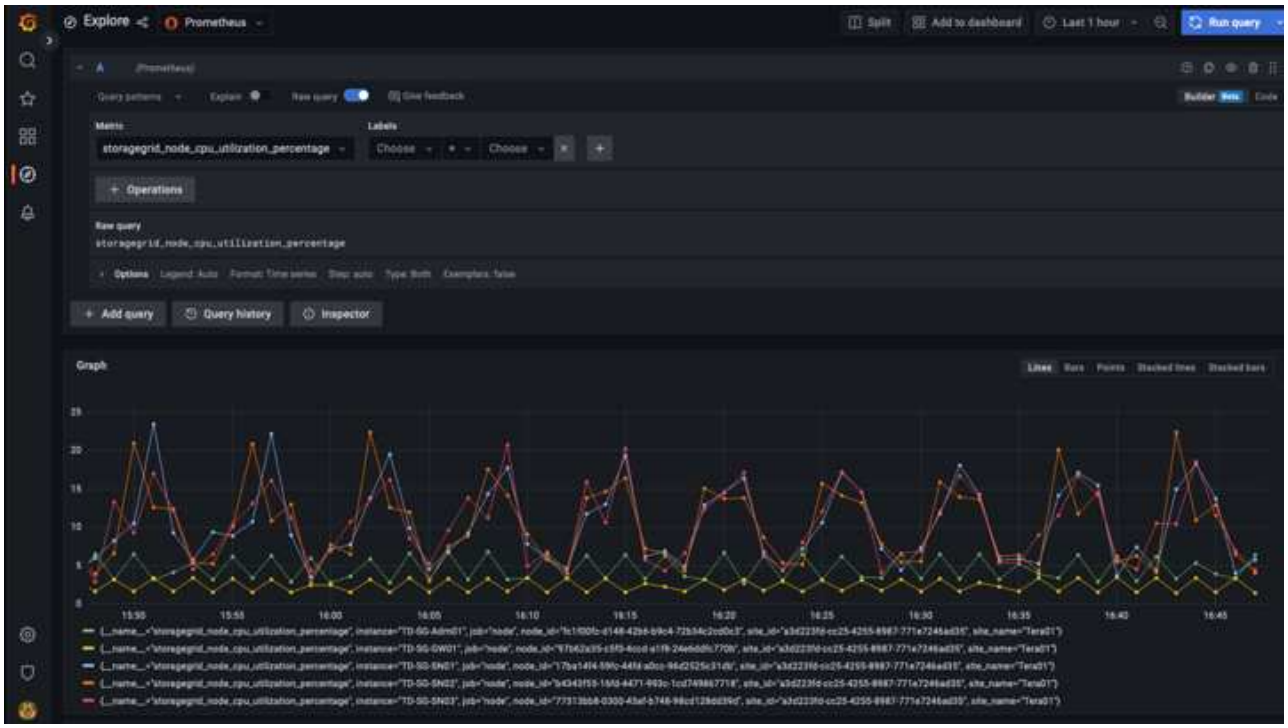
Erstellen eines Grafana Dashboards für StorageGRID

Mit der Installation und dem Betrieb von Grafana und Prometheus ist es jetzt an der Zeit, beide zu verbinden. Dazu wird eine Datenquelle erstellt und ein Dashboard erstellt

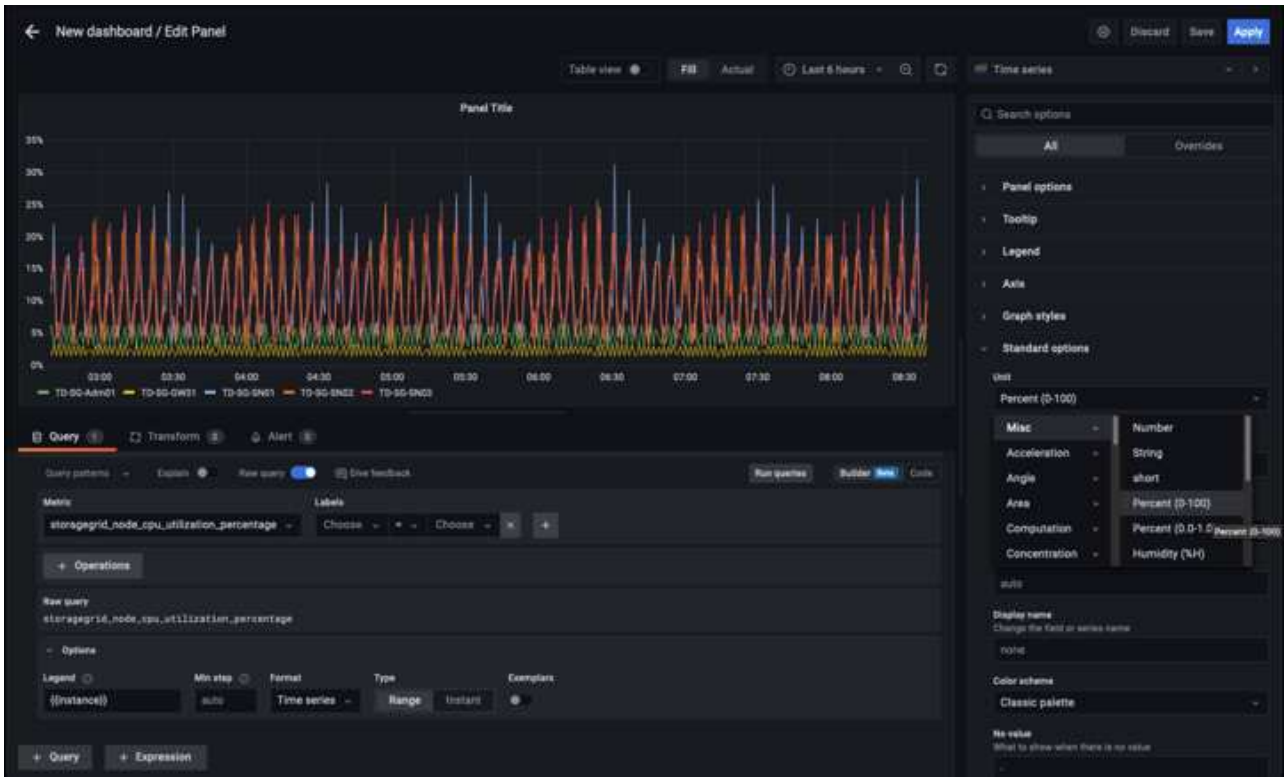
1. Erweitern Sie im linken Fensterbereich „Konfiguration“ und wählen Sie „Datenquellen“, und klicken Sie dann auf die Schaltfläche „Datenquelle hinzufügen“
2. Prometheus wird eine der wichtigsten Datenquellen zur Auswahl sein. Wenn nicht, dann verwenden Sie die Suchleiste zu finden "Prometheus"
3. Konfigurieren Sie die Prometheus-Quelle, indem Sie die URL der prometheus-Instanz und das Scrape-Intervall eingeben, um das Prometheus-Intervall zu entsprechen. Ich habe auch den Abschnitt „Warnungen“ deaktiviert, da ich den Alarmmanager auf prometheus nicht konfiguriert habe.



4. Blättern Sie nach unten, und klicken Sie auf „Speichern & Testen“, wenn Sie die gewünschten Einstellungen eingegeben haben.
5. Nachdem der Konfigurationstest erfolgreich abgeschlossen wurde, klicken Sie auf die Schaltfläche Explore.
   a. Im Erkundungs-Fenster können Sie die gleiche Metrik verwenden, die wir Prometheus mit „storagegrid_Node_cpu_Utifficienty_percenty“ getestet haben, und auf die Schaltfläche „Run query“ klicken



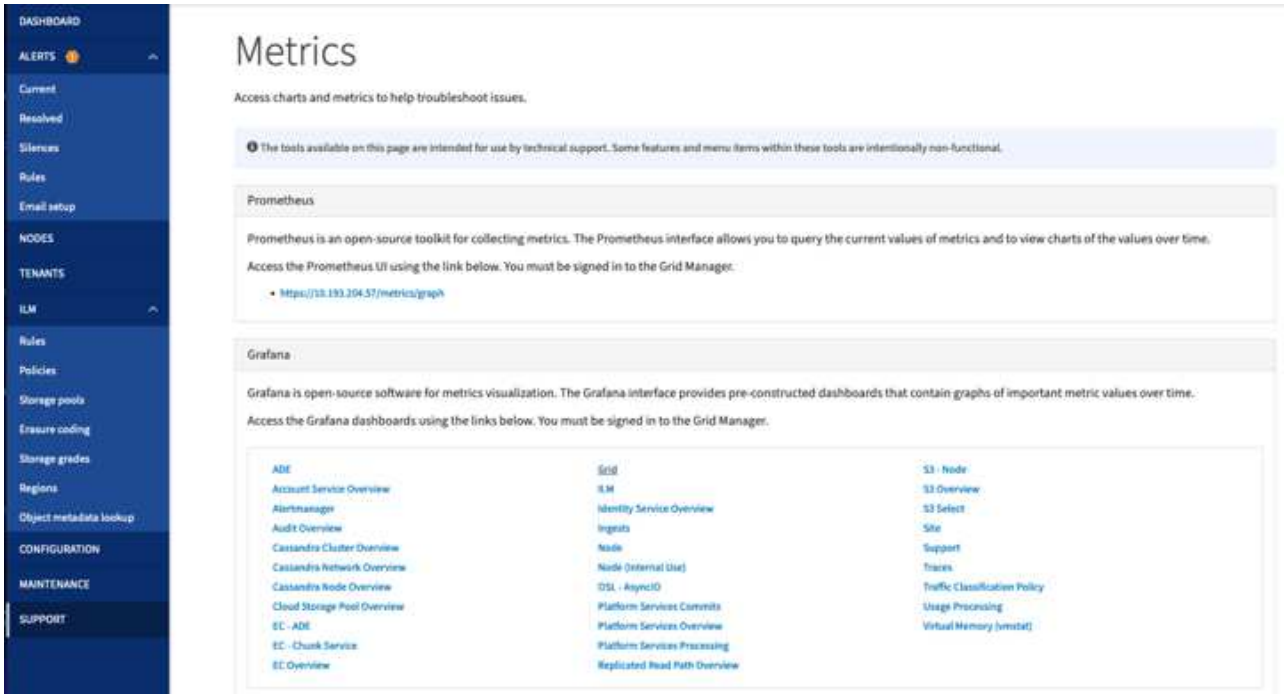
6. Nachdem die Datenquelle konfiguriert ist, können wir jetzt ein Dashboard erstellen.
   a. Erweitern Sie im linken Fensterbereich „Dashboards“ und wählen Sie „+ neues Dashboard“ aus.
   b. Wählen Sie „Neues Bedienfeld hinzufügen“ aus.
   c. Konfigurieren Sie das neue Panel durch Auswahl einer Metrik, wieder werde ich "storagegrid_Node_cpu_Utilement_percenty" verwenden, einen Titel für das Panel eingeben, unten "Optionen" erweitern und für Legende ändern zu Custom und geben Sie "{{instance}}" ein, um die Knotennamen zu definieren, und im rechten Fensterbereich unter "Standardoptionen" setzen "Einheit" auf "Misc/Prozent(0-100)". Klicken Sie dann auf „Übernehmen“, um das Panel im Dashboard zu speichern.



7. Wir könnten unser Dashboard für jede gewünschte Metrik weiter ausbauen, aber glücklicherweise verfügt StorageGRID bereits über Dashboards mit Panels, die wir in unsere benutzerdefinierten Dashboards kopieren können.

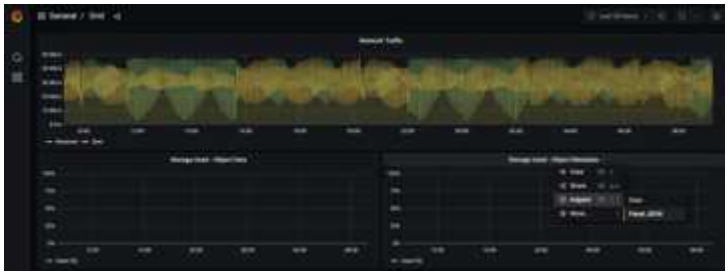
   a. Wählen Sie im linken Fensterbereich der StorageGRID-Managementoberfläche „Support“ und klicken Sie unten in der Spalte „Tools“ auf „Metriken“.

   b. Innerhalb von Kennzahlen wähle ich den Link „Grid“ oben in der mittleren Spalte aus.



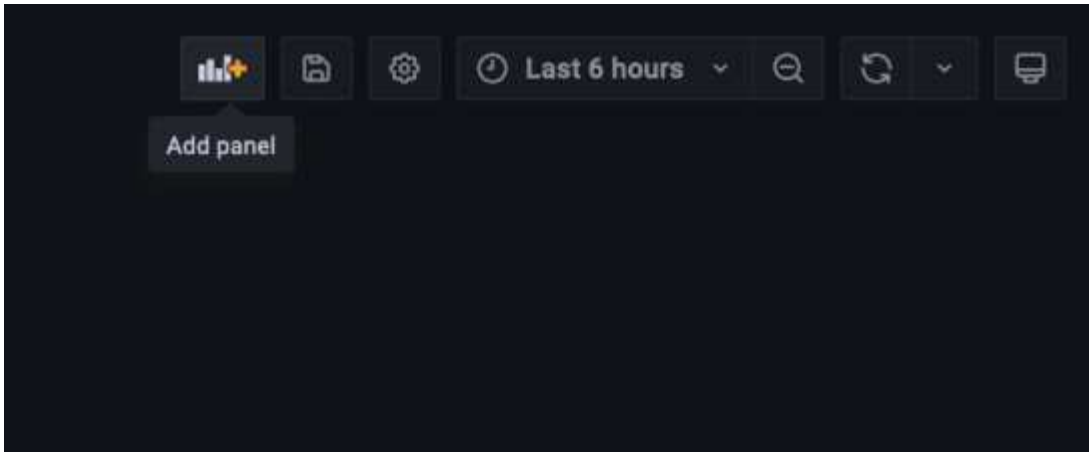
   c. Wählen Sie im Grid-Dashboard den Bereich „Storage Used - Object Metadata“ aus. Klicken Sie auf

den kleinen Pfeil nach unten und auf das Ende des Bedienfeldtitels, um ein Menü zu öffnen. Wählen Sie in diesem Menü „Inspect“ und „Panel JSON“ aus.



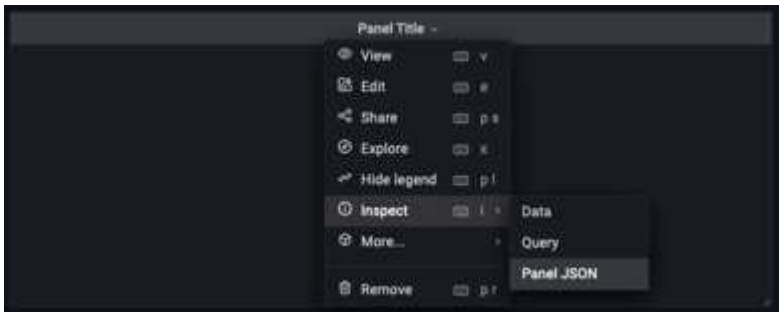
d. Kopieren Sie den JSON-Code und schließen Sie das Fenster.

e. Klicken Sie in unserem neuen Dashboard auf das Symbol, um ein neues Panel hinzuzufügen.

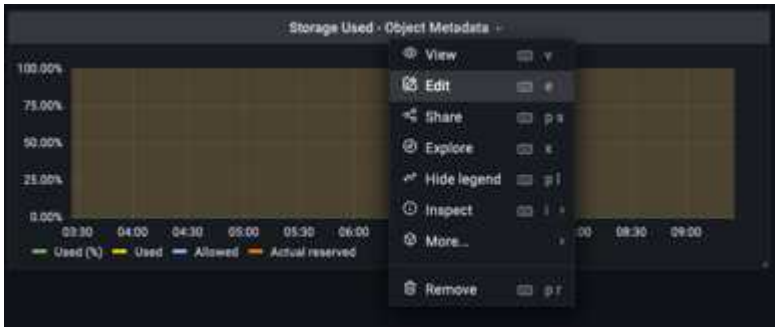


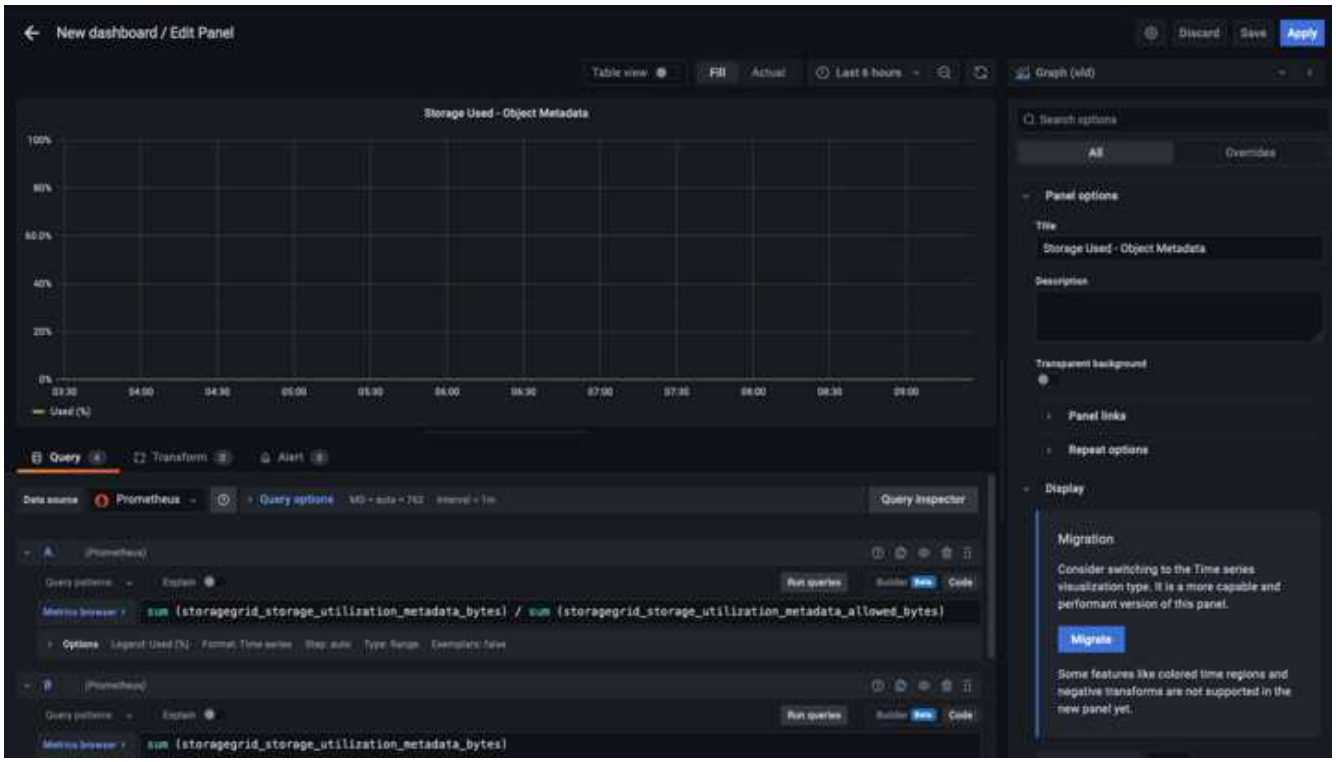
f. Wenden Sie das neue Bedienfeld an, ohne Änderungen vorzunehmen

g. Wie bei dem StorageGRID-Panel sollten Sie auch die JSON überprüfen. Entfernen Sie den gesamten JSON-Code, und ersetzen Sie ihn durch den kopierten Code aus dem StorageGRID-Fenster.



h. Bearbeiten Sie das neue Bedienfeld, und auf der rechten Seite sehen Sie eine Migrationsmeldung mit einem "Migrate"-Button. Klicken Sie auf die Schaltfläche und dann auf die Schaltfläche „Übernehmen“.





8. Sobald Sie alle Panels eingerichtet und so konfiguriert haben, wie Sie möchten. Speichern Sie das Dashboard, indem Sie oben rechts auf das Festplatten-Symbol klicken und Ihrem Dashboard einen Namen geben.

Schlussfolgerung

Jetzt verfügen wir über einen Prometheus Server mit anpassbarer Datenaufbewahrung und Storage-Kapazität. Damit können wir unsere eigenen Dashboards mit den für unsere Betriebsabläufe wichtigsten Kennzahlen weiterentwickeln. Weitere Informationen zu den in der erfassten Prometheus-Kennzahlen finden Sie unter "StorageGRID-Dokumentation".

Von Aron Klein

Datadog SNMP-Konfiguration

Konfigurieren Sie Datadog, um StorageGRID-snmp-Metriken und Traps zu erfassen.

Konfigurieren Sie Das Datadog

Datadog ist eine Überwachungslösung, die Metriken, Visualisierungen und Warnmeldungen bereitstellt. Die folgende Konfiguration wurde mit linux Agent Version 7.43.1 auf einem lokalen Ubuntu 22.04.1-Host auf dem StorageGRID-System implementiert.

Datadog-Profil- und Trap-Dateien, die aus der StorageGRID-MIB-Datei generiert wurden

Datadog bietet eine Methode zum Konvertieren von Produkt-MIB-Dateien in Datadog-Referenzdateien, die für die Zuordnung der SNMP-Meldungen erforderlich sind.

Diese StorageGRID-yaml-Datei für die Datadog-Trap-Auflösungszuordnung wurde nach der gefundenen Anweisung erstellt "Hier". + Platzieren Sie diese Datei in /etc/datadog-Agent/conf.d/snmp.d/Traps_db/ +

- "Laden Sie die Trap yaml-Datei herunter" +
  - **md5-Prüfsumme** 42e27e4210719945a46172b98c379517 +
  - **Sha256 Prüfsumme** d0fe5c8e6ca3c902d054f854b70a85f928cba8b7c76391d356f05d2cf73b6887 +

Diese yaml-Datei für das StorageGRID-Profil für die Datadog-Metrikzuordnung wurde nach der gefundenen Anweisung generiert "Hier". + Platzieren Sie diese Datei in /etc/datadog-Agent/conf.d/snmp.d/profiles/ +

- "Laden Sie die Profil-yaml-Datei herunter" +
  - **md5-Prüfsumme** 72bb7784f4801adda4e0c3ea77df19aa +
  - **Sha256 Prüfsumme** b6b7fadd33063422a8bb8e39b3ead8ab38349ee0229926eadc8585f0087b8cee +

SNMP-Datadog-Konfiguration für Metriken

Die Konfiguration von SNMP für Metriken kann auf zwei Arten verwaltet werden. Sie können für die automatische Erkennung konfigurieren, indem Sie einen Netzwerkadressbereich bereitstellen, der die StorageGRID-Systeme enthält, oder die IP-Adressen der einzelnen Geräte definieren. Der Konfigurationsposition unterscheidet sich je nach getroffenen Entscheidungen. Die automatische Erkennung wird in der Datei des Datadog-Agenten yaml definiert. Explizite Gerätedefinitionen werden in der snmp-yaml-Konfigurationsdatei konfiguriert. Im Folgenden finden Sie Beispiele für jedes System eines StorageGRID.

Automatische Erkennung

Konfiguration befindet sich in /etc/datadog-Agent/datadog.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
```

Einzelne Geräte

/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'
```

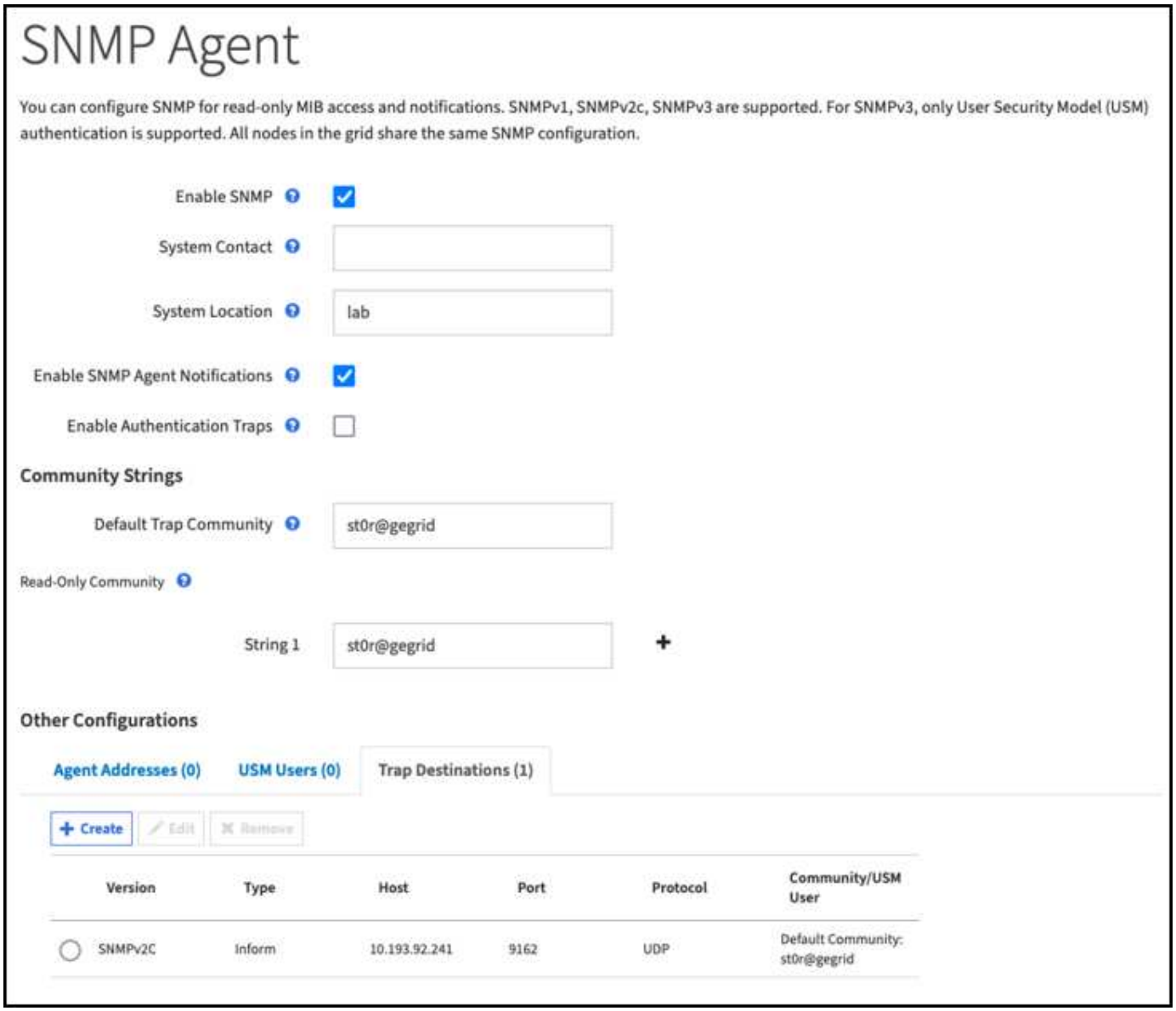
SNMP-Konfiguration für Traps

Die Konfiguration für SNMP-Traps wird in der Datei /etc/datadog-Agent/datadog.yaml der Datendog-Konfiguration definiert

```
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
```

Beispiel für eine StorageGRID-SNMP-Konfiguration

Der SNMP-Agent in Ihrem StorageGRID-System befindet sich auf der Registerkarte Konfiguration in der Spalte Überwachung. Aktivieren Sie SNMP und geben Sie die gewünschten Informationen ein. Wenn Sie Traps konfigurieren möchten, wählen Sie „Traps-Ziele“ und erstellen Sie ein Ziel für den Datadog-Agent-Host, der die Trap-Konfiguration enthält.



Von Aron Klein

Mit rclone können Sie Objekte auf StorageGRID migrieren, VERSCHIEBEN und LÖSCHEN

Rclone ist ein kostenloses Kommandozeilen-Tool und Client für S3-Vorgänge. Sie können rclone verwenden, um Objektdaten auf StorageGRID zu migrieren, zu kopieren und zu löschen. Rclone bietet die Möglichkeit, Buckets auch dann zu löschen, wenn es nicht leer ist. Die Funktion „purge“ ist in einem Beispiel unten dargestellt.

Installieren und Konfigurieren von rclone

Um rclone auf einer Workstation oder einem Server zu installieren, laden Sie es von herunter "rclone.org".

Erste Konfigurationsschritte

1. Erstellen Sie die rclone-Konfigurationsdatei, indem Sie entweder das Konfigurationsskript ausführen oder die Datei manuell erstellen.
2. In diesem Beispiel verwende ich sgdemo für den Namen des entfernten StorageGRID S3-Endpunkts in der rclone-Konfiguration.
   a. Erstellen Sie die Konfigurationsdatei ~/.config/rclone/rclone.conf

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

   b. Führen Sie rclone config aus

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
```

Beispiele für grundlegende Befehle

- **Erstellen Sie einen Eimer:**

  `rclone mkdir remote:bucket`

  ```
  # Rclone mkdir sgdemo:test01
  ```

  Verwenden Sie --no-Check-Certificate, wenn Sie SSL-Zertifikate ignorieren müssen.

- **Alle Buckets auflisten:**

```
rclone lsd remote:
```

```
# Rclone lsd sgdemo:
```

- **Objekte in einem bestimmten Bucket auflisten:**

```
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
```

- **Ein Eimer löschen:**

```
rclone rmdir remote:bucket
```

```
# Rclone rmdir sgdemo:test02
```

- **Legen Sie ein Objekt:**

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

```
# Rclone copy ~/Test/testfile.txt sgdemo:test01
```

- **Holen Sie sich ein Objekt:**

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

```
# Rclone copy sgdemo:test01/testfile.txt ~/Test/testfileS3.txt
```

- **Ein Objekt löschen:**

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

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

- **Objekte in einen Bucket migrieren**

```
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
```

Verwenden Sie --progress oder -P, um den Fortschritt der Aufgabe anzuzeigen. Andernfalls gibt es keine Ausgabe.

- **Löschen eines Buckets und aller Objektinhalte**

```
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
```

Von Siegfried Hepp und Aron Klein

StorageGRID Best Practices für die Implementierung mit Veeam Backup and Replication

Dieser Leitfaden konzentriert sich auf die Konfiguration von NetApp StorageGRID und teilweise Veeam Backup and Replication. Dieses Dokument richtet sich an Storage- und Netzwerkadministratoren, die mit Linux-Systemen vertraut sind und mit der Wartung oder Implementierung eines NetApp StorageGRID-Systems in Kombination mit Veeam Backup and Replication betraut sind.

Überblick

Speicheradministratoren möchten das Wachstum ihrer Daten mit Lösungen managen, die die Anforderungen an Verfügbarkeit, schnelle Wiederherstellung erfüllen, an ihre Bedürfnisse anpassen und ihre Richtlinien für die langfristige Aufbewahrung von Daten automatisieren. Diese Lösungen sollten auch Schutz vor Verlust oder böswilligen Angriffen bieten. Veeam und NetApp haben gemeinsam eine Datensicherungslösung entwickelt, die Veeam Backup & Recovery mit NetApp StorageGRID kombiniert und damit Objekt-Storage vor Ort ermöglicht.

Veeam und NetApp StorageGRID bieten eine benutzerfreundliche Lösung, die zusammen die Anforderungen eines schnellen Datenwachstums und die zunehmenden Vorschriften weltweit erfüllt. Cloud-basierter Objekt-Storage ist für seine Ausfallsicherheit, Skalierbarkeit, betriebliche Effizienz und Kosteneffizienz bekannt, die ihn zur ersten Wahl für Ihre Backups machen. Dieses Dokument enthält Anleitungen und Empfehlungen für die Konfiguration Ihrer Veeam Backup-Lösung und des StorageGRID Systems.

Der Objekt-Workload von Veeam erstellt eine große Anzahl von gleichzeitigen PUT-, DELETE- und LISTENVORGÄNGEN für kleine Objekte. Durch die Unveränderlichkeit wird die Anzahl der Anfragen an den Objektspeicher zur Festlegung von Aufbewahrungs- und Listenversionen weiter addiert. Der Prozess eines Backup-Jobs umfasst das Schreiben von Objekten für die tägliche Änderung. Nach Abschluss der neuen Schreibvorgänge löscht der Job alle Objekte, die auf der Aufbewahrungsrichtlinie des Backups basieren. Die Planung von Backup-Jobs wird sich fast immer überschneiden. Diese Überschneidung führt zu einem großen Teil des Backup-Fensters, das aus 50/50 PUT/DELETE Workloads auf dem Objektspeicher besteht. Anpassungen an die Anzahl gleichzeitiger Vorgänge mit der Einstellung für den Task-Slot vornehmen und die Objektgröße dadurch erhöhen, dass die Blockgröße für den Backup-Job erhöht wird, die Anzahl der Objekte in

den Anforderungen zum Löschen mehrerer Objekte reduziert wird, und wenn Sie das maximale Zeitfenster für die Fertigstellung der Jobs auswählen, wird die Lösung hinsichtlich Performance und Kosten optimiert.

Lesen Sie unbedingt die Produktdokumentation für "Veeam Backup und Replication" Und "StorageGRID" Bevor Sie beginnen. Veeam bietet Rechner mit Informationen zur Dimensionierung der Veeam Infrastruktur und den Kapazitätsanforderungen, die vor der Dimensionierung Ihrer StorageGRID Lösung verwendet werden sollten. Informationen zu validierten Veeam-NetApp Konfigurationen finden Sie auf der Veeam Ready Program Website für "Veeam Ready Objekt, Unveränderlichkeit von Objekten und Repository".

Veeam Konfiguration

Empfohlene Version

Es wird empfohlen, immer auf dem neuesten Stand zu bleiben und die neuesten Hotfixes für Ihr Veeam Backup & Replication 12-System anzuwenden. Derzeit wird empfohlen, mindestens den Veeam Patch P20230718 zu installieren.

S3-Repository-Konfiguration

Ein Scale-out-Backup-Repository (SOBR) ist die Kapazitäts-Tier des S3-Objekt-Storage. Die Kapazitäts-Tier ist eine Erweiterung des primären Repositorys mit längeren Aufbewahrungszeiträumen und einer kostengünstigeren Storage-Lösung. Veeam ermöglicht Unveränderlichkeit über die S3 Object Lock API. Veeam 12 kann mehrere Buckets in einem Scale-out-Repository verwenden. StorageGRID hat keine Obergrenze für die Anzahl der Objekte oder Kapazität in einem einzelnen Bucket. Die Verwendung mehrerer Buckets verbessert möglicherweise die Performance beim Backup von sehr großen Datensätzen, bei denen die Backup-Daten in Objekten bis in den Petabyte-Bereich skaliert werden können.

Je nach Dimensionierung Ihrer spezifischen Lösung und den Anforderungen können Sie gleichzeitige Tasks beschränken. In den Standardeinstellungen wird für jeden CPU-Kern ein Repository-Tasksteckplatz und für jeden Task-Steckplatz ein Limit von 64 gleichzeitig ausgeführten Tasksteckplätzen festgelegt. Wenn Ihr Server beispielsweise 2 CPU-Kerne hat, werden insgesamt 128 gleichzeitige Threads für den Objektspeicher verwendet. Dies beinhaltet PUT, GET und Batch Delete. Es wird empfohlen, einen konservativen Grenzwert für die Taskslots auszuwählen, um mit zu beginnen und diesen Wert einzustellen, sobald Veeam-Backups einen stabilen Zustand von neuen Backups und auslaufenden Backupdaten erreicht haben. Arbeiten Sie mit Ihrem NetApp Account Team zusammen, um die Größe des StorageGRID Systems entsprechend anzupassen, damit die gewünschten Zeitfenster und Leistungen eingehalten werden. Um die optimale Lösung zu bieten, müssen Sie die Anzahl der Aufgabenplätze und die Anzahl der Aufgaben pro Steckplatz anpassen.

Konfiguration des Backupjobs

Veeam Backup-Jobs können mit anderen Blockgrößen konfiguriert werden, die besonders sorgfältig geprüft werden sollten. Die standardmäßige Blockgröße beträgt 1 MB und mit der Speichereffizienz, die Veeam mit Komprimierung und Deduplizierung bietet, erzeugt Objektgrößen von ca. 500 KB für das erste vollständige Backup und 100-200-KB-Objekte für die inkrementellen Jobs. Wir können die Performance des Objektspeichers deutlich steigern und die Anforderungen reduzieren, indem wir eine größere Blockgröße für Backups auswählen. Obwohl die größere Blockgröße zu großen Verbesserungen der Objektspeicher-Performance führt, geht dies zu Kosten, da potenziell höhere Anforderungen an die Kapazität des primären Storage aufgrund niedrigerer Storage-Effizienz-Performance entstehen. Es wird empfohlen, die Backup-Jobs mit einer Blockgröße von 4 MB zu konfigurieren, die ca. 2 MB Objekte für die vollständigen Backups und 700kB-1MB Objektgrößen für inkrementelle Backups erzeugt. Kunden ziehen möglicherweise sogar die Konfiguration von Backup-Jobs mit einer Blockgröße von 8 MB in Betracht, die mithilfe des Veeam Supports aktiviert werden kann.

Bei der Implementierung von unveränderlichen Backups wird auf S3 Object Lock im Objektspeicher gesetzt.

Die Option „Unveränderlichkeit“ generiert eine größere Anzahl von Anfragen an den Objektspeicher zur Auflistung und Aktualisierung der Aufbewahrung der Objekte.

Wenn Backup-Retentions ablaufen, verarbeiten die Backup-Jobs das Löschen von Objekten. Veeam sendet die Löschanforderungen je Anforderung in 1000 Objekten an den Objektspeicher. Bei kleinen Lösungen muss diese ggf. angepasst werden, um die Anzahl der Objekte pro Anfrage zu verringern. Eine Senkung dieses Werts hat den zusätzlichen Vorteil, dass die Löschanforderungen gleichmäßig auf die Knoten im StorageGRID-System verteilt werden. Es wird empfohlen, die Werte in der folgenden Tabelle als Ausgangspunkt für die Konfiguration der Grenze für das Löschen mehrerer Objekte zu verwenden. Multiplizieren Sie den Wert in der Tabelle mit der Anzahl der Knoten für den ausgewählten Gerätetyp, um den Wert für die Einstellung in Veeam zu erhalten. Wenn dieser Wert gleich oder größer als 1000 ist, muss der Standardwert nicht angepasst werden. Wenn dieser Wert angepasst werden muss, wenden Sie sich an den Veeam-Support, um die Änderung vorzunehmen.

Appliance-Modell	S3MultiObjectDeleteLimit pro Knoten
SG5712	34
SG5760	75
SG6060	200

Wenden Sie sich an Ihr NetApp Account Team, um die empfohlene Konfiguration basierend auf Ihren spezifischen Anforderungen zu erhalten. Die Empfehlungen für die Veeam-Konfigurationseinstellungen umfassen:

- Blockgröße des Backupjobs = 4 MB
- SOBR-Task-Slot-Limit= 2-16
- Limit Für Mehrere Objekte Löschen = 34-1000

StorageGRID-Konfiguration

Empfohlene Version

NetApp StorageGRID 11.6 oder 11.7 mit dem aktuellen Hotfix sind die empfohlenen Versionen für Veeam-Bereitstellungen. In StorageGRID 11.6.0.11 und 11.7.0.4 wurden zahlreiche Optimierungsfunktionen eingeführt, von denen Veeam-Workloads profitieren. Es wird empfohlen, immer auf dem neuesten Stand zu bleiben und die neuesten Hotfixes für Ihr StorageGRID-System anzuwenden.

Load Balancer und S3-Endpunktkonfiguration

Für Veeam muss der Endpunkt nur über HTTPS verbunden sein. Eine nicht verschlüsselte Verbindung wird von Veeam nicht unterstützt. Das SSL-Zertifikat kann ein selbstsigniertes Zertifikat, eine private vertrauenswürdige Zertifizierungsstelle oder eine öffentliche vertrauenswürdige Zertifizierungsstelle sein. Um den kontinuierlichen Zugriff auf das S3-Repository zu gewährleisten, wird die Verwendung von mindestens zwei Load Balancern in einer HA-Konfiguration empfohlen. Beim Lastausgleich kann es sich um einen von StorageGRID bereitgestellten integrierten Load Balancer handeln, der sich auf jedem Administrator-Node und Gateway-Node oder bei Lösungen von Drittanbietern wie F5, Kemp, HAProxy, Loadbalanacer.org usw. befindet Mithilfe eines StorageGRID Load Balancer kann man Traffic-Klassifikatoren (QoS-Regeln) festlegen, die den Veeam Workload priorisieren können oder Veeam auf Workloads mit höherer Priorität im StorageGRID System beschränken.

S3-Bucket

StorageGRID ist ein sicheres mandantenfähiges Storage-System. Es wird empfohlen, einen dedizierten Mandanten für den Veeam Workload zu erstellen. Optional kann ein Storage-Kontingent zugewiesen werden. Aktivieren Sie als Best Practice „eigene Identitätsquelle verwenden“. Sichern Sie den Mandanten-Root-Managementbenutzer mit einem geeigneten Passwort. Veeam Backup 12 erfordert eine hohe Konsistenz für S3 Buckets. StorageGRID bietet mehrere Konsistenzoptionen, die auf Bucket-Ebene konfiguriert sind. Implementierungen an mehreren Standorten, bei denen Veeam von diversen Standorten auf Daten zugreifen kann, wählen Sie „Strong Global“. Wenn Veeam-Backups und -Restores nur an einem einzigen Standort durchgeführt werden, sollte das Konsistenzniveau auf „Strong-Site“ gesetzt werden. Weitere Informationen zu Bucket-Konsistenzstufen finden Sie im "Dokumentation". Um StorageGRID Backups zur Unveränderlichkeit von Veeam zu nutzen, muss S3 Object Lock global aktiviert und während der Bucket-Erstellung auf dem Bucket konfiguriert werden.

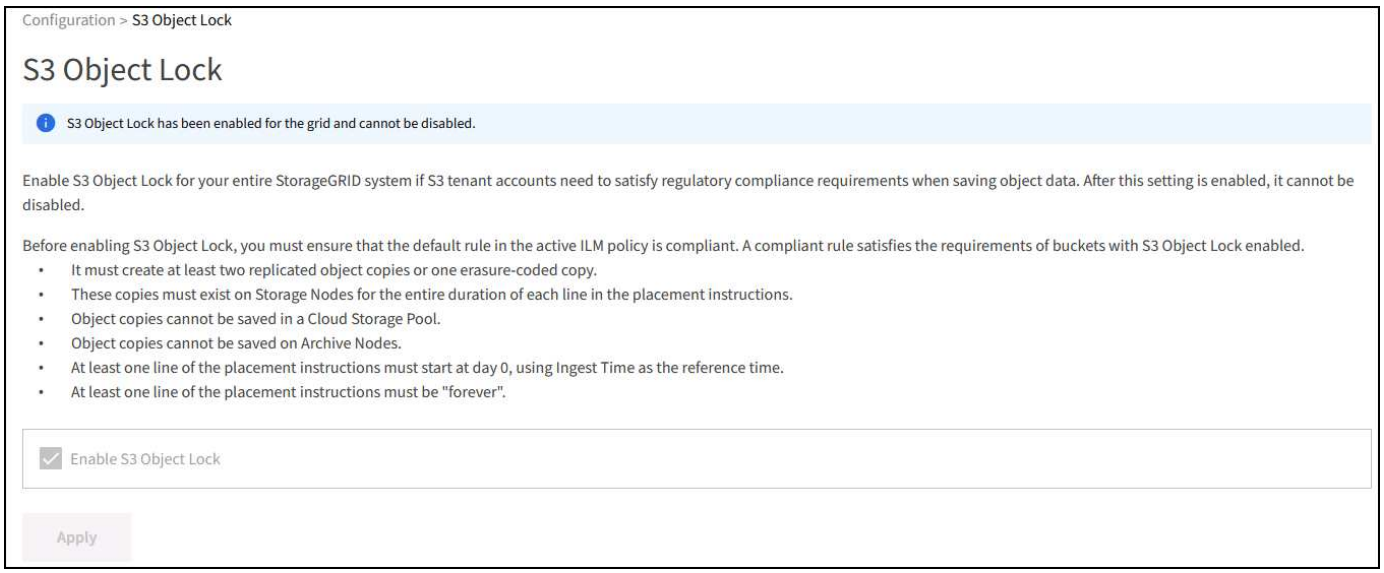
Lifecycle Management

StorageGRID unterstützt Replizierung und Erasure Coding für eine Sicherung auf Objektebene über StorageGRID Nodes und Standorte hinweg. Erasure Coding erfordert mindestens eine Objektgröße von 200 kB. Die standardmäßige Blockgröße für Veeam von 1 MB erzeugt Objektgrößen, die oft unter dieser empfohlenen Mindestgröße von 200 KB liegen können, nachdem Veeam die Storage-Effizienz erreicht hat. Für die Performance der Lösung wird empfohlen, kein Erasure Coding-Profil für mehrere Standorte zu verwenden, es sei denn, die Verbindung zwischen den Standorten reicht aus, um keine Latenz hinzuzufügen oder die Bandbreite des StorageGRID-Systems zu beschränken. Bei einem StorageGRID System mit mehreren Standorten kann die ILM-Regel so konfiguriert werden, dass eine einzige Kopie an jedem Standort gespeichert wird. Um die ultimative Aufbewahrungszeit zu gewährleisten, kann eine Regel für die Speicherung einer Kopie, die nach dem Verfahren zur Fehlerkorrektur codiert wurde, an jedem Standort konfiguriert werden. Die am besten empfohlene Implementierung für diesen Workload ist der lokale Einsatz von zwei Kopien auf den Veeam Backup Servern.

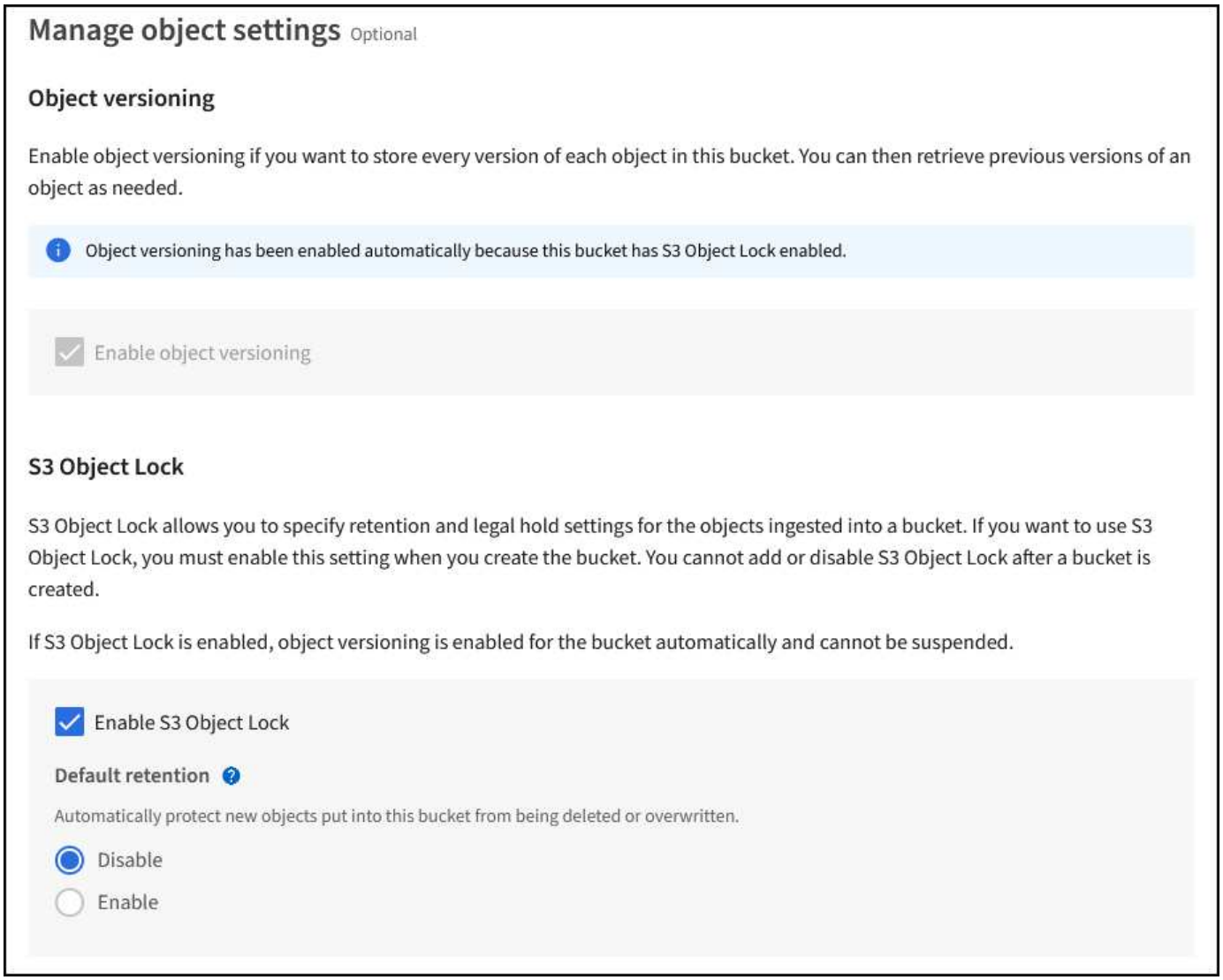
Zentrale Punkte bei der Implementierung

StorageGRID

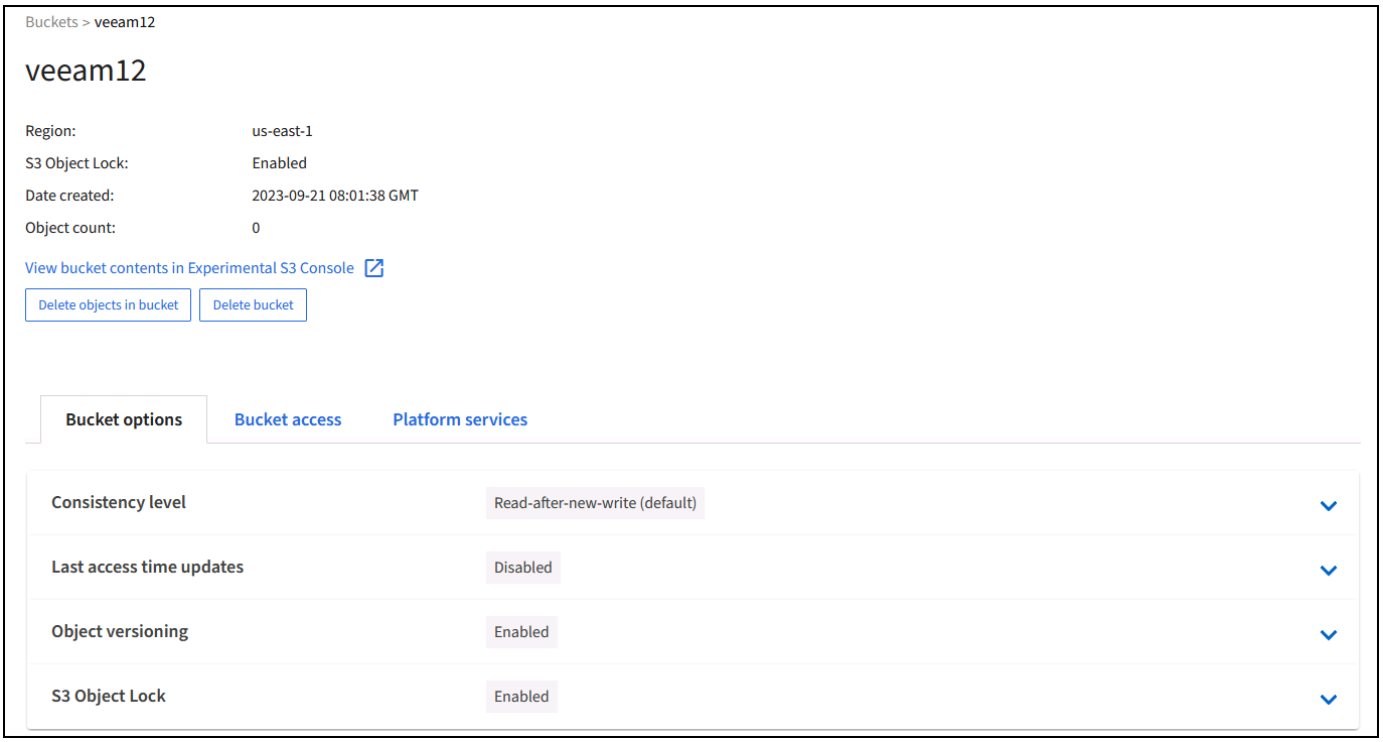
Stellen Sie sicher, dass die Objektsperre auf dem StorageGRID System aktiviert ist, falls eine Unveränderlichkeit erforderlich ist. Suchen Sie die Option in der Management-UI unter Configuration/S3 Object Lock.



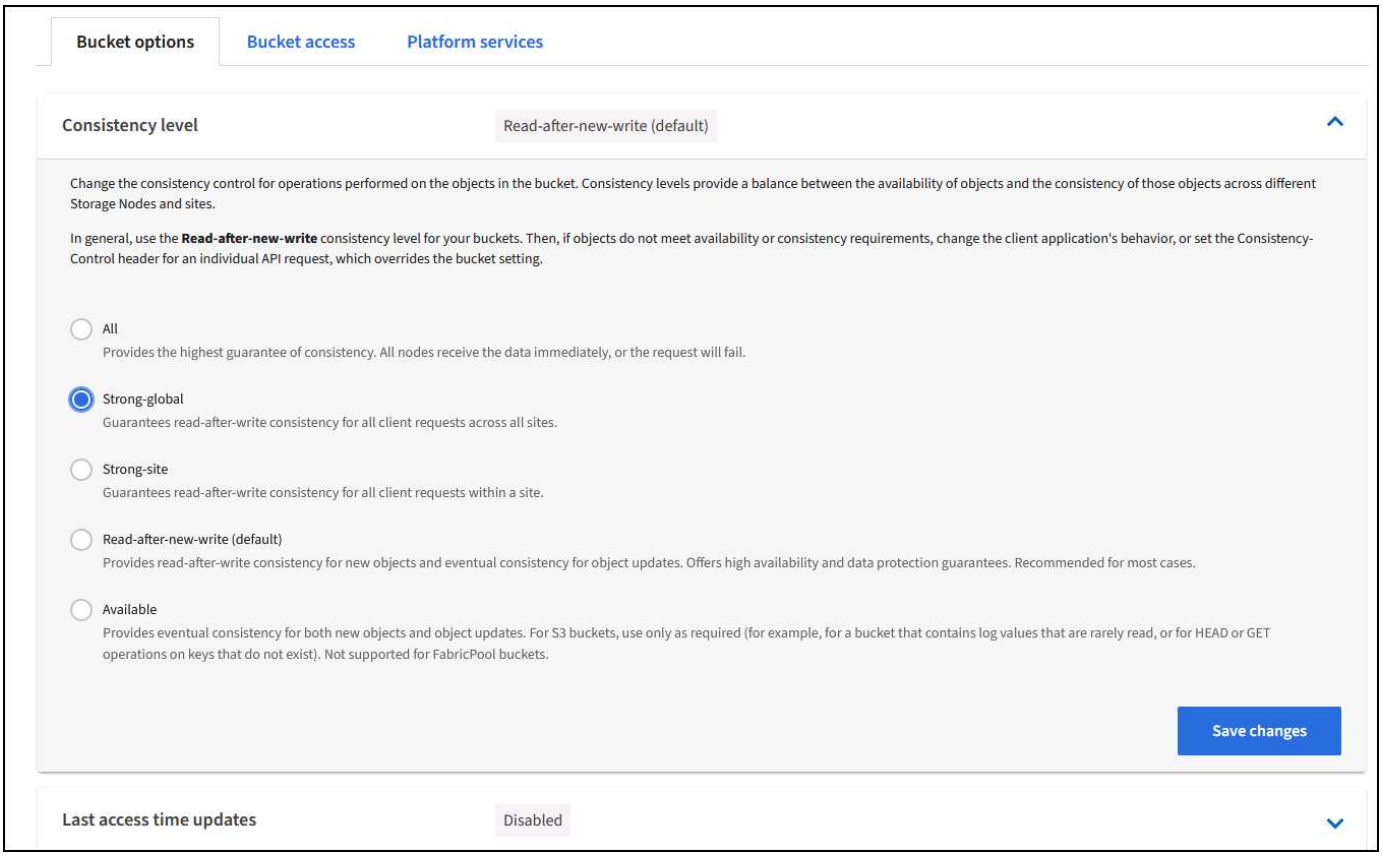
Wählen Sie bei der Erstellung des Buckets die Option „S3 Object Lock aktivieren“ aus, wenn dieser Bucket zur Unveränderlichkeit von Backups verwendet werden soll. Dadurch wird die Bucket-Versionierung automatisch aktiviert. Die Standardaufbewahrung bleibt deaktiviert, da Veeam die Objektaufbewahrung explizit festlegt. Versionierung und S3 Object Lock sollten nicht ausgewählt werden, wenn Veeam keine unveränderlichen Backups erstellt.



Sobald der Bucket erstellt wurde, gehen Sie zur Detailseite des erstellten Buckets. Wählen Sie die Konsistenzstufe aus.



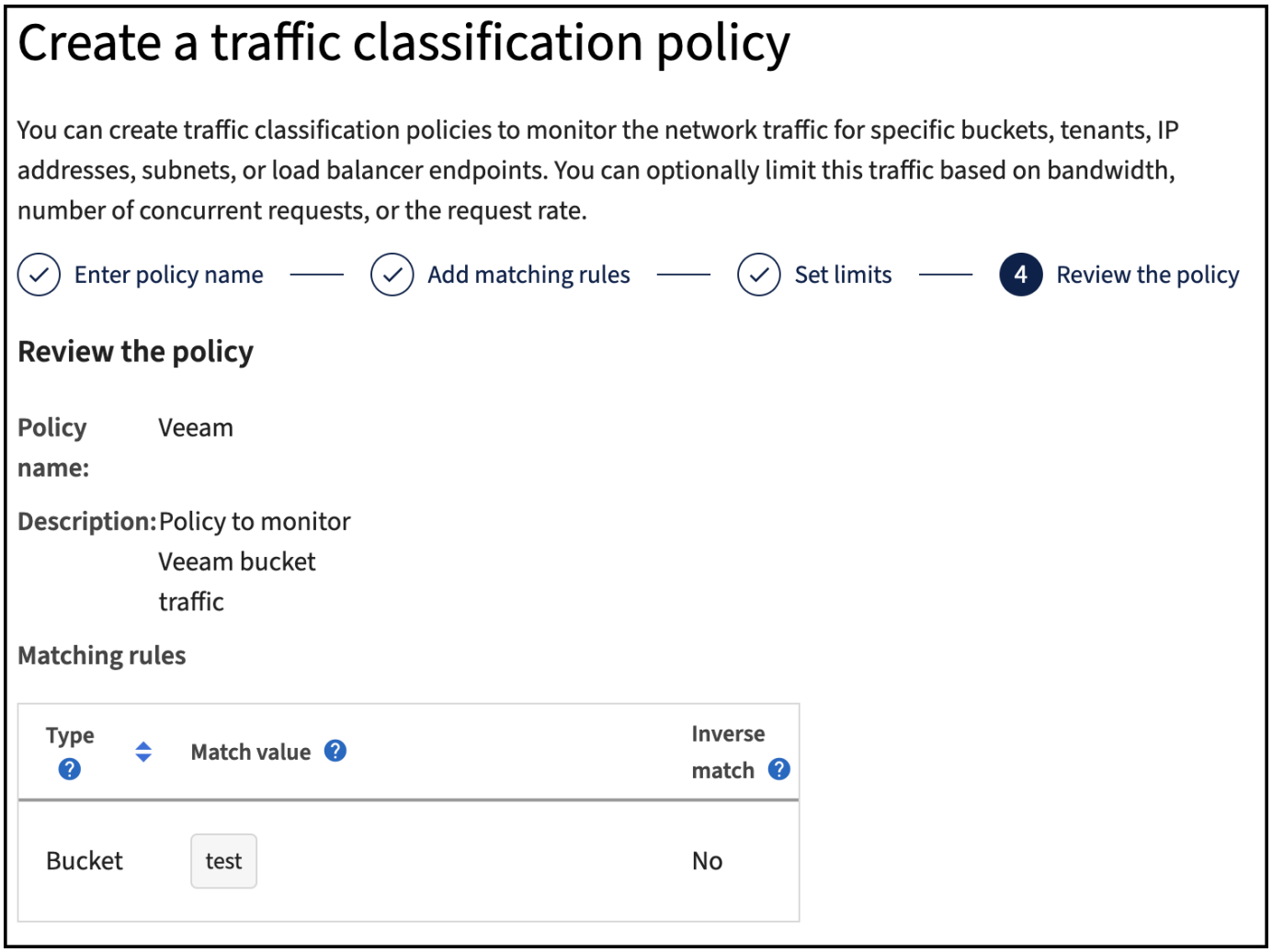
Veeam erfordert eine hohe Konsistenz für S3-Buckets. Wenn also Implementierungen an mehreren Standorten implementiert werden, bei denen Veeam von diversen Standorten auf die Daten zugreifen kann, wählen Sie „Strong Global“. Wenn Veeam-Backups und -Restores nur an einem einzigen Standort durchgeführt werden, sollte das Konsistenzniveau auf „Strong-Site“ gesetzt werden. Speichern Sie die Änderungen.



StorageGRID bietet einen integrierten Load Balancer auf jedem Admin-Node und dedizierten Gateway-Nodes.

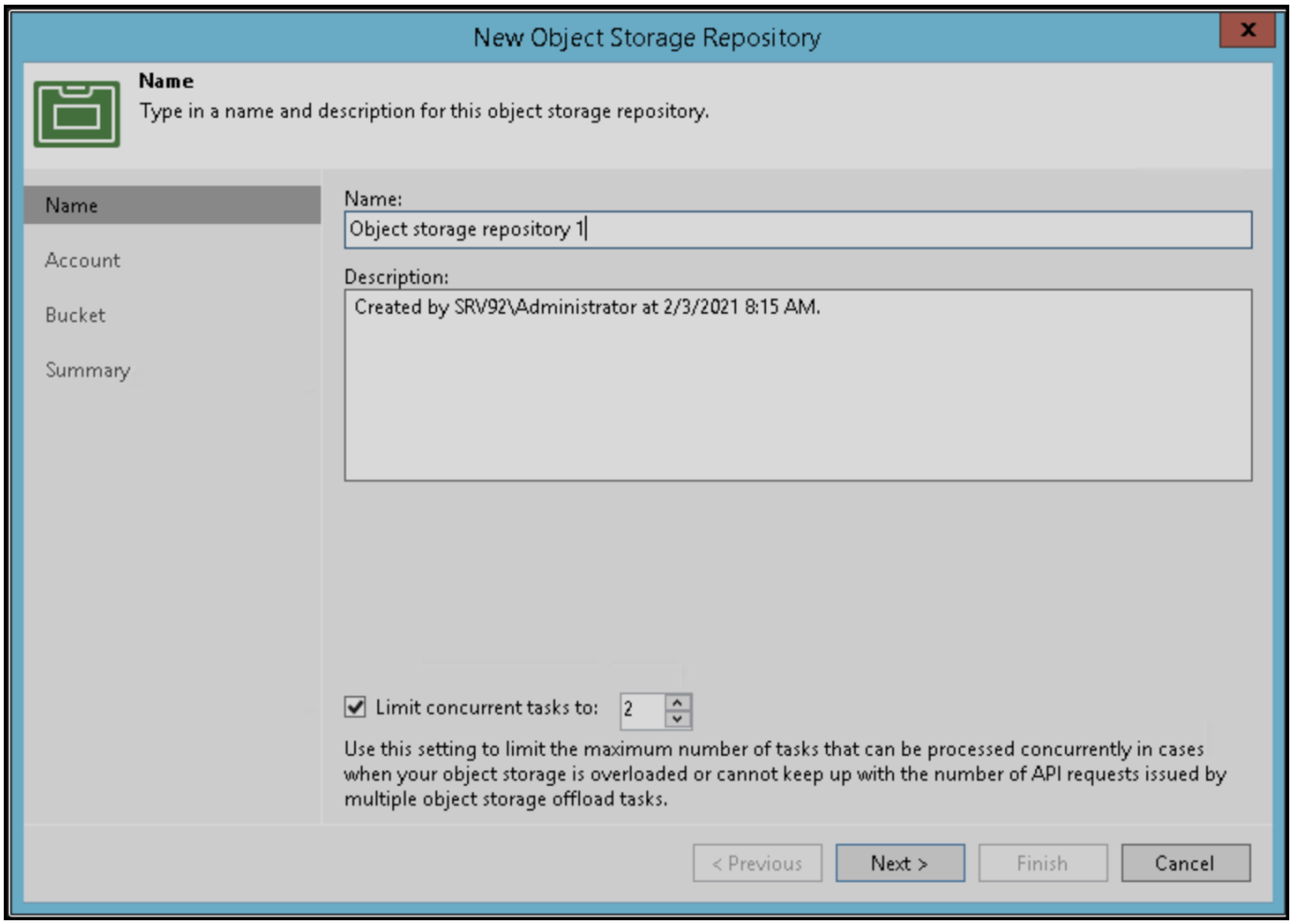
Einer der vielen Vorteile dieser Load Balancer ist die Möglichkeit zur Konfiguration von Richtlinien zur Traffic-Klassifizierung (QoS). Diese dienen hauptsächlich der Beschränkung der Auswirkungen von Applikationen auf andere Client-Workloads oder der Priorisierung von Workloads gegenüber anderen. Sie bieten jedoch auch einen Bonus bei der Erfassung zusätzlicher Metriken zur Unterstützung des Monitorings.

Wählen Sie auf der Registerkarte „Konfiguration“ die Option „Traffic Classification“ aus, und erstellen Sie eine neue Richtlinie. Benennen Sie die Regel, und wählen Sie entweder den/die Bucket(s) oder den Mandanten als Typ aus. Geben Sie die Namen der Bucket(s) oder Tenant ein. Falls QoS erforderlich ist, legen Sie eine Grenze fest. Bei den meisten Implementierungen jedoch möchten wir nur die Monitoring-Vorteile hinzufügen, damit Sie keine Obergrenze festlegen können.

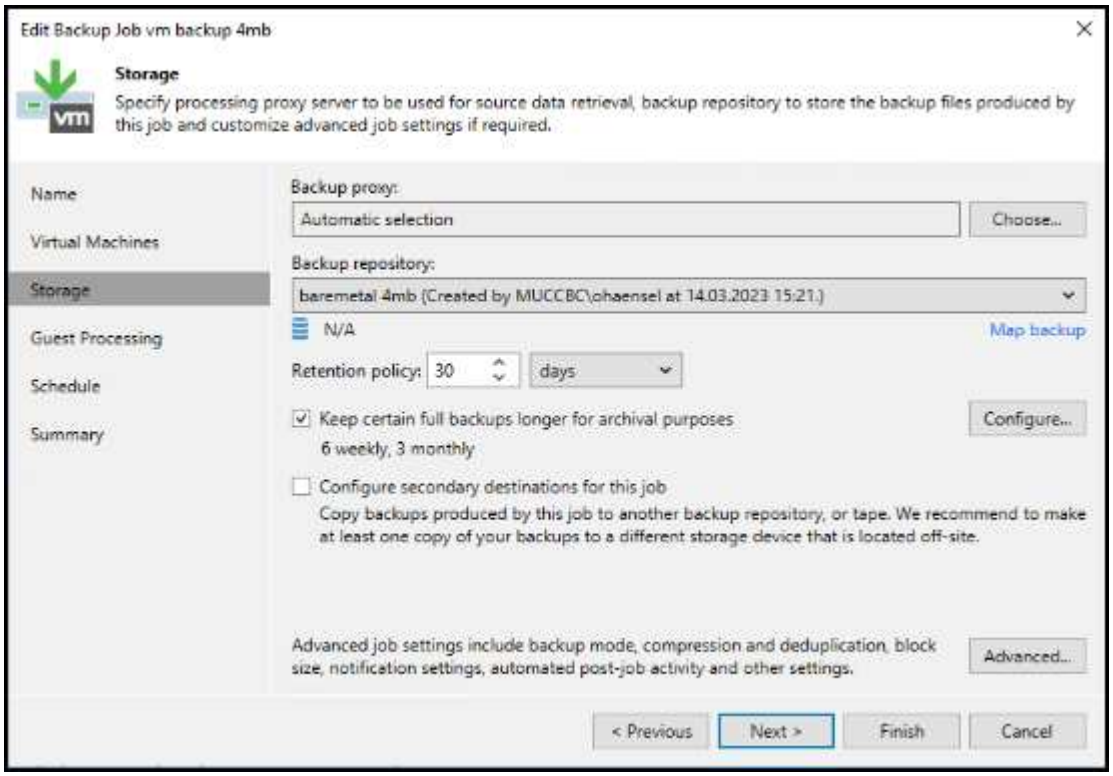


Veeam

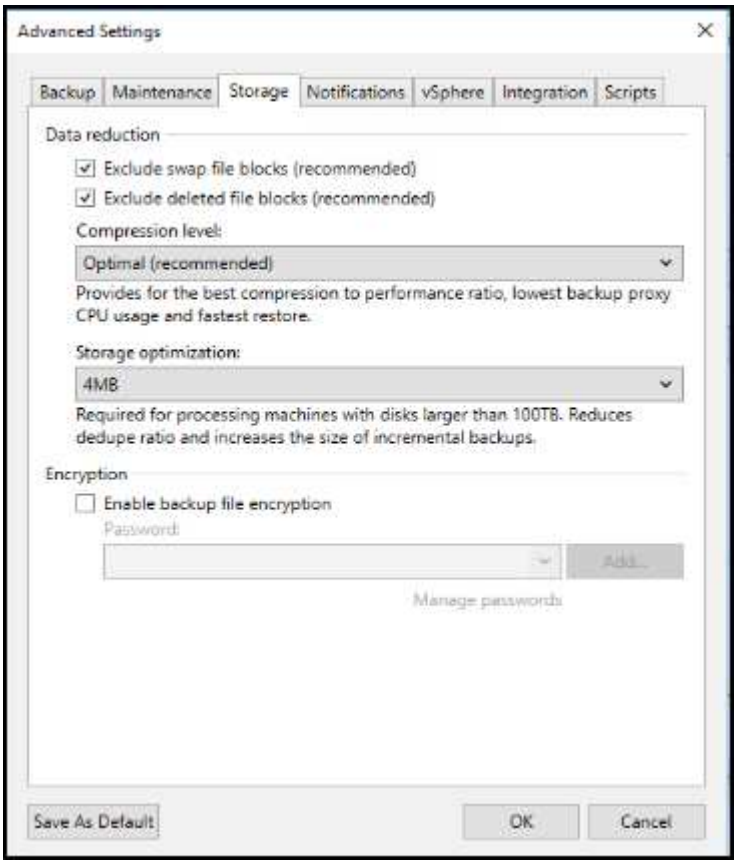
Je nach Modell und Anzahl der StorageGRID Appliances kann es erforderlich sein, eine Begrenzung der Anzahl gleichzeitiger Operationen auf dem Bucket auszuwählen und zu konfigurieren.



Folgen Sie der Veeam Dokumentation zur Konfiguration des Backup-Jobs in der Veeam Konsole, um den Assistenten zu starten. Wählen Sie nach dem Hinzufügen von VMs das SOBR-Repository aus.



Klicken Sie auf Erweiterte Einstellungen, und ändern Sie die Einstellungen für die Speicheroptimierung auf 4 MB oder mehr. Komprimierung und Deduplizierung sollen aktiviert werden. Ändern Sie die Gasteinstellungen entsprechend Ihren Anforderungen und konfigurieren Sie den Zeitplan für den Backupjob.



Monitoring von StorageGRID

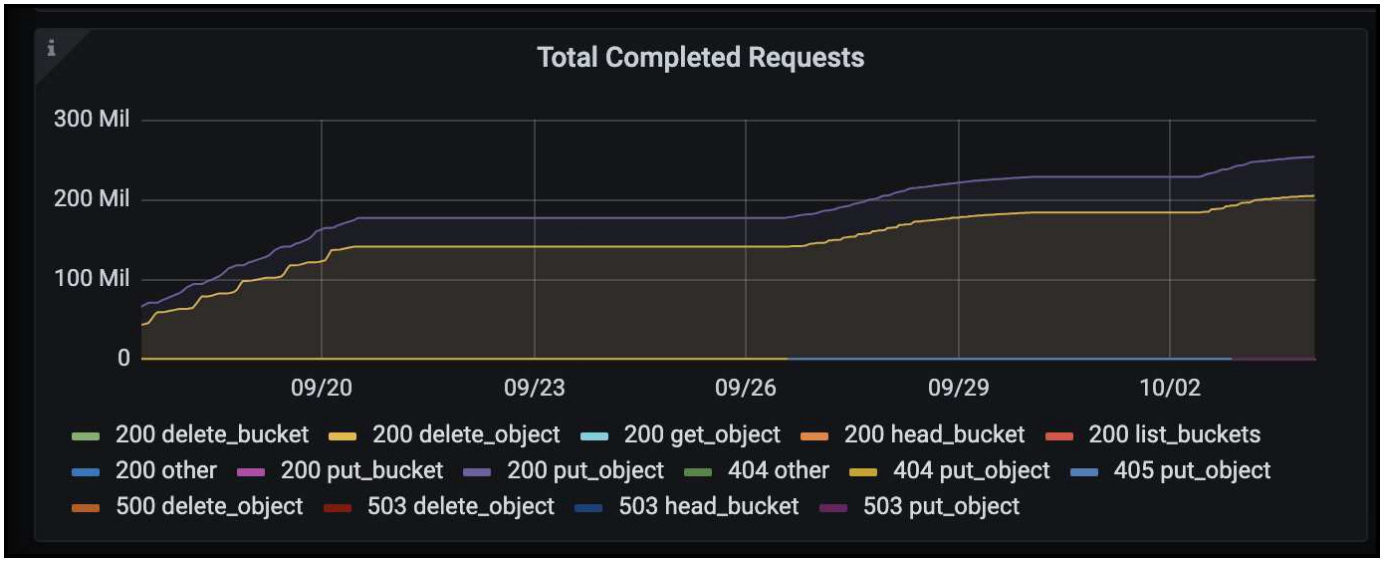
Um sich ein vollständiges Bild davon zu machen, wie Veeam und StorageGRID zusammenarbeiten, müssen Sie warten, bis die Aufbewahrungszeit der ersten Backups abgelaufen ist. Bis zu diesem Zeitpunkt besteht der Veeam-Workload in erster Linie aus PUT-Vorgängen und es sind keine Löschungen aufgetreten. Sobald Sicherungsdaten ablaufen und Clean-ups durchgeführt werden, können Sie jetzt die vollständige konsistente Nutzung im Objektspeicher sehen und die Einstellungen in Veeam bei Bedarf anpassen.

StorageGRID bietet bequeme Diagramme zur Überwachung des Betriebs des Systems auf der Registerkarte „Support“ auf der Seite „Kennzahlen“. Sie sehen sich primär die S3 Übersicht, ILM und die Richtlinie zur Klassifizierung von Datenverkehr an, wenn eine Richtlinie erstellt wurde. Im S3-Übersichts-Dashboard erhalten Sie Informationen zu den S3-Betriebsraten, Latenzen und Anfragenreaktionen.

Bei Blick auf die S3-Raten und aktiven Anfragen sehen Sie, wie viel von der Last die einzelnen Nodes verarbeiten, und wie viele Anfragen insgesamt nach Typ verarbeitet werden.

Im Diagramm „Durchschnittliche Dauer“ wird die durchschnittliche Zeit angezeigt, die jeder Knoten für jeden Anforderungstyp einnimmt. Dies ist die durchschnittliche Latenz der Anfrage und kann ein guter Indikator dafür sein, dass möglicherweise zusätzliche Anpassungen erforderlich sind, oder dass das StorageGRID-System mehr Last aufnehmen kann.

Im Diagramm „abgeschlossene Anforderungen gesamt“ werden die Anforderungen nach Typ und Antwortcodes angezeigt. Wenn Sie andere Antworten als 200 (OK) für die Antworten sehen, kann dies auf ein Problem hinweisen, wie das StorageGRID-System wird stark geladen Senden 503 (Slow Down) Antworten und einige zusätzliche Tuning erforderlich sein, oder die Zeit ist gekommen, um das System für die erhöhte Last zu erweitern.



Im ILM Dashboard können Sie die Performance beim Löschen des StorageGRID Systems überwachen.

StorageGRID verwendet eine Kombination aus synchronen und asynchronen Löschungen auf jedem Node, um die Gesamtleistung für alle Anforderungen zu optimieren.

Mithilfe einer Richtlinie zur Traffic-Klassifizierung können wir Kennzahlen zum Load Balancer Anforderungsdurchsatz, zu Raten, zur Dauer sowie zu den Objektgrößen anzeigen, die Veeam sendet und empfängt.

Wo Sie weitere Informationen finden

Sehen Sie sich die folgenden Dokumente und/oder Websites an, um mehr über die in diesem Dokument beschriebenen Informationen zu erfahren:

- "NetApp StorageGRID 11.7 Produktdokumentation"
- "Veeam Backup und Replication"

Von Oliver Haensel und Aron Klein

Dremio Datenquelle mit StorageGRID konfigurieren

Dremio unterstützt eine Vielzahl von Datenquellen, einschließlich Cloud-basiertem oder lokalem Objektspeicher. Sie können Dremio so konfigurieren, dass StorageGRID als Objektspeicher-Datenquelle verwendet wird.

Dremio-Datenquelle konfigurieren

Voraussetzungen

- Eine StorageGRID S3-Endpunkt-URL, eine s3-Zugriffsschlüssel-ID des Mandanten und ein geheimer Zugriffsschlüssel.
- StorageGRID-Konfigurationsempfehlung: Deaktivieren Sie die Komprimierung (standardmäßig deaktiviert).

  Dremio verwendet Byte-Bereich GET, um während der Abfrage verschiedene Byte-Bereiche aus demselben Objekt gleichzeitig abzurufen. Die typische Größe für Anforderungen im Byte-Bereich beträgt 1 MB. Komprimiertes Objekt beeinträchtigt die GET-Performance im Byte-Bereich.

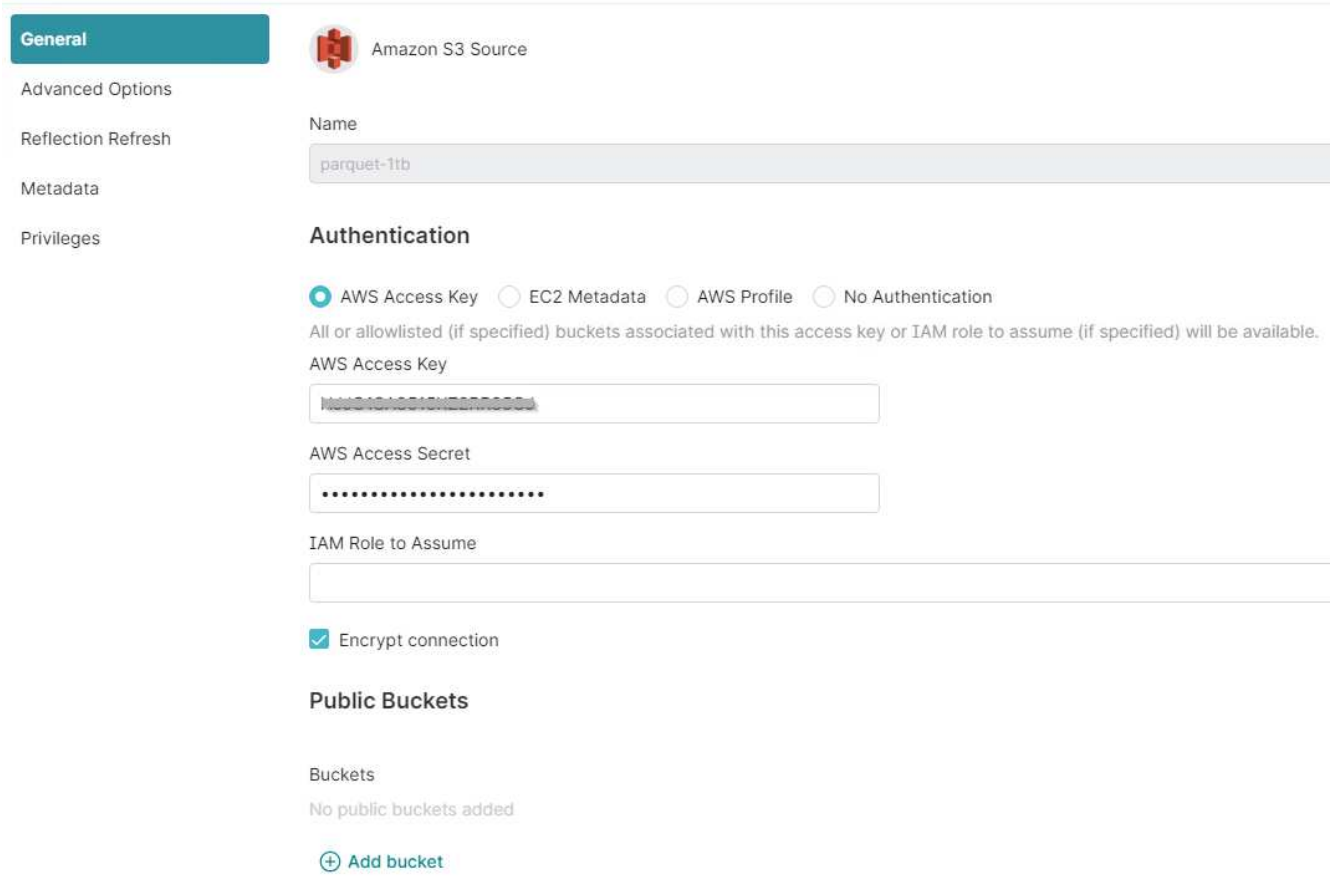
Dremio-Führer

"Connecting to Amazon S3 - Configuring S3-Compatible Storage".

Anweisung

1. Klicken Sie auf der Seite Dremio Datasets auf + signieren, um eine Quelle hinzuzufügen, und wählen Sie „Amazon S3“.
2. Geben Sie einen Namen für diese neue Datenquelle ein: StorageGRID S3-Mandanten-Zugriffsschlüssel-ID und geheimer Zugriffsschlüssel.

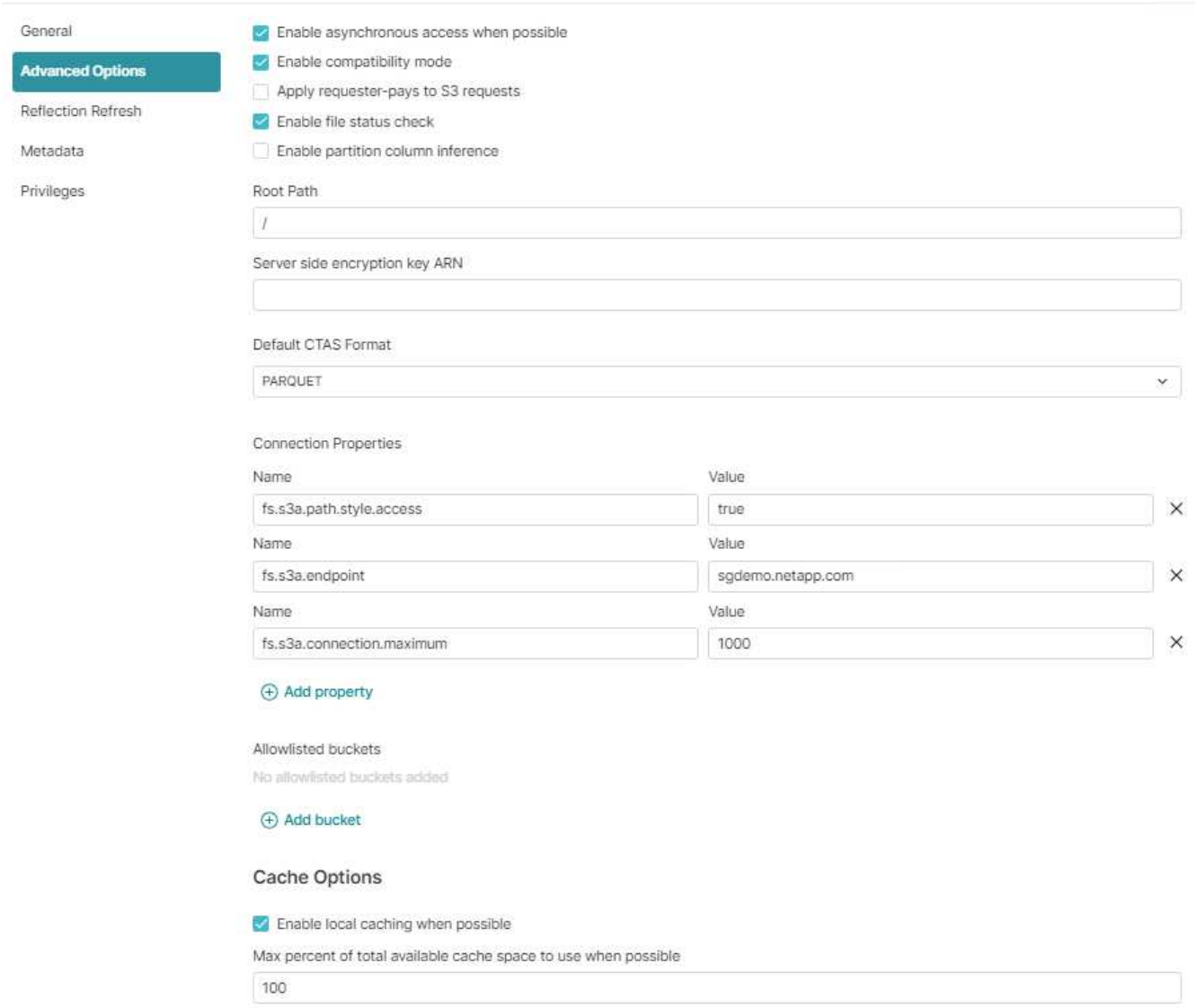
3. Aktivieren Sie das Kontrollkästchen „Verbindung verschlüsseln“, wenn HTTPS für die Verbindung zum StorageGRID S3-Endpunkt verwendet wird.
   Wenn Sie ein selbstsigniertes CA-Zertifikat für diesen s3-Endpunkt verwenden, folgen Sie der Dremio-Anleitung, um dieses CA-Zertifikat in den <JAVA_HOME>/jre/lib/Security + des Dremio-Servers hinzuzufügen
   Beispiel Screenshot



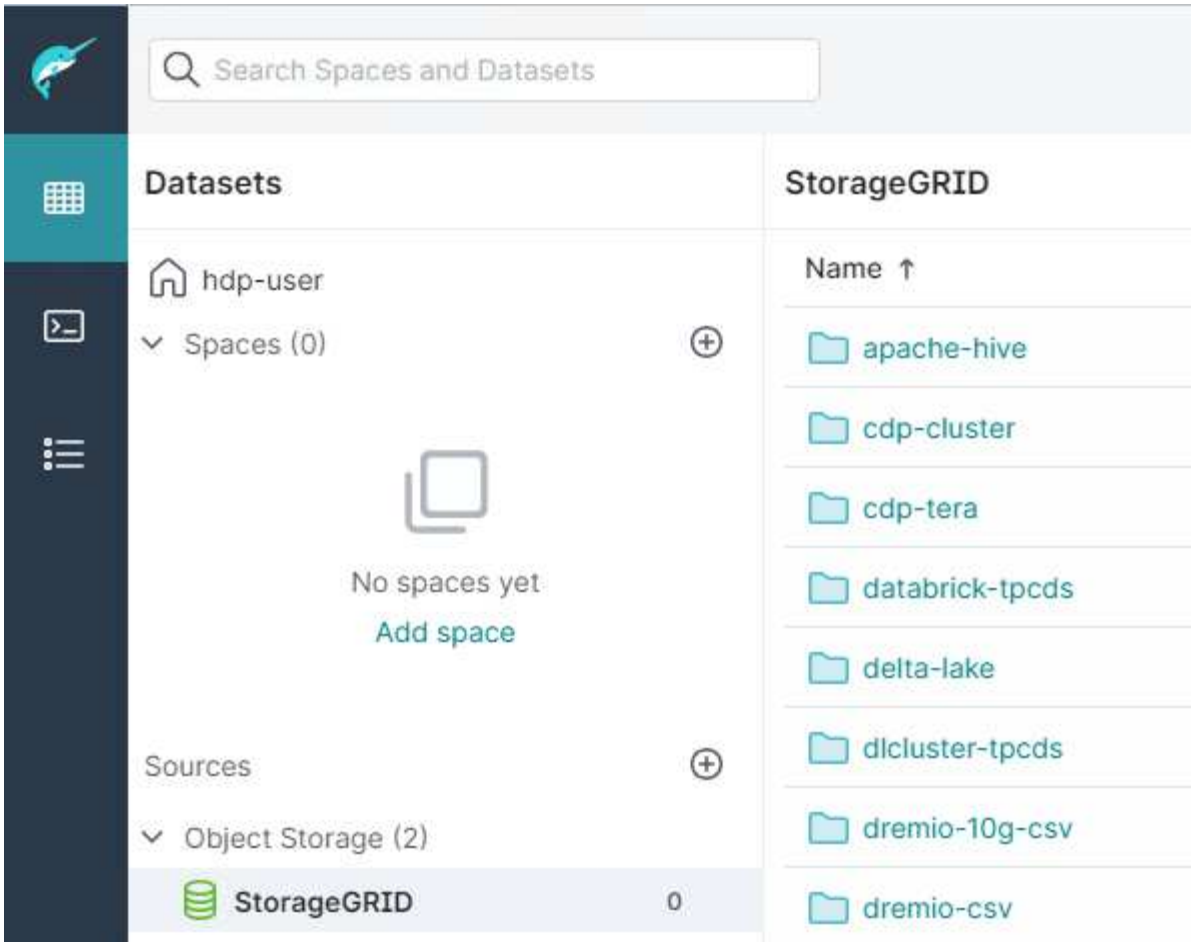
4. Klicken Sie auf „Erweiterte Optionen“, und aktivieren Sie „Kompatibilitätsmodus aktivieren“.
5. Klicken Sie unter Verbindungseigenschaften auf + Eigenschaften hinzufügen, und fügen Sie diese s3a-Eigenschaften hinzu.
6. fs.s3a.Connection.die Standardeinstellung ist 100. Wenn Ihre s3-Datensätze große Parkett-Dateien mit 100 oder mehr Spalten enthalten, muss ein Wert größer als 100 eingegeben werden. Diese Einstellung finden Sie im Dremio-Handbuch.

Name	Wert
fs.s3a.Endpunkt	*<StorageGRID S3 Endpunkt:Port>*
fs.s3a.path.style.Access	Richtig
fs.s3a.Verbindung.Maximum	*<ein Wert größer als 100>*

Beispiel Screenshot



7. Konfigurieren Sie andere Dremio-Optionen gemäß Ihren Unternehmens- oder Anwendungsanforderungen.
8. Klicken Sie auf die Schaltfläche Speichern, um diese neue Datenquelle zu erstellen.
9. Sobald die StorageGRID-Datenquelle erfolgreich hinzugefügt wurde, wird im linken Bereich eine Liste der Buckets angezeigt.
   Beispiel Screenshot



Von Angela Cheng

NetApp StorageGRID mit GitLab

NetApp hat StorageGRID mit GitLab getestet. Siehe Beispielkonfiguration für GitLab unten. Siehe "Gitlab Leitfaden zur Konfiguration von Objektspeicher" Entsprechende Details.

Beispiel für eine Objekt-Storage-Verbindung

Für Linux Package-Installationen ist dies ein Beispiel für das `connection` Einstellung im konsolidierten Formular. Bearbeiten `/etc/gitlab/gitlab.rb` Und fügen Sie die folgenden Zeilen hinzu, um die gewünschten Werte zu ersetzen:

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

Beispiele für Verfahren und APIs

Testen und Demonstrieren der S3 Verschlüsselungsoptionen auf StorageGRID

StorageGRID und die S3-API bieten verschiedene Methoden zur Verschlüsselung von Daten im Ruhezustand. Weitere Informationen finden Sie unter "Prüfen Sie die StorageGRID Verschlüsselungsmethoden".

In diesem Leitfaden werden die S3-API-Verschlüsselungsmethoden demonstriert.

Serverseitige Verschlüsselung (SSE)

Mit SSE kann der Client ein Objekt speichern und mit einem eindeutigen Schlüssel verschlüsseln, der von StorageGRID verwaltet wird. Wenn das Objekt angefordert wird, wird das Objekt durch den in StorageGRID gespeicherten Schlüssel entschlüsselt.

Beispiel: SSE

- SETZEN Sie ein Objekt mit SSE

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

- LEITEN Sie das Objekt, um die Verschlüsselung zu überprüfen

```
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": {}
}
```

- GET das Objekt

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

Serverseitige Verschlüsselung mit vom Kunden bereitgestellten Schlüsseln (SSE-C)

Mit SSE kann der Client ein Objekt speichern und mit einem eindeutigen Schlüssel verschlüsseln, der vom Client mit dem Objekt bereitgestellt wird. Wenn das Objekt angefordert wird, muss derselbe Schlüssel bereitgestellt werden, um das Objekt zu entschlüsseln und zurückzugeben.

Beispiel SSE-C

- Sie können zu Test- oder Demonstrationszwecken einen Schlüssel erstellen
  - Erstellen eines Verschlüsselungsschlüssels

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

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

- Legen Sie ein Objekt mit dem generierten Schlüssel

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

- Das Objekt in den Kopf stellen

```
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=="
}
```

Wenn Sie den Verschlüsselungsschlüssel nicht angeben, erhalten Sie einen Fehler „ein Fehler ist aufgetreten (404), wenn Sie den HeadObject-Vorgang aufrufen: Nicht gefunden“.

- Get das Objekt

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

Wenn Sie den Verschlüsselungsschlüssel nicht bereitstellen, erhalten Sie beim Aufruf des GetObject-Vorgangs einen Fehler „aufgetreten (InvalidRequest): Das Objekt wurde mit einer Form von serverseitiger Verschlüsselung gespeichert. Zum Abrufen des Objekts müssen die richtigen Parameter angegeben werden.“

Bucket-serverseitige Verschlüsselung (SSE-S3)

Mit SSE-S3 kann der Client ein Standardverschlüsselungsverhalten für alle in einem Bucket gespeicherten Objekte definieren. Die Objekte werden mit einem eindeutigen Schlüssel verschlüsselt, der von StorageGRID gemanagt wird. Wenn das Objekt angefordert wird, wird das Objekt von dem in StorageGRID gespeicherten Schlüssel entschlüsselt.

Beispiel für Bucket SSE-S3

- Erstellen eines neuen Buckets und Festlegen einer Standardverschlüsselungsrichtlinie
  - Erstellen eines neuen Buckets

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

  - Put Bucket-Verschlüsselung

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

- Legen Sie ein Objekt in den Bucket

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

- Das Objekt in den Kopf stellen

```
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": "\"82e8bfb872e778a4687a26e6c0b36bc1\"",
"ContentType": "binary/octet-stream",
"ServerSideEncryption": "AES256",
"Metadata": {}
}
```

- GET das Objekt

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

Von Aron Klein

S3-Objektsperre auf StorageGRID testen und demonstrieren

Object Lock bietet ein WORM-Modell, um das Löschen oder Überschreiben von Objekten zu verhindern. Die StorageGRID Implementierung von Objektsperrung wird auf Cohasset überprüft, um gesetzliche Vorgaben einzuhalten, den gesetzlichen Aufbewahrungs- und Compliance-Modus für Objektspeicherung zu unterstützen und standardmäßige Bucket-Aufbewahrungsrichtlinien einzuhalten.

In diesem Handbuch wird die S3-Objekt-Lock-API demonstriert.

Gesetzliche Aufbewahrungspflichten

- Object Lock Legal Hold ist ein einfacher ein-/Ausschaltstatus, der auf ein Objekt angewendet wird.

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

- Überprüfen Sie es mit EINEM GET-Vorgang.

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

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

- Deaktivieren Sie die gesetzliche Aufbewahrungspflichten

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

- Überprüfen Sie es mit EINEM GET-Vorgang.

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

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

Compliance-Modus

- Die Objektspeicherung erfolgt mit einer Aufbewahrung bis zum Zeitstempel.

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

- Überprüfen Sie den Aufbewahrungsstatus

```
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"
    }
}
```

Standardaufbewahrung

- Legen Sie den Aufbewahrungszeitraum in Tagen und Jahren als Aufbewahrungsdatum fest, das mit der API pro Objekt definiert wurde.

```
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
```

- Überprüfen Sie den Aufbewahrungsstatus

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

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

- Legen Sie ein Objekt in den Bucket

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

- Die auf dem Bucket festgelegte Aufbewahrungsdauer wird in einen Aufbewahrungszeitstempel des Objekts konvertiert.

```
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"
    }
}
```

Testen Löschen eines Objekts mit einer definierten Aufbewahrung

Objekt Lock basiert auf der Versionierung. Die Aufbewahrung ist für eine Version des Objekts definiert. Wenn versucht wird, ein Objekt mit einer definierten Aufbewahrung zu löschen, und keine Version angegeben wird, wird als aktuelle Version des Objekts eine Löschmarkierung erstellt.

- Löschen Sie das Objekt mit definierter Aufbewahrung

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

- Listen Sie die Objekte im Bucket auf

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

  - Beachten Sie, dass das Objekt nicht aufgeführt ist.

- Listen Sie Versionen auf, um die Löschen-Markierung und die ursprüngliche gesperrte Version anzuzeigen

```
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"
        }
    ]
}
```

- Löschen Sie die gesperrte Version des Objekts

```
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
```

Von Aron Klein

Beispiel für Bucket- und Gruppenrichtlinien (IAM)

Hier sind Beispiele für Bucket-Richtlinien und Gruppenrichtlinien (IAM-Richtlinien).

Gruppenrichtlinien (IAM)

Bucket-Zugriff im Home Directory-Stil

Diese Gruppenrichtlinie erlaubt Benutzern nur den Zugriff auf Objekte im Bucket mit dem Namen „username“.

```
"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}/*"
    }

  ]
}
```

Erstellung von Bucket-Objektsperrung verweigern

Diese Gruppenrichtlinie schränkt Benutzer am Erstellen eines Buckets ein, für den die Objektsperre für den Bucket aktiviert ist.

> Diese Richtlinie wird in der StorageGRID-Benutzeroberfläche nicht durchgesetzt, sie wird nur durch die S3-API durchgesetzt.

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

Aufbewahrungslimit für Objektsperre

Diese Bucket-Richtlinie beschränkt die Aufbewahrungsdauer der Objektsperre auf maximal 10 Tage

```
{
 "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"
      }
    }
   }
  ]
}
```

Benutzer daran hindern, Objekte mit VersionID zu löschen

Diese Gruppenrichtlinie schränkt Benutzer davon ab, versionierte Objekte nach VersionID zu löschen

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

Diese Bucket-Richtlinie beschränkt das Löschen versionierter Objekte durch einen Benutzer (identifiziert durch Benutzer-ID „56622399308951294926“) nach 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"
        ]
      }
    }
  ]
}
```

Bucket auf einzelnen Benutzer mit schreibgeschütztem Zugriff beschränken

Diese Richtlinie erlaubt einem einzelnen Benutzer, schreibgeschützten Zugriff auf einen Bucket zu haben und explizit allen anderen Benutzern den zugriff zu verweigert. Die Gruppierung der Ablehenserklärungen an der Spitze der Richtlinie ist eine gute Methode für eine schnellere Bewertung.

```
{
    "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/*"
            ]
        }
    ]
}
```

Beschränken Sie eine Gruppe auf ein einzelnes Unterverzeichnis (Präfix) mit Lesezugriff

Diese Richtlinie ermöglicht Mitgliedern der Gruppe schreibgeschützten Zugriff auf ein Unterverzeichnis (Präfix) innerhalb eines Buckets. Der Bucket-Name lautet „Study“ und das Unterverzeichnis lautet „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/*"
            ]
        }
    ]
}
```

Technische Berichte

NetApp StorageGRID und Big Data Analytics

Anwendungsfälle für NetApp StorageGRID

Die NetApp StorageGRID Objekt-Storage-Lösung bietet Skalierbarkeit, Datenverfügbarkeit, Sicherheit und hohe Performance. Unternehmen jeder Größe und Branche nutzen StorageGRID S3 für zahlreiche Anwendungsfälle. Sehen wir uns einige typische Szenarien an:

Big Data Analytics: StorageGRID S3 wird häufig als Data Lake verwendet, wo Unternehmen mit Tools wie Apache Spark, Splunk SmartStore und Dremio große Mengen an strukturierten und unstrukturierten Daten für Analysen speichern.

Daten-Tiering: NetApp Kunden nutzen die FabricPool Funktion von ONTAP, um Daten automatisch zwischen einem hochperformanten lokalen Tier zu StorageGRID zu verschieben. Durch Tiering wird teurer Flash-Storage für häufig abgerufene Daten frei. Kalte Daten werden auf kostengünstigem Objekt-Storage bereitgehalten. Dadurch werden Performance und Einsparungen maximiert.

Daten-Backup und Disaster Recovery: Unternehmen können StorageGRID S3 als zuverlässige und kostengünstige Lösung für Backup und Recovery kritischer Daten im Notfall einsetzen.

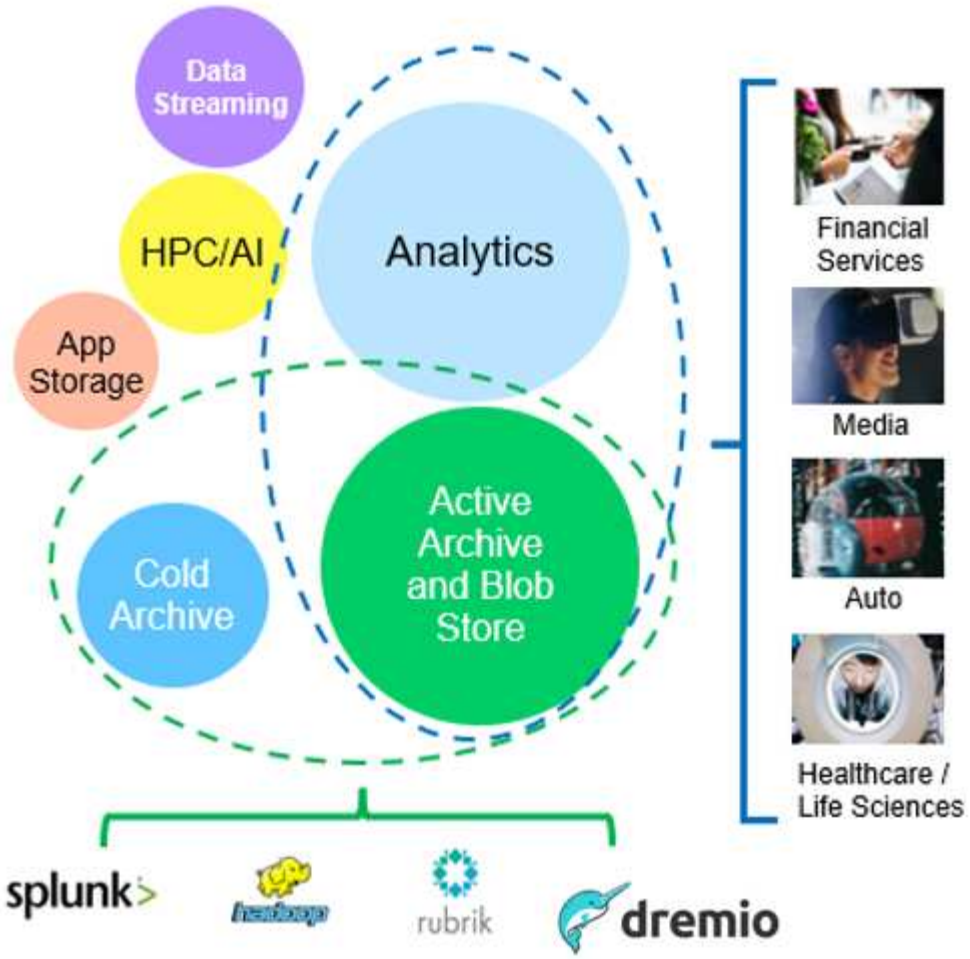
Datenspeicher für Anwendungen: StorageGRID S3 kann als Speicher-Backend für Anwendungen verwendet werden, so dass Entwickler Dateien, Bilder, Videos und andere Arten von Daten einfach speichern und abrufen können.

Inhaltsbereitstellung: StorageGRID S3 kann verwendet werden, um statische Website-Inhalte, Mediendateien und Software-Downloads für Benutzer auf der ganzen Welt zu speichern und bereitzustellen. Dabei wird die geografische Distribution und der globale Namespace von StorageGRID für eine schnelle und zuverlässige Content-Bereitstellung genutzt.

Daten-Tiering: NetApp-Kunden nutzen die ONTAP FabricPool-Funktion, um Daten automatisch zwischen einem leistungsstarken lokalen Tier zu StorageGRID zu verschieben. Durch Tiering wird teurer Flash-Storage für heiße Daten frei, während weniger oft benötigte Daten von kostengünstigem Objekt-Storage verfügbar bleiben. Dadurch werden Performance und Einsparungen maximiert.

Datenarchiv: StorageGRID bietet verschiedene Speichertypen und unterstützt Tiering zu öffentlichen, langfristigen und kostengünstigen Speicheroptionen. Damit ist es eine ideale Lösung für die Archivierung und langfristige Aufbewahrung von Daten, die für Compliance- oder historische Zwecke aufbewahrt werden müssen.

Anwendungsfälle für Objekt-Storage

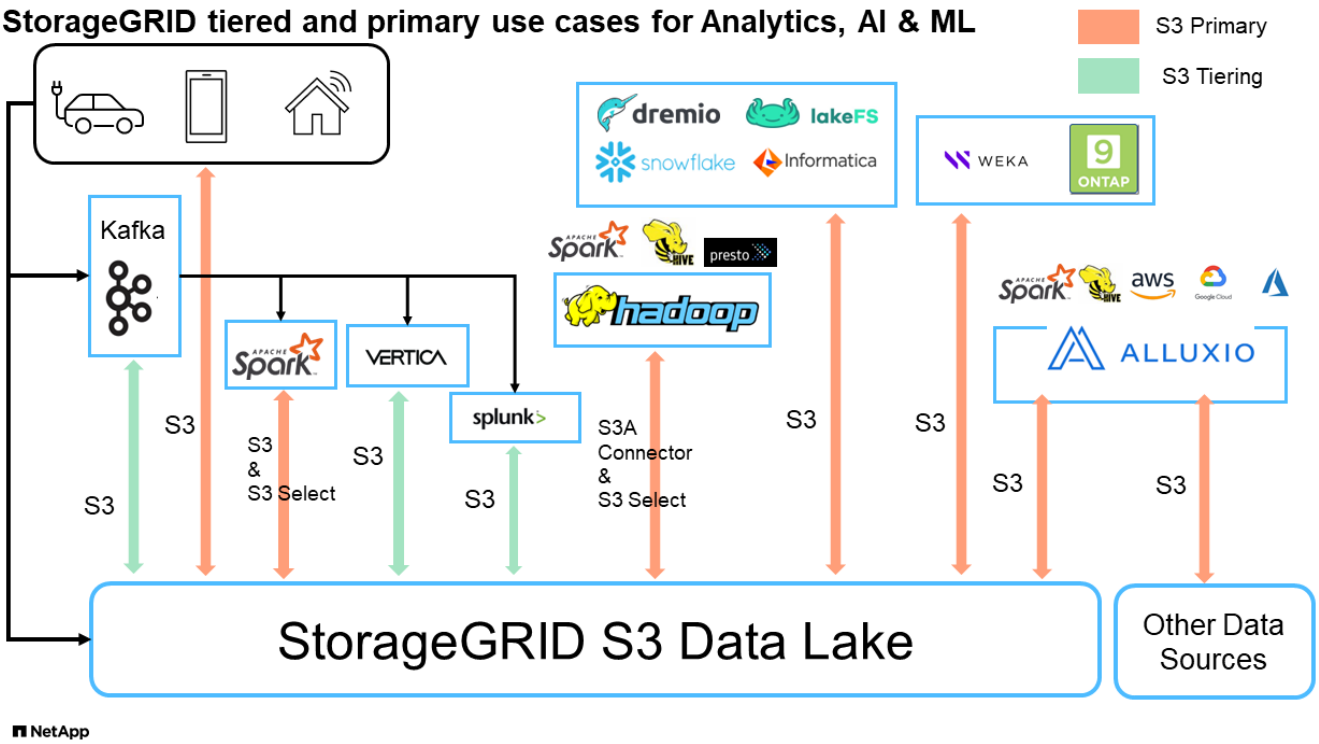


Zu den oben genannten Fällen gehört Big-Data-Analysen zu den häufigsten Nutzungsfällen und die Nutzung dieser Daten ist mit einem Aufwärtstrend verbunden.

Warum StorageGRID für Data Lakes?

- Verstärkte Zusammenarbeit: Enorme, gemeinsam genutzte Mandantenfähigkeit mit branchenüblicher API-Zugriff
- Niedrigere Betriebskosten: Einfacher Betrieb in einer einzelnen, automatisierten Scale-out-Architektur mit Selbstreparatur
- Skalierbarkeit – im Gegensatz zu herkömmlichen Hadoop- und Data-Warehouse-Lösungen entkoppelt der StorageGRID S3 Objekt-Storage den Storage von Computing und Daten. So können Unternehmen ihre Storage-Anforderungen mit wachsendem Bedarf skalieren.
- Langlebigkeit und Zuverlässigkeit: StorageGRID bietet eine Lebensdauer von 99.999999999 %, was bedeutet, dass die gespeicherten Daten sehr resistent gegen Datenverlust sind. Darüber hinaus ist Hochverfügbarkeit gewährleistet, sodass die Daten jederzeit abrufbar sind.
- Sicherheit – StorageGRID bietet verschiedene Sicherheitsfunktionen, darunter Verschlüsselung, Zugriffssteuerungsrichtlinien, Daten-Lifecycle-Management, Objektsperre und Versionierung zum Schutz der in S3 Buckets gespeicherten Daten

StorageGRID S3 Data Lakes



Welches Data Warehouse oder Data Lake eignet sich am besten für S3 Objekt-Storage

NetApp benchmarked StorageGRID mit drei Data Warehouse/Lake House Ökosysteme - Hive, Delta Lake und Dremio. "Apache Iceberg: Der Endgültige Führer" Enthält eine kurze Einführung in Data Warehouse und Data Lake House sowie vor- und Nachteile dieser beiden Architekturen.

- Benchmark-Tool - TPC-DS - https://www.tpc.org/tpcds/
- Big-Data-Ecosysteme
  - Cluster mit 5 VMs, jeweils mit 128 GB RAM und 24 vCPUs, SSD-Storage für Systemfestplatte
  - Hadoop 3.3.5 mit Hive 3.1.3 (1 Name Node + 4 Daten-Nodes)
  - Delta Lake mit Spark 3.2.0 (1 Master + 4 Workers) und Hadoop 3.3.5
  - Dremio v23 (1 Master + 4 Ausführende)
- Objekt-Storage
  - NetApp® StorageGRID® 11.6 mit 3 SG6060 + 1 SG1000 Load Balancer
  - Objektschutz - 2 Kopien
- Datenbankgröße 1.000 GB
- Cache in allen 3 Ökosystemen deaktiviert, um für jeden Abfragetest ein konsistentes Ergebnis zu erhalten.

TPC-DS verfügt über 99 komplexe SQL-Abfragen für das Abfrage-Benchmarking. Wir haben die Gesamtzahl der Minuten gemessen, um alle 99 Abfragen zu beantworten, und wir sind tiefer in die Tiefe gegangen, indem wir die Art und Anzahl der S3-Anfragen für die Analyse des Ergebnisses aufteilen. Die erste Tabelle unten zeigt die Gesamtdauer aller 99 Abfragen und die zweite Tabelle fasst die Anzahl und die Typen der S3-Anforderungen zusammen, die jedes Ökosystem an StorageGRID sendet.

TPC-DS Abfrageergebnis

Ecosystem	Hive	Delta Lake	Dremio
Storage-Ebene	NetApp® StorageGRID®	NetApp® StorageGRID®	NetApp® StorageGRID®
Laufwerkstyp	HDD	HDD	HDD
Tabellenformat	Parkett	Parkett	Parkett [1]
Datenbankgröße	1000 G	1000 G	1000 G
TPCDS 99 Abfragen Minuten gesamt	1084 [2]	55	47

[1] getestet sowohl Parkett- als auch Iceberg-Tischformat, Ergebnis ist ähnlich.

[2] Hive konnte die Abfragenummer 72 nicht abschließen.

TPC-DS Abfragen - S3 fordert Aufschlüsselung an

S3-Anfragen	Hive	Delta Lake	Dremio
GET	1,117,184	2,074,610	4,414,227
Beobachtung: Alle Reichweite ERHALTEN	80% Bereich von 2 KB bis 2 MB von 32 MB Objekten, 50 - 100 Anfragen/Sek.	73% Bereich unter 100 KB von 32-MB-Objekten, 1000 - 1400 Anforderungen/Sek.	90 % 1 MB-Bereich von Objekten mit 256 MB, 2000 bis 2300 Anforderungen/Sek.
Objekte auflisten	312,053	24,158	240
KOPF (Nicht vorhandenes Objekt)	156,027	12,103	192
KOPF (Vorhandenes Objekt)	982,126	922,732	1,845
Gesamtanforderungen	2,567,390	3,033,603	4,416,504

Vom ersten Tisch aus sehen wir Delta Lake und Dremio sind viel schneller als Hive. Aus der zweiten Tabelle geht hervor, dass Hive viele Anfragen zu S3 Listenobjekten gesendet hat, die in der Regel auf allen Objekt-Storage-Plattformen langsam sind, insbesondere dann, wenn es um einen Bucket mit vielen Objekten geht. Dies erhöht die gesamte Abfragedauer deutlich. Eine weitere Beobachtung ist, dass Dremio in der Lage war, eine hohe Anzahl von GET-Anfragen parallel zu senden, 2,000 bis 2,300 Anfragen pro Sekunde gegenüber 50 bis 100 Anfragen pro Sekunde in Hive. Hive und Hadoop S3A imitieren das Standarddateisystem und tragen zur Hive-Langsamkeit auf S3-Objekt-Storage bei.

Bei der Nutzung von Hadoop (entweder auf HDFS oder S3 Objekt-Storage) mit Hive oder Spark sind umfassende Kenntnisse zu Hadoop und Hive/Spark sowie die Interaktion der Einstellungen der einzelnen Services erforderlich – zusammen verfügen diese über mehr als 1000 Einstellungen. Sehr oft sind die Einstellungen miteinander verknüpft und können nicht allein geändert werden. Es erfordert enorm viel Zeit und Aufwand, um die optimale Kombination von Einstellungen und Werten zu finden.

Dremio ist eine Data-Lake-Engine, die mithilfe von End-to-End-Apache Arrow die Abfrage-Performance drastisch steigert. Apache Arrow bietet ein standardisiertes spaltenbasierte Speicherformat für effizientes Daten-Sharing und schnelle Analysen. Arrow verwendet einen sprachunabhängigen Ansatz, der die Notwendigkeit einer Datenserialisierung und -Deserialisierung eliminiert und die Performance und

Interoperabilität zwischen komplexen Datenprozessen und -Systemen verbessert.

Die Leistung von Dremio wird hauptsächlich durch die Rechenleistung des Dremio Clusters angetrieben. Obwohl Dremio für die S3-Objektspeicher-Verbindung den S3A-Connector von Hadoop verwendet, ist Hadoop nicht erforderlich und die meisten der fs.s3a-Einstellungen von Hadoop werden von Dremio nicht verwendet. Damit ist die Optimierung der Leistung von Dremio ganz einfach, ohne Zeit zum Erlernen und Testen verschiedener Hadoop s3a-Einstellungen zu benötigen.

Aus diesem Benchmark-Ergebnis können wir schließen, dass Big-Data-Analysesysteme mit Optimierung für S3-basierte Workloads zu einem wesentlichen Performance-Faktor werden. Dremio optimiert die Abfrageausführung, verwendet Metadaten effizient und bietet nahtlosen Zugriff auf S3-Daten. Dies ermöglicht eine bessere Performance im Vergleich zu Hive bei der Arbeit mit S3-Storage. Weitere Informationen finden Sie hier "Seite" Zur Konfiguration der Dremio S3 Datenquelle mit StorageGRID.

Unter den folgenden Links erfahren Sie mehr darüber, wie StorageGRID und Dremio gemeinsam eine moderne und effiziente Data-Lake-Infrastruktur bereitstellen und wie NetApp von Hive + HDFS auf Dremio + StorageGRID migrierte, um die Analyseeffizienz von Big Data drastisch zu steigern.

- "Mehr Performance für Big Data mit NetApp StorageGRID"
- "Moderne, leistungsstarke und effiziente Data-Lake-Infrastruktur mit StorageGRID und Dremio"
- "Wie NetApp die Kundenerfahrung mit Produktanalysen neu definiert"

Hadoop S3A-Tuning

Der Hadoop S3A Connector ermöglicht die nahtlose Interaktion zwischen Hadoop-basierten Applikationen und S3 Objektspeicher. Um die Performance bei der Arbeit mit S3-Objektspeicher zu optimieren, ist die Anpassung des Hadoop S3A Connector unerlässlich. Bevor wir uns mit der Feinabstimmung befassen, wollen wir zunächst ein grundlegendes Verständnis von Hadoop und seinen Komponenten haben.

Was ist Hadoop?

Hadoop ist ein leistungsfähiges Open-Source-Framework, das für die Verarbeitung und Speicherung großer Datenmengen entwickelt wurde. Sie ermöglicht verteilte Speicherung und parallele Verarbeitung über Cluster von Computern hinweg.

Die drei Kernkomponenten von Hadoop sind:

- **Hadoop HDFS (Hadoop Distributed File System)**: Dies verarbeitet Speicherung, teilt Daten in Blöcke auf und verteilt sie über Knoten.
- **Hadoop MapReduce**: Verantwortlich für die Verarbeitung der Daten durch Aufteilen von Aufgaben in kleinere Blöcke und deren parallele Ausführung.
- **Hadoop YARN (noch ein weiterer Resource Negotiator):** "Managt Ressourcen und plant Aufgaben effizient"

Hadoop HDFS- und S3A-Steckverbinder

HDFS ist eine wichtige Komponente des Hadoop Ecosystems und spielt eine entscheidende Rolle bei der effizienten Verarbeitung von Big Data. HDFS ermöglicht zuverlässige Speicherung und Verwaltung. Sie ermöglicht die parallele Verarbeitung und optimiert den Datenspeicher, was zu schnellerem Datenzugriff und schnelleren Analysen führt.

Bei der Big Data-Verarbeitung überzeugt HDFS durch fehlertoleranten Storage für große Datensätze. Es

erreicht dies durch Datenreplikation. Die IT kann große Mengen strukturierter und unstrukturierter Daten in einer Data Warehouse-Umgebung speichern und managen. Darüber hinaus lässt sich die Software nahtlos in führende Big Data Processing Frameworks wie Apache Spark, Hive, Pig und Flink integrieren und ermöglicht so eine skalierbare und effiziente Datenverarbeitung. Er ist mit Unix-basierten (Linux) Betriebssystemen kompatibel und somit eine ideale Wahl für Unternehmen, die Linux-basierte Umgebungen für ihre Big-Data-Verarbeitung bevorzugen.

Mit der Zeit wuchs das Datenvolumen. Daher ist es ineffizient, dem Hadoop Cluster neue Maschinen mit eigenen Computing- und Storage-Ressourcen hinzuzufügen. Lineare Skalierung führt zu Herausforderungen bei der effizienten Nutzung von Ressourcen und dem Management der Infrastruktur.

Als Antwort auf diese Herausforderungen bietet der Hadoop S3A Connector hochperformante I/O für S3 Objekt-Storage. Durch die Implementierung eines Hadoop Workflows mit S3A können Sie Objekt-Storage als Daten-Repository nutzen und Computing- und Storage-Ressourcen separat voneinander skalieren. Dadurch wiederum können Sie Computing- und Storage-Ressourcen unabhängig voneinander skalieren. Die Abkopplung von Computing und Storage erlaubt es Ihnen auch, die richtige Menge an Ressourcen für Ihre Compute-Jobs bereitzustellen und die Kapazität abhängig von der Größe des Datensatzes bereitzustellen. Somit lassen sich die Gesamtbetriebskosten für Hadoop Workflows verringern.

Hadoop S3A Connector Tuning

S3 verhält sich anders als HDFS, und einige Versuche, das Aussehen eines Filesystems beizubehalten, sind aggressiv suboptimal. Um die S3-Ressourcen möglichst effizient zu nutzen, sind sorgfältiges Tuning/Testen/Experimentieren erforderlich.

Die Hadoop Optionen in diesem Dokument basieren auf Hadoop 3.3.5, siehe "Hadoop 3.3.5 core-site.xml" Für alle verfügbaren Optionen.

Hinweis – einige Hadoop fs.s3a Einstellungen sind in den jeweiligen Hadoop Versionen unterschiedlich. Überprüfen Sie unbedingt den Standardwert Ihrer aktuellen Hadoop Version. Werden diese Einstellungen nicht in Hadoop core-site.xml angegeben, so wird als Standardwert verwendet. Sie können den Wert zur Laufzeit mithilfe der Konfigurationsoptionen Spark oder Hive überschreiben.

Sie müssen zu diesem gehen "Apache Hadoop Seite aufrufen" Um die einzelnen Optionen von fs.s3a zu verstehen. Testen Sie diese nach Möglichkeit im nicht produktiven Hadoop Cluster, um die optimalen Werte zu ermitteln.

Sie sollten lesen "Maximale Leistung bei der Arbeit mit dem S3A-Steckverbinder" Für weitere Optimierungsempfehlungen.

Sehen wir uns einige wichtige Überlegungen an:

1. Datenkomprimierung

Aktivieren Sie die StorageGRID-Komprimierung nicht. Die meisten Big-Data-Systeme verwenden Byte-Bereich get, anstatt das gesamte Objekt abzurufen. Die Verwendung von Byte Range Get mit komprimierten Objekten beeinträchtigt die GET-Performance erheblich.

2. S3A Committer

Generell wird Magic s3a Committer empfohlen. Weitere Informationen finden Sie hier "Allgemeine Seite mit den Optionen für den S3A-Committer" Um ein besseres Verständnis von Magic Committer und den damit verbundenen s3a-Einstellungen zu bekommen.

Magic Committer:

Der Magic Committer setzt speziell auf S3Guard, um konsistente Verzeichnisauflistungen im S3-Objektspeicher zu bieten.

Mit konsistenten S3 (was jetzt der Fall ist), kann der Magic Committer sicher mit jedem S3-Bucket verwendet werden.

Auswahl und Experimentieren:

Je nach Anwendungsfall können Sie zwischen dem Staging Committer (der auf einem Cluster HDFS-Dateisystem basiert) und dem Magic Committer wählen.

Experimentieren Sie mit beiden, um zu ermitteln, welche Lösung am besten zu Ihrem Workload und Ihren Anforderungen passt.

Zusammenfassend stellen die S3A Committers eine Lösung für die grundlegende Herausforderung dar, die ein konsistentes, leistungsstarkes und zuverlässiges Leistungsengagement für S3 darstellt. Das interne Design gewährleistet einen effizienten Datentransfer bei gleichzeitiger Wahrung der Datenintegrität.

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, Größe des Verbindungspools und Blockgröße

- Jeder **S3A**-Client, der mit einem einzelnen Bucket interagiert, hat einen eigenen dedizierten Pool von offenen HTTP 1.1-Verbindungen und Threads für Upload- und Kopiervorgänge.
- "Sie können diese Poolgrößen so anpassen, dass ein ausgewogenes Verhältnis zwischen Leistung und Speicher-/Thread-Nutzung erzielt wird".
- Beim Hochladen von Daten in S3 werden sie in Blöcke unterteilt. Die standardmäßige Blockgröße beträgt 32 MB. Sie können diesen Wert anpassen, indem Sie die Eigenschaft fs.s3a.Block.size festlegen.
- Größere Blockgrößen verbessern die Performance beim Hochladen großer Daten, da sich der Managementaufwand für mehrteilige Teile während des Uploads verringert. Der empfohlene Wert ist 256 MB oder höher für große Datensätze.

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. Mehrteiliges Hochladen

s3a-Committer **Always** Verwenden Sie MPU (mehrteilige Uploads) zum Hochladen von Daten in s3-Buckets. Dies ist erforderlich, um Folgendes zu ermöglichen: Task Failure, spekulative Ausführung von Aufgaben und Job Abbrüche vor Commit. Hier sind einige wichtige Spezifikationen für mehrteilige Uploads:

- Maximale Objektgröße: 5 tib (Terabyte).
- Maximale Anzahl von Teilen pro Upload: 10,000.
- Teilenummern: Von 1 bis 10,000 (inklusive).
- Größe des Teils: Zwischen 5 MiB und 5 gib. Insbesondere gibt es keine Mindestgröße für den letzten Teil Ihres mehrteiligen Uploads.

Die Verwendung kleinerer Teilgröße für S3-Multipart-Uploads hat sowohl vor- als auch Nachteile.

Vorteile:

- Schnelle Wiederherstellung von Netzwerkproblemen: Wenn Sie kleinere Teile hochladen, werden die Auswirkungen des Neustarts eines fehlgeschlagenen Uploads aufgrund eines Netzwerkfehlers minimiert. Wenn ein Teil fehlschlägt, müssen Sie nur dieses Teil neu hochladen, nicht das gesamte Objekt.

- Bessere Parallelisierung: Mehr Teile können parallel hochgeladen werden, wobei Multi-Threading oder gleichzeitige Verbindungen genutzt werden können. Diese Parallelisierung verbessert die Performance, insbesondere bei der Verarbeitung großer Dateien.

Nachteil:

- Netzwerk-Overhead: Kleinere Teilegröße bedeutet, dass mehr Teile hochgeladen werden müssen, jedes Teil benötigt eine eigene HTTP-Anforderung. Mehr HTTP-Anfragen erhöhen den Overhead beim Initiieren und Abschließen einzelner Anfragen. Die Verwaltung einer großen Anzahl von Kleinteilen kann die Leistung beeinträchtigen.
- Komplexität: Die Verwaltung der Bestellung, die Nachverfolgung von Teilen und die Sicherstellung erfolgreicher Uploads können umständlich sein. Wenn der Upload abgebrochen werden muss, müssen alle bereits hochgeladenen Teile nachverfolgt und gelöscht werden.

Für Hadoop wird eine Teilegröße von 256 MB oder höher für fs.s3a.Multipart.size empfohlen. Stellen Sie immer den Wert fs.s3a.mutlipart.threshold auf 2 x fs.s3a.multipart.size ein. Beispiel: Fs.s3a.multipart.size = 256M, fs.s3a.mutlipart.threshold sollte 512M sein.

Größere Teilegröße für großen Datensatz verwenden Es ist wichtig, eine Teilegröße zu wählen, die diese Faktoren auf der Grundlage Ihres spezifischen Anwendungsfalls und der Netzwerkbedingungen ausgleicht.

Ein mehrteiliges Hochladen ist ein "Prozess in drei Schritten":

1. Der Upload wird gestartet, StorageGRID gibt eine Upload-ID zurück.
2. Die Objektteile werden mit der Upload-ID hochgeladen.
3. Sobald alle Objektteile hochgeladen sind, sendet die komplette mehrteilige Upload-Anfrage mit Upload-ID. StorageGRID erstellt das Objekt aus den hochgeladenen Teilen, und der Client kann auf das Objekt zugreifen.

Wenn die Anfrage zum vollständigen Hochladen mehrerer Teile nicht erfolgreich gesendet wird, bleiben die Teile in StorageGRID und erstellen kein Objekt. Dies geschieht, wenn Jobs unterbrochen, fehlgeschlagen oder abgebrochen werden. Die Teile verbleiben im Raster, bis der Upload mehrerer Teile abgeschlossen ist oder abgebrochen wird oder StorageGRID diese Teile löscht, wenn 15 Tage nach dem Upload vergangen sind. Wenn sich viele (einige Hunderttausend bis Millionen) mehrteilige Uploads in einem Bucket befinden und Hadoop ‘list-Multipart-Uploads’ sendet (diese Anfrage filtert nicht nach Upload-id), kann die Bearbeitung der Anfrage sehr viel Zeit in Anspruch nehmen oder eventuell eine bestimmte Zeit in Anspruch nehmen. Sie können die Einstellung fs.s3a.mutlipart.purge mit dem entsprechenden Wert fs.s3a.Multipart.purge.age (z. B. 5 bis 7 Tage, verwenden Sie den Standardwert 86400, d. h. 1 Tag) auf true setzen. Oder wenden Sie sich an den NetApp Support, um die Situation zu untersuchen.

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. Pufferschreibdaten im Speicher

Zur Verbesserung der Performance können Sie Schreibdaten vor dem Hochladen in S3 zwischenspeichern. Dies kann die Anzahl kleiner Schreibvorgänge reduzieren und die Effizienz verbessern.

Option	Meaning	Default (Hadoop 3.3.5)
fs.s3a.fast.upload.buffer	The buffering mechanism to for data being written. Values: disk, array, bytebuffer. "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

S3 und HDFS funktionieren jedoch auf unterschiedliche Weise. Um die S3-Ressourcen optimal zu nutzen, sind

sorgfältiges Tuning/Test/Experiment nötig.

NetApp StorageGRID-Blogs

Hier finden Sie einige der großartigen NetApp StorageGRID Blogs:

- Mai 10: "Lab on Demand ist Ihr bestes Vertriebstool für StorageGRID"
- Mai 24: "Modernisieren Sie Ihre Analyse-Workloads mit NetApp und Alluxio"
- Mai 26: "StorageGRID – Speichern und Managen von Backup- und Replizierungsdaten On-Premises"
- Juni 9: "Nutzen Sie Hadoop S3A-Connector von Cloudera mit StorageGRID"
- Juli 26: "Die wachsende Liste validierter Partnerlösungen für StorageGRID nimmt zu"
- August 5: "NetApp StorageGRID erhält eine Common Criteria Security-Zertifizierung"
- August 16: "Integration von StorageGRID in den Open-Source-ELK Stack für mehr Benutzerfreundlichkeit"
- August 17: "Es beginnt alles mit Objekt sperren… Aufbau eines S3-Storage-Ecosystems für kritische Backup-Applikationen"
- August 23: "Data-Lake-Lösung auf Basis von StorageGRID"
- September: "Nehmen Sie diese Metriken und graten Sie sie"
- September 19: "Unterstützung von DataLock und Ransomware-Schutz für StorageGRID"
- September 26: "NetApp StorageGRID für Service Provider"
- Okt. 5: "Auftauen Sie Ihre Daten auf StorageGRID für Snowflake"
- Okt. 5: "NetApp Cloud Insights bietet StorageGRID Galerie-Dashboards"
- Nov. 7: "Unterstützung von StorageGRID und ONTAP S3: Unterschiede, Gemeinsamkeiten und Integration"
- Nov. 23: "Erklärbare AI mit MLOps von NetApp und Modzy"
- Dezember 6: "StorageGRID erhält die KPMG-Compliance-Zertifizierung"
- Januar 16: "StorageGRID erneuert die NF203- und ISO/IEC 25051-Compliance-Zertifizierung"
- Januar 18: "StorageGRID S3 Object Lock validiert für Veritas NetBackup"
- Februar 14: "Was haben Schokolade, Skifahren, Uhren und Mainframes gemeinsam?"
- März 14: "So sichern Sie EHR-Datenbanken von Epic Systems mit einem Befehl in einer 3:2:1-konformen Architektur"
- März 30: "Mit BlueXP schützen Sie Epic EHR durch eine Backup-Richtlinie, die 3:2:1-konform ist"
- März 30: "Mountpoint für Amazon S3 Alpha-Version mit StorageGRID"
- Mai 16: "Neuerungen bei der StorageGRID Objekt-Storage-Produktfamilie"
- Mai 16: "Wir stellen vor: StorageGRID 11.7 und die neue All-Flash Objekt-Storage Appliance SGF6112"
- August 30: "Bereitstellungspunkt für Amazon S3 File System ist jetzt allgemein verfügbar"
- September: "Mithilfe von Cloud Insights können Sie Protokolle mithilfe von Fluent Bit überwachen und erfassen"
- Okt. 17: "Weiter aus Hadoop: Datenanalysen modernisieren mit Dremio und StorageGRID"
- Nov. 7: "Spectra Logic On-Premises-Gletscher mit StorageGRID"
- Dezember 12: "Big Data Analytics auf StorageGRID: Dremio arbeitet 23-mal schneller als Apache Hive"
- Februar 2: "Wir stellen vor: Die StorageGRID + LakeFS Lösung im Überblick"

- Februar 16: "Neu: StorageGRID 11.8: Mehr Sicherheit, Einfachheit und Benutzerfreundlichkeit"
- Februar 16: "Wir stellen vor: StorageGRID 11.8"

NetApp StorageGRID-Dokumentation

Die vollständige Dokumentation zu jeder NetApp StorageGRID Version finden Sie hier:

- "StorageGRID Appliances"
- "StorageGRID 11.8"
- "StorageGRID 11.7"
- "StorageGRID 11.6"
- "StorageGRID 11.5"
- "StorageGRID 11.4"
- "StorageGRID 11.3"
- "StorageGRID 11.2"

Rechtliche Hinweise

Rechtliche Hinweise ermöglichen den Zugriff auf Copyright-Erklärungen, Marken, Patente und mehr.

Urheberrecht

"https://www.netapp.com/company/legal/copyright/"

Marken

NetApp, das NETAPP Logo und die auf der NetApp Markenseite aufgeführten Marken sind Marken von NetApp Inc. Andere Firmen- und Produktnamen können Marken der jeweiligen Eigentümer sein.

"https://www.netapp.com/company/legal/trademarks/"

Patente

Eine aktuelle Liste der NetApp Patente finden Sie unter:

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

Datenschutzrichtlinie

"https://www.netapp.com/company/legal/privacy-policy/"

Open Source

In den Benachrichtigungsdateien finden Sie Informationen zu Urheberrechten und Lizenzen von Drittanbietern, die in der NetApp Software verwendet werden.

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

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

Copyright-Informationen

Copyright © 2024 NetApp. Alle Rechte vorbehalten. Gedruckt in den USA. Dieses urheberrechtlich geschützte Dokument darf ohne die vorherige schriftliche Genehmigung des Urheberrechtsinhabers in keiner Form und durch keine Mittel – weder grafische noch elektronische oder mechanische, einschließlich Fotokopieren, Aufnehmen oder Speichern in einem elektronischen Abrufsystem – auch nicht in Teilen, vervielfältigt werden.

Software, die von urheberrechtlich geschütztem NetApp Material abgeleitet wird, unterliegt der folgenden Lizenz und dem folgenden Haftungsausschluss:

DIE VORLIEGENDE SOFTWARE WIRD IN DER VORLIEGENDEN FORM VON NETAPP ZUR VERFÜGUNG GESTELLT, D. H. OHNE JEGLICHE EXPLIZITE ODER IMPLIZITE GEWÄHRLEISTUNG, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDE GEWÄHRLEISTUNG DER MARKTGÄNGIGKEIT UND EIGNUNG FÜR EINEN BESTIMMTEN ZWECK, DIE HIERMIT AUSGESCHLOSSEN WERDEN. NETAPP ÜBERNIMMT KEINERLEI HAFTUNG FÜR DIREKTE, INDIREKTE, ZUFÄLLIGE, BESONDERE, BEISPIELHAFTE SCHÄDEN ODER FOLGESCHÄDEN (EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE BESCHAFFUNG VON ERSATZWAREN ODER -DIENSTLEISTUNGEN, NUTZUNGS-, DATEN- ODER GEWINNVERLUSTE ODER UNTERBRECHUNG DES GESCHÄFTSBETRIEBS), UNABHÄNGIG DAVON, WIE SIE VERURSACHT WURDEN UND AUF WELCHER HAFTUNGSTHEORIE SIE BERUHEN, OB AUS VERTRAGLICH FESTGELEGTER HAFTUNG, VERSCHULDENSUNABHÄNGIGER HAFTUNG ODER DELIKTSHAFTUNG (EINSCHLIESSLICH FAHRLÄSSIGKEIT ODER AUF ANDEREM WEGE), DIE IN IRGENDEINER WEISE AUS DER NUTZUNG DIESER SOFTWARE RESULTIEREN, SELBST WENN AUF DIE MÖGLICHKEIT DERARTIGER SCHÄDEN HINGEWIESEN WURDE.

NetApp behält sich das Recht vor, die hierin beschriebenen Produkte jederzeit und ohne Vorankündigung zu ändern. NetApp übernimmt keine Verantwortung oder Haftung, die sich aus der Verwendung der hier beschriebenen Produkte ergibt, es sei denn, NetApp hat dem ausdrücklich in schriftlicher Form zugestimmt. Die Verwendung oder der Erwerb dieses Produkts stellt keine Lizenzierung im Rahmen eines Patentrechts, Markenrechts oder eines anderen Rechts an geistigem Eigentum von NetApp dar.

Das in diesem Dokument beschriebene Produkt kann durch ein oder mehrere US-amerikanische Patente, ausländische Patente oder anhängige Patentanmeldungen geschützt sein.

ERLÄUTERUNG ZU „RESTRICTED RIGHTS“: Nutzung, Vervielfältigung oder Offenlegung durch die US-Regierung unterliegt den Einschränkungen gemäß Unterabschnitt (b)(3) der Klausel „Rights in Technical Data – Noncommercial Items“ in DFARS 252.227-7013 (Februar 2014) und FAR 52.227-19 (Dezember 2007).

Die hierin enthaltenen Daten beziehen sich auf ein kommerzielles Produkt und/oder einen kommerziellen Service (wie in FAR 2.101 definiert) und sind Eigentum von NetApp, Inc. Alle technischen Daten und die Computersoftware von NetApp, die unter diesem Vertrag bereitgestellt werden, sind gewerblicher Natur und wurden ausschließlich unter Verwendung privater Mittel entwickelt. Die US-Regierung besitzt eine nicht ausschließliche, nicht übertragbare, nicht unterlizenzierbare, weltweite, limitierte unwiderrufliche Lizenz zur Nutzung der Daten nur in Verbindung mit und zur Unterstützung des Vertrags der US-Regierung, unter dem die Daten bereitgestellt wurden. Sofern in den vorliegenden Bedingungen nicht anders angegeben, dürfen die Daten ohne vorherige schriftliche Genehmigung von NetApp, Inc. nicht verwendet, offengelegt, vervielfältigt, geändert, aufgeführt oder angezeigt werden. Die Lizenzrechte der US-Regierung für das US-Verteidigungsministerium sind auf die in DFARS-Klausel 252.227-7015(b) (Februar 2014) genannten Rechte beschränkt.

Markeninformationen

NETAPP, das NETAPP Logo und die unter http://www.netapp.com/TM aufgeführten Marken sind Marken von NetApp, Inc. Andere Firmen und Produktnamen können Marken der jeweiligen Eigentümer sein.