# Activation de StorageGRID dans votre environnement

How to enable StorageGRID in your environment

NetApp
April 26, 2024

# Sommaire

# Activation de StorageGRID dans votre environnement

# Des solutions tierces validées

## Solutions tierces validées : présentation

NetApp, en collaboration avec ses partenaires, a validé ces solutions pour StorageGRID. Consultez les informations de cette section pour savoir quelles solutions ont été validées et pour obtenir des instructions supplémentaires, le cas échéant.

Ensemble, nous pouvons accélérer l'innovation, sensibiliser davantage le marché et augmenter les ventes grâce à la création de solutions NetApp testées et de pointe. "Devenir un partenaire Alliance".

## Solutions tierces validées par StorageGRID 11.8

L'utilisation des solutions tierces suivantes a été validée avec StorageGRID 11.8.
Si la solution que vous recherchez n'est pas répertoriée, contactez votre représentant de compte NetApp.

### Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Almuxio
- Apache Kafka
- Point de montage AWS
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- Collibra (qualité minimale des données de Collibra version 2024.02)
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Disquette de données
- Dremio
- Emam
- Archive d'objets Fujifilm
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect plus

- Interica
- Komprise
- Clusters de Big Data Microsoft SQL Server
- Modèle 9
- Modzy
- La promenade au clair de lune Universal
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- PixitMedia ngenea
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- Reveille v10 construire 220706 ou plus
- CDM Rubrik
- s3a
- Signiant
- Flocon de neige
- Spectra Logic On-sur-site Glacier
- SmartStore Splunk
- Pour un stockage simplifié
- Trino
- Vernis Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric
- Weka v3.10 ou ultérieure

### Solutions tierces validées sur StorageGRID avec verrouillage objet

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Fonctionnalité CommVault 11 version 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 et versions ultérieures

### Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d'archivage
- Communications d'axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d'objets Fujifilm
- Archive GE Centricity Enterprise
- Gitlab
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Solutions tierces validées par StorageGRID 11.7

L'utilisation des solutions tierces suivantes a été validée avec StorageGRID 11.7. + si la solution que vous recherchez n'est pas répertoriée, contactez votre ingénieur commercial NetApp.

## Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Almuxio
- Apache Kafka
- Point de montage AWS
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- Collibra (qualité minimale des données de Collibra version 2024.02)
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Disquette de données
- Dremio
- Emam
- Archive d’objets Fujifilm
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect plus
- Interica
- Komprise
- Clusters de Big Data Microsoft SQL Server
- Modèle 9
- Modzy
- La promenade au clair de lune Universal
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura

- PixitMedia ngenea
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- Reveille v10 construire 220706 ou plus
- CDM Rubrik
- s3a
- Signiant
- Flocon de neige
- Spectra Logic On-sur-site Glacier
- SmartStore Splunk
- Pour un stockage simplifié
- Trino
- Vernis Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric
- Weka v3.10 ou ultérieure

## Solutions tierces validées sur StorageGRID avec verrouillage objet

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Fonctionnalité CommVault 11 version 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 et versions ultérieures

## Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d’archivage
- Communications d’axe
- Congruity360

- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d’objets Fujifilm
- Archive GE Centricity Enterprise
- Gitlab
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Des solutions tierces validées pour StorageGRID 11.6

Les solutions tierces suivantes ont été validées pour une utilisation avec StorageGRID 11.6. + si la solution que vous recherchez n’est pas répertoriée, contactez votre ingénieur commercial NetApp.

## Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Almuxio
- Apache Kafka
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Disquette de données
- Dremio

- Emam
- Archive d’objets Fujifilm
- GitHub Enterprise Server
- IBM FileNet
- IBM Spectrum Protect plus
- Interica
- Komprise
- Clusters de Big Data Microsoft SQL Server
- Modèle 9
- Modzy
- La promenade au clair de lune Universal
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- PixitMedia ngenea
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- Reveille v10 construire 220706 ou plus
- CDM Rubrik
- s3a
- Signiant
- Flocon de neige
- Spectra Logic On-sur-site Glacier
- SmartStore Splunk
- Pour un stockage simplifié
- Trino
- Vernis Enterprise 6.0.4
- Veeam 12
- Veritas Enterprise Vault 14
- Veritas NetBackup 8.0
- Vertica 10.x

- Vidispine
- Virtalica StorageFabric
- Weka v3.10 ou ultérieure

### Solutions tierces validées sur StorageGRID avec verrouillage objet

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Fonctionnalité CommVault 11 version 26
- IBM FileNet
- OpenText Documentum 21.4
- Veeam 12
- Veritas Enterprise Vault 14.2.2
- Veritas NetBackup 10.1.1 et versions ultérieures

### Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d’archivage
- Communications d’axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d’objets Fujifilm
- Archive GE Centricity Enterprise
- Gitlab
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Des solutions tierces validées pour StorageGRID 11.5

Les solutions tierces suivantes ont été validées pour une utilisation avec StorageGRID 11.5. + si la solution que vous recherchez n’est pas répertoriée, contactez votre ingénieur

commercial NetApp.

## Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Almuxio
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Interica
- Komprise
- La promenade au clair de lune Universal
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Documentum 21.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- CDM Rubrik
- s3a
- Signiant
- SmartStore Splunk
- Trino
- Vernis Enterprise 6.0.4
- Veeam 11
- Veritas Enterprise Vault 11

- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine
- Virtalica StorageFabric

### Solutions tierces validées sur StorageGRID avec verrouillage objet

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- OpenText Documentum 21.4
- Veeam 11

### Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d'archivage
- Communications d'axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d'objets Fujifilm
- Archive GE Centricity Enterprise
- Gitlab
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Des solutions tierces validées pour StorageGRID 11.4

Les solutions tierces suivantes ont été validées pour une utilisation avec StorageGRID 11.4. + si la solution que vous recherchez n'est pas répertoriée, contactez votre ingénieur commercial NetApp.

## Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Interica
- Komprise
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText InfoArchive 16 EP7
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- CDM Rubrik
- Signiant
- SmartStore Splunk
- Vernis Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vertica 10.x
- Vidispine

### Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d'archivage
- Communications d'axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d'objets Fujifilm
- Archive GE Centricity Enterprise
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

## Des solutions tierces validées pour StorageGRID 11.3

Les solutions tierces suivantes ont été validées pour une utilisation avec StorageGRID 11.3. + si la solution que vous recherchez n'est pas répertoriée, contactez votre ingénieur commercial NetApp.

### Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données

- DétendX
- Interica
- Komprise
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- Passerelle d’archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- SmartStore Splunk
- Vernis Enterprise 6.0.4
- Veeam 9.5.4
- Veritas Enterprise Vault 11
- Veritas Enterprise Vault 12
- Veritas NetBackup 8.0
- Vidispine

## Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d’archivage
- Communications d’axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d’objets Fujifilm
- Archive GE Centricity Enterprise
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH

- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Des solutions tierces validées pour StorageGRID 11.2

Les solutions tierces suivantes ont été validées pour une utilisation avec StorageGRID 11.2. + si la solution que vous recherchez n'est pas répertoriée, contactez votre ingénieur commercial NetApp.

## Solutions tierces validées sur StorageGRID

Ces solutions ont été testées en collaboration avec les partenaires respectifs.

- Actifio
- Bridgesdor
- Cantemo
- Collaboration de contenu Citrix
- CommVault 11
- Portail CTERA 6
- Dalet
- Datobi
- Le stockage dynamique des données
- DétendX
- Interica
- Komprise
- BIEN
- Nasuni
- OpenText Documentum 16.4
- OpenText Media Management 16.5 avec CyanGate Cloud
- Panzura
- Passerelle d'archivage de points 2.0
- Point Storage Manager 6.4
- Primestream
- StorNext de Quantum 5.4.0.1
- Rubrik CDM 5.0.1 p1-1342
- Signiant
- SmartStore Splunk
- Vernis Enterprise 6.0.4

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

## Solutions tierces prises en charge sur StorageGRID

Ces solutions ont été testées.

- Logiciel d’archivage
- Communications d’axe
- Congruity360
- DataFrameworks
- Plate-forme DIVA EcoDigital
- Encoding.com
- Archive d’objets Fujifilm
- Archive GE Centricity Enterprise
- Acuo Hyland
- IBM Aspera
- Systèmes Milestone
- RSSI
- Moteur REACH
- SilverTrak
- SoftNAS
- QSTAR
- Velasea

# Guides des fonctionnalités des produits

## Création d'un pool de stockage cloud pour AWS ou Google Cloud

Vous pouvez utiliser un pool de stockage cloud pour déplacer des objets StorageGRID vers un compartiment S3 externe. Le compartiment externe peut appartenir à Amazon S3 (AWS) ou à Google Cloud.

**Ce dont vous avez besoin**

- StorageGRID 11.6 a été configuré.
- Vous avez déjà configuré un compartiment S3 externe sur AWS ou Google Cloud.

**Étapes**

1. Dans Grid Manager, accédez à **ILM** > **Storage pools**.
2. Dans la section Cloud Storage pools de la page, sélectionnez **Create**.

   La fenêtre contextuelle Créer un pool de stockage cloud s'affiche.

3. Entrez un nom d'affichage.
4. Sélectionnez **Amazon S3** dans la liste déroulante Type de fournisseur.

   Ce type de fournisseur fonctionne pour AWS S3 ou Google Cloud.

5. Entrez l'URI du compartiment S3 à utiliser pour le pool de stockage cloud.

   Deux formats sont autorisés :

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

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

6. Entrez le nom du compartiment S3.

   Le nom que vous spécifiez doit correspondre exactement au nom du compartiment S3. Sinon, la création du pool de stockage cloud échoue. Vous ne pouvez pas modifier cette valeur après l'enregistrement du pool de stockage cloud.

7. Vous pouvez également saisir l'ID de clé d'accès et la clé d'accès secrète.
8. Sélectionnez **ne pas vérifier le certificat** dans la liste déroulante.
9. Cliquez sur **Enregistrer**.

**Résultat attendu**

Assurez-vous qu'un pool de stockage cloud a été créé pour Amazon S3 ou Google Cloud.

*Par Jonathan Wong*

# Création d’un pool de stockage cloud pour le stockage Azure Blob

Vous pouvez utiliser un pool de stockage cloud pour déplacer des objets StorageGRID vers un conteneur Azure externe.

**Ce dont vous avez besoin**

- StorageGRID 11.6 a été configuré.
- Vous avez déjà configuré un conteneur Azure externe.

**Étapes**

1. Dans Grid Manager, accédez à **ILM** > **Storage pools**.
2. Dans la section Cloud Storage pools de la page, sélectionnez **Create**.

   La fenêtre contextuelle Créer un pool de stockage cloud s’affiche.

3. Entrez un nom d’affichage.
4. Sélectionnez **Azure Blob Storage** dans la liste déroulante Type de fournisseur.
5. Entrez l’URI du compartiment S3 à utiliser pour le pool de stockage cloud.

   Deux formats sont autorisés :

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

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

6. Entrez le nom du conteneur Azure.

   Le nom que vous spécifiez doit correspondre exactement au nom du conteneur Azure. Sinon, la création du pool de stockage cloud échoue. Vous ne pouvez pas modifier cette valeur après l’enregistrement du pool de stockage cloud.

7. Vous pouvez également saisir le nom de compte et la clé de compte associés du conteneur Azure pour l’authentification.
8. Sélectionnez **ne pas vérifier le certificat** dans la liste déroulante.
9. Cliquez sur **Enregistrer**.

**Résultat attendu**

Confirmation de la création d’un pool de stockage cloud pour Azure Blob Storage

*Par Jonathan Wong*

# Utilisation d’un pool de stockage cloud pour la sauvegarde

Vous pouvez créer une règle ILM pour déplacer des objets dans Cloud Storage Pool à des fins de sauvegarde.

**Ce dont vous avez besoin**

- StorageGRID 11.6 a été configuré.
- Vous avez déjà configuré un conteneur Azure externe.

**Étapes**

1. Dans Grid Manager, accédez à **ILM** > **règles** > **Créer**.
2. Entrez une description.
3. Entrez un critère pour déclencher la règle.
4. Cliquez sur **Suivant**.
5. Répliquez l'objet dans les nœuds de stockage.
6. Ajoutez une règle de placement.
7. Réplication de l'objet vers le pool de stockage cloud
8. Cliquez sur **Suivant**.
9. Cliquez sur **Enregistrer**.

**Résultat attendu**

Vérifiez que le diagramme de conservation affiche les objets stockés localement dans StorageGRID et dans un Cloud Storage Pool pour la sauvegarde.

Confirmez que, lorsque la règle ILM est déclenchée, une copie existe dans le pool de stockage cloud et vous pouvez récupérer l'objet localement sans effectuer de restauration d'objet.

*Par Jonathan Wong*

# Configurez le service d'intégration de recherche StorageGRID

Ce guide fournit des instructions détaillées sur la configuration du service d'intégration de recherche NetApp StorageGRID 11.6 avec Amazon OpenSearch Service ou avec Elasticsearch sur site.

## Introduction

StorageGRID prend en charge trois types de services de plateforme.

- **Réplication StorageGRID CloudMirror**. Mettre en miroir des objets spécifiques d'un compartiment StorageGRID vers une destination externe spécifiée.
- **Notifications**. Notifications d'événements par compartiment pour envoyer des notifications sur des actions spécifiques réalisées sur des objets vers un Amazon simple notification Service (Amazon SNS) externe spécifié.
- **Service d'intégration de recherche**. Envoyez les métadonnées d'objet S3 (simple Storage Service) à un index Elasticsearch spécifique où vous pouvez rechercher ou analyser les métadonnées à l'aide du service externe.

Les services de plateforme sont configurés par le locataire S3 via l'interface du gestionnaire des locataires. Pour plus d'informations, voir "Considérations relatives à l'utilisation des services de plate-forme".

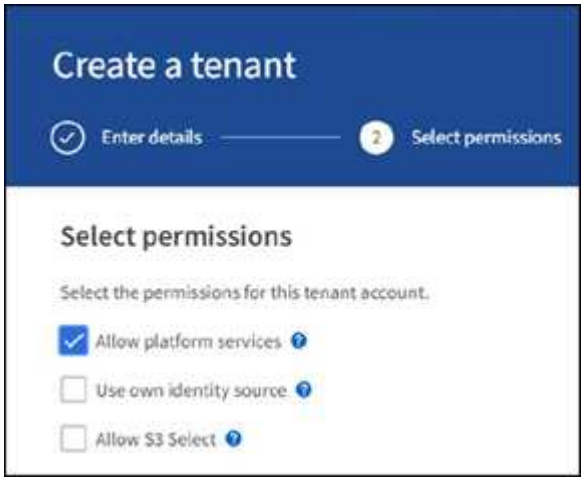
Ce document est un supplément au "Guide des locataires StorageGRID 11.6" et fournit des instructions

détaillées et des exemples de configuration du terminal et des compartiments pour les services d'intégration de la recherche. Les instructions d'installation d'Amazon Web Services (AWS) ou de Elasticsearch sur site indiquées ici sont fournies à des fins de test ou de démonstration uniquement.

Les participants doivent maîtriser Grid Manager et le Gestionnaire de locataires, et avoir accès au navigateur S3 pour effectuer des opérations de chargement (PUT) et de téléchargement (GET) de base pour les tests d'intégration de la recherche StorageGRID.

## Créez des locataires et activez les services de plateforme

1. Créez un locataire S3 à l'aide de Grid Manager, entrez un nom d'affichage et sélectionnez le protocole S3.
2. Sur la page d'autorisation, sélectionnez l'option Autoriser les services de plate-forme. Vous pouvez également sélectionner d'autres autorisations, si nécessaire.




3. Configurez le mot de passe initial de l'utilisateur root du locataire ou, si la fédération identifier est activée sur la grille, sélectionnez le groupe fédéré disposant d'une autorisation d'accès racine pour configurer le compte du locataire.
4. Cliquez sur se connecter en tant que racine et sélectionnez godet : créer et gérer des godets.

   Vous accédez alors à la page Gestionnaire de locataires.

5. Dans le Gestionnaire des locataires, sélectionnez Mes clés d'accès pour créer et télécharger la clé d'accès S3 pour des tests ultérieurs.

## Services d'intégration de recherche avec Amazon OpenSearch

### Configuration du service Amazon OpenSearch (anciennement Elasticsearch

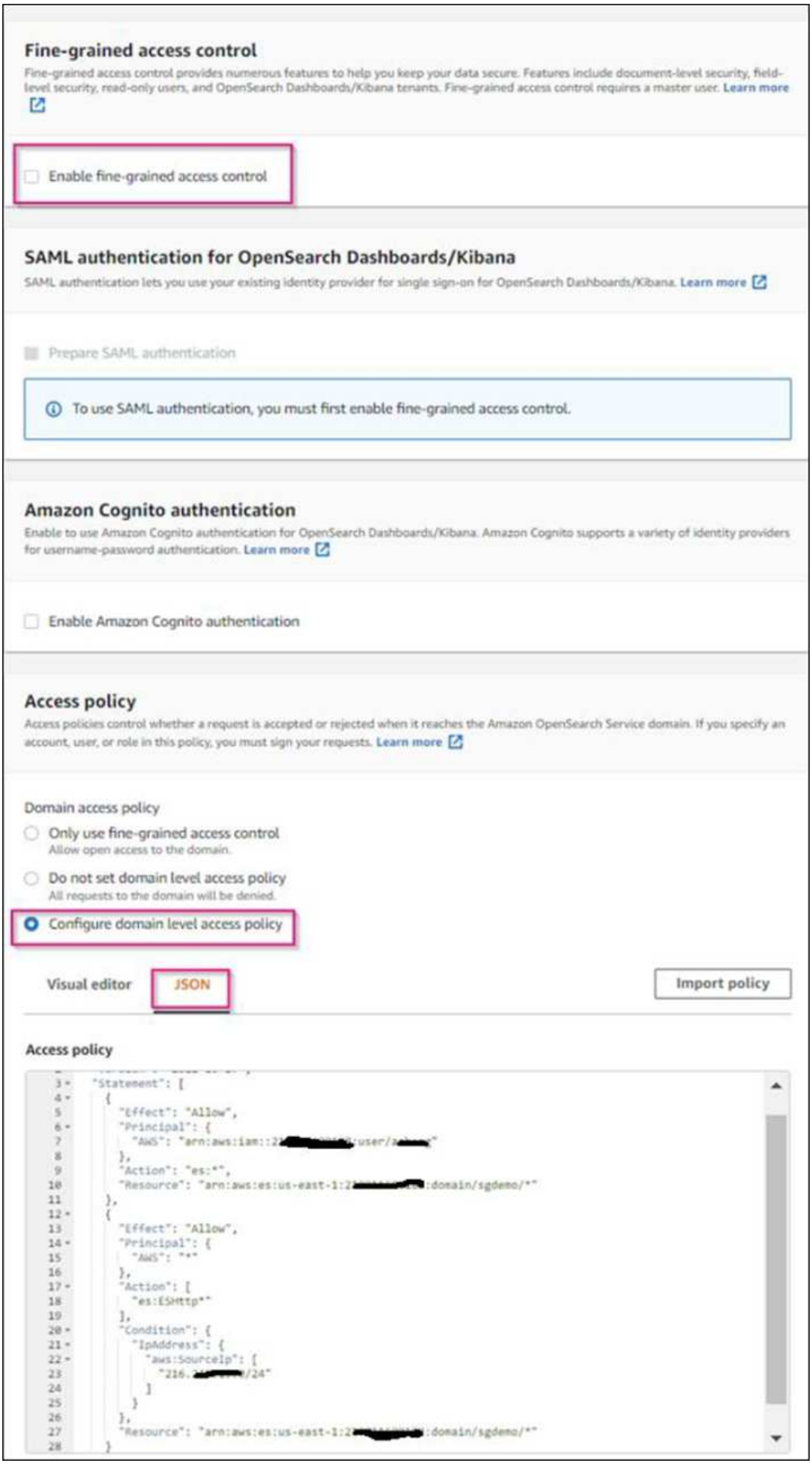
Utilisez cette procédure pour une configuration rapide et simple du service OpenSearch à des fins de test/démonstration uniquement. Si vous utilisez Elasticsearch sur site pour des services d'intégration de la recherche, consultez la section Services d'intégration de recherche avec Elasticsearch sur site.

Vous devez disposer d'un identifiant de console AWS valide, d'une clé d'accès, d'une clé d'accès secrète et d'une autorisation pour vous abonner au service OpenSearch.

1. Créez un nouveau domaine à l'aide des instructions de "Mise en route du service OpenSearch d'AWS", à l'exception de ce qui suit :

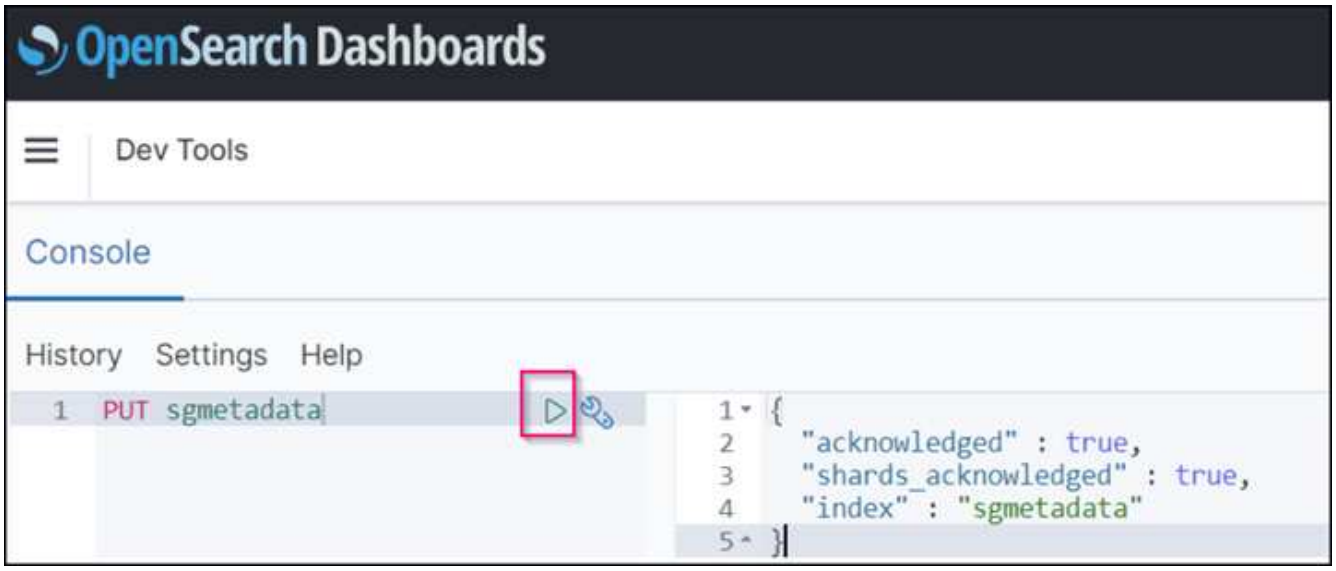
- Étape 4. Nom de domaine : sgdemo
- Étape 10. Contrôle d'accès de grain fin : désélectionnez l'option Activer le contrôle d'accès de grain fin.
- Étape 12. Règle d'accès : sélectionnez configurer la stratégie d'accès de niveau, sélectionnez l'onglet JSON pour modifier la stratégie d'accès en utilisant l'exemple suivant :
  - Remplacez le texte surligné par votre propre ID et nom d'utilisateur AWS Identity and Access Management (IAM).
  - Remplacez le texte en surbrillance (adresse IP) par l'adresse IP publique de votre ordinateur local utilisé pour accéder à la console AWS.
  - Ouvrez un onglet de navigateur pour "https://checkip.amazonaws.com" Pour trouver votre IP publique.

```
{

    "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/*"
        }
    ]
}
```

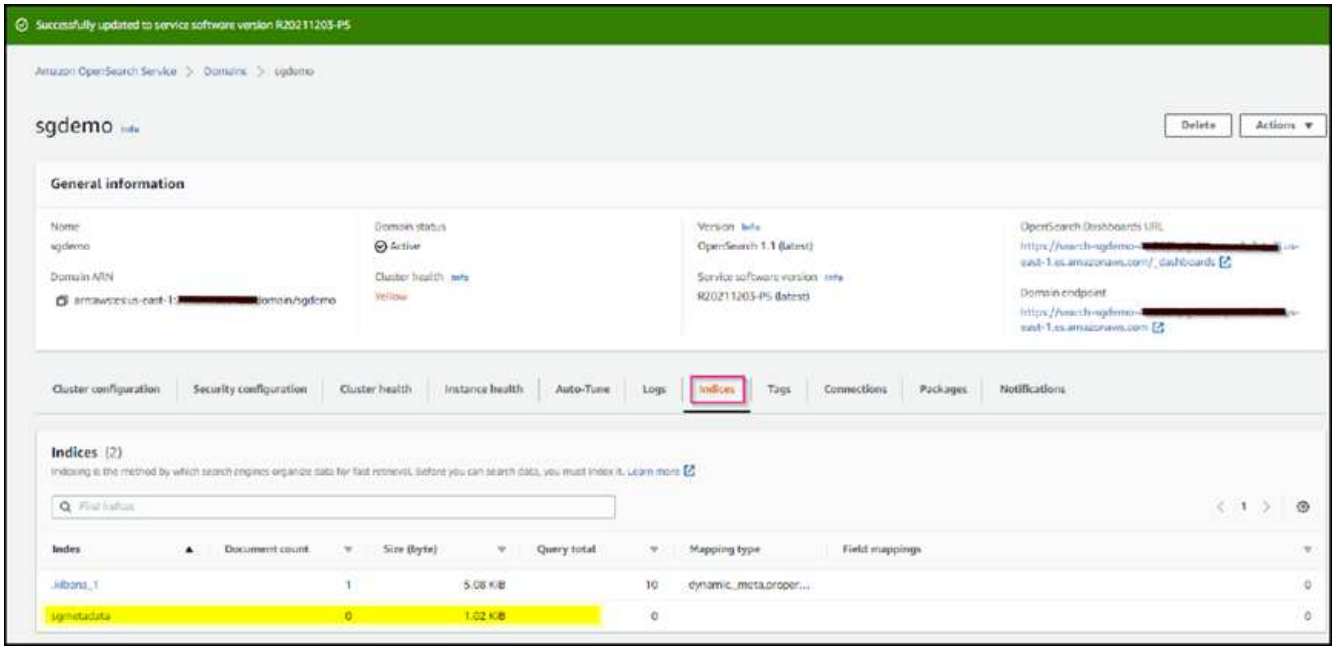

Fine-grained access control
Fine-grained access control provides numerous features to help you keep your data secure. Features include document-level security, field-level security, read-only users, and OpenSearch Dashboards/Kibana tenants. Fine-grained access control requires a master user. Learn more
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. Attendez 15 à 20 minutes pour que le domaine devienne actif.


3. Cliquez sur OpenSearch tableaux de bord URL pour ouvrir le domaine dans un nouvel onglet pour accéder au tableau de bord. Si vous obtenez une erreur d'accès refusé, vérifiez que l'adresse IP source de la stratégie d'accès est correctement définie sur l'adresse IP publique de votre ordinateur pour autoriser l'accès au tableau de bord du domaine.
4. Sur la page d'accueil du tableau de bord, sélectionnez Explorer de votre choix. Dans le menu, accédez à Management → Dev Tools
5. Sous Outils de développement → Console , entrez `PUT <index>` Où vous utilisez l'index pour le stockage des métadonnées d'objet StorageGRID. Nous utilisons le nom d'index 'gmetadatas' dans l'exemple suivant. Cliquez sur le petit symbole de triangle pour exécuter la commande PUT. Le résultat attendu s'affiche dans le panneau de droite comme indiqué dans l'exemple d'écran suivant.




6. Vérifiez que l'index est visible depuis l'interface utilisateur Amazon OpenSearch sous sgdomain > indices.

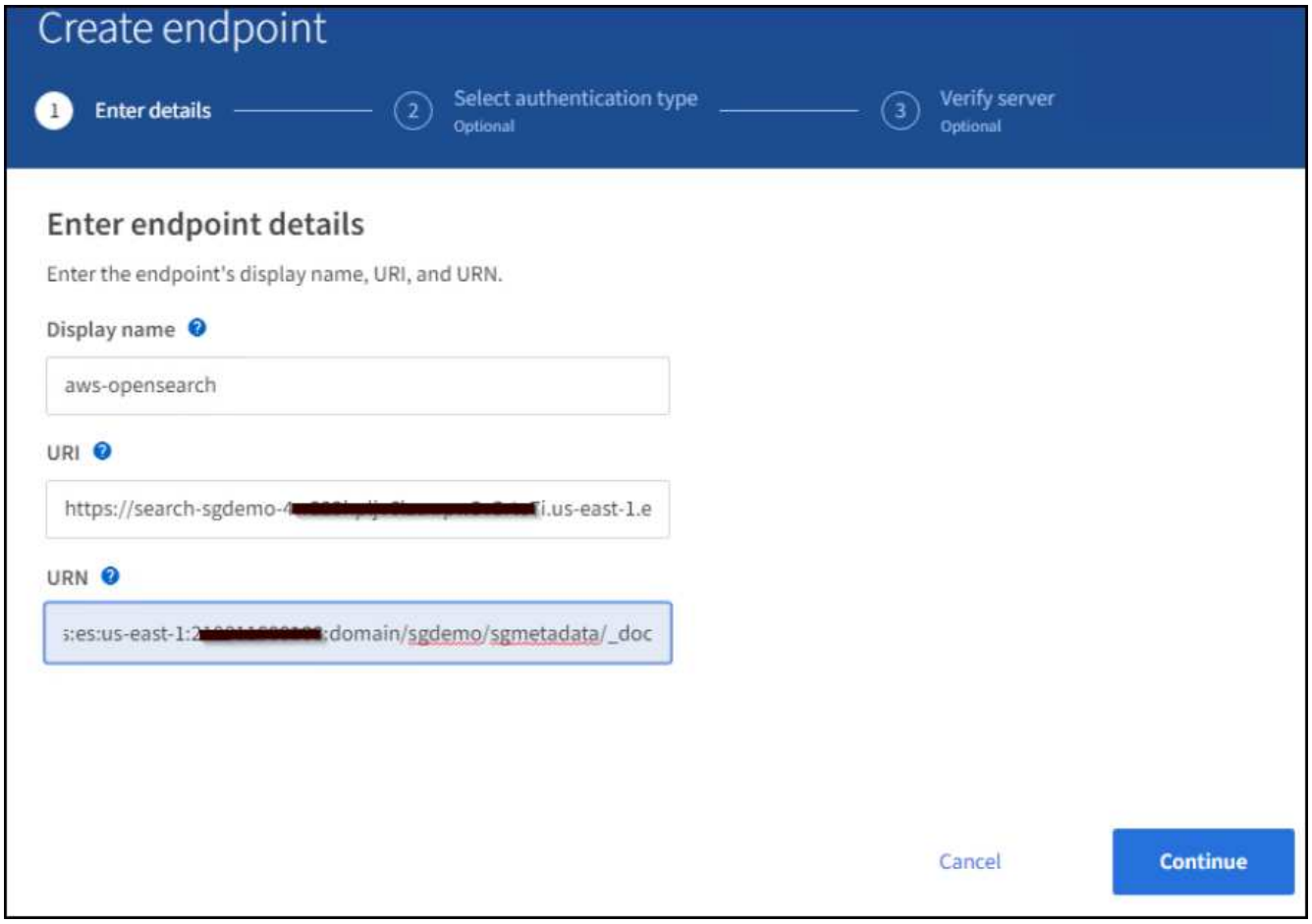



## Configuration du terminal des services de plate-forme

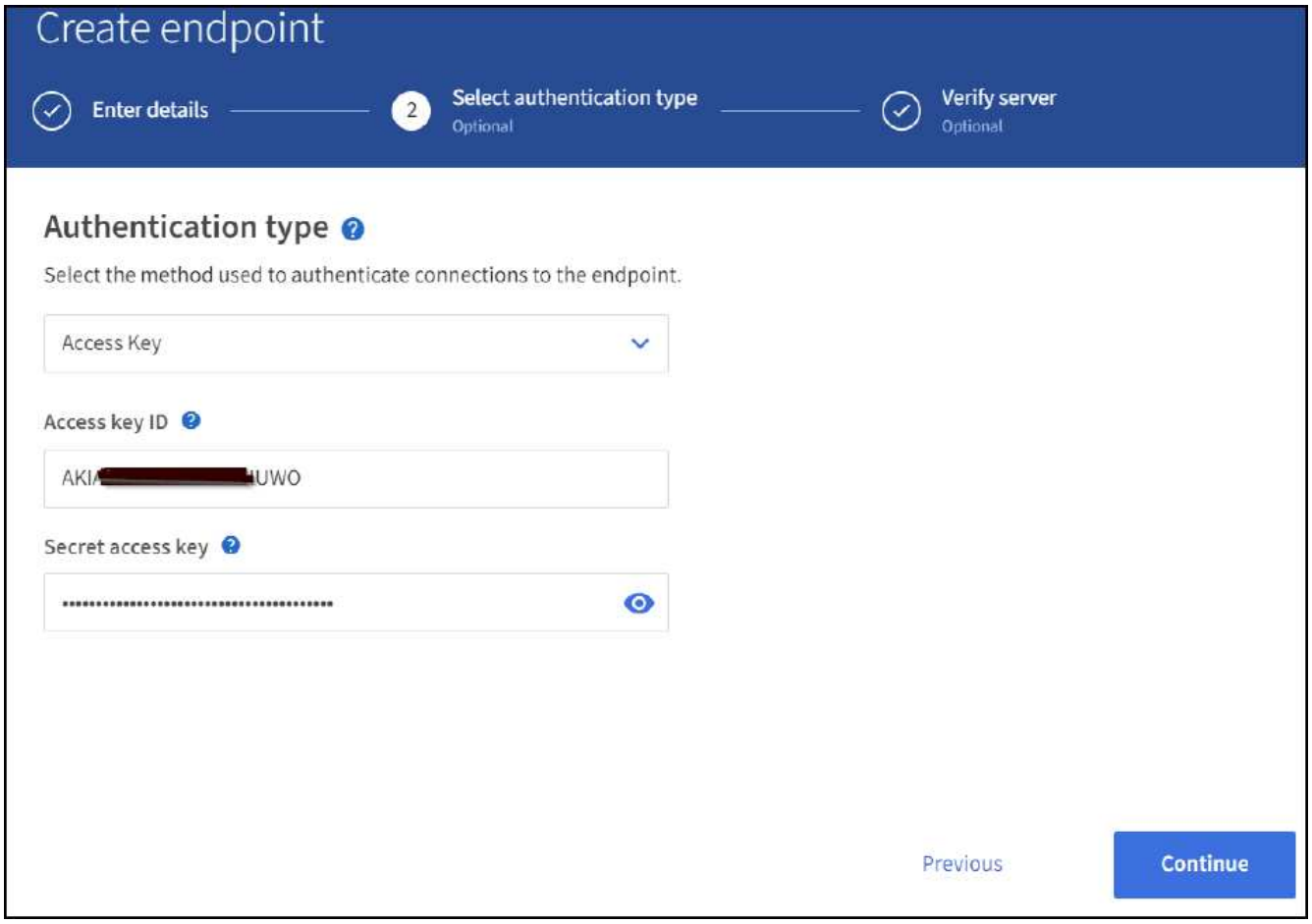
Pour configurer les terminaux des services de plate-forme, procédez comme suit :

1. Dans tenant Manager, accédez à STORAGE(S3) > terminaux des services de plateforme.
2. Cliquez sur Créer un point final, entrez les informations suivantes, puis cliquez sur Continuer :
   - Exemple de nom d'affichage `aws-opensearch`
   - Le noeud final du domaine dans la capture d'écran de l'exemple sous l'étape 2 de la procédure précédente dans le champ URI.
   - Le domaine ARN utilisé à l'étape 2 de la procédure précédente dans le champ URN et ajouter `/<index>/_doc` Jusqu'à la fin de l'ARN.

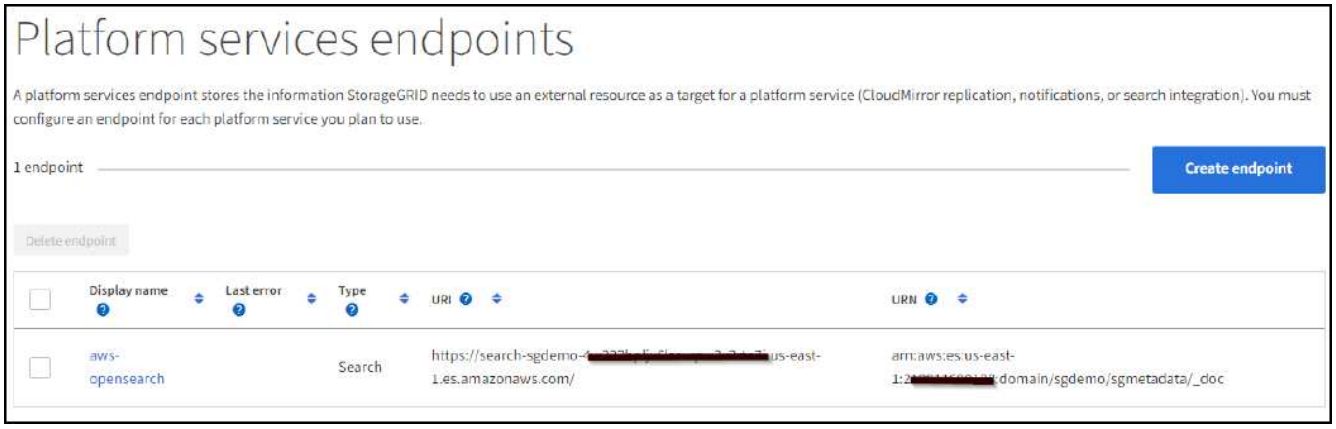
     Dans cet exemple, l'URN devient `arn:aws:es:us-east-1:211234567890:domain/sgdemo/sgmedata/_doc`.




3. Pour accéder au domaine Amazon OpenSearch sgdomain, choisissez Access Key comme type d'authentification, puis entrez la clé d'accès Amazon S3 et la clé secrète. Pour passer à la page suivante, cliquez sur Continuer.




4. Pour vérifier le noeud final, sélectionnez utiliser le certificat CA du système d'exploitation et tester et Créer un noeud final. Si la vérification réussit, un écran de point final similaire à la figure suivante s'affiche. En cas d'échec de la vérification, vérifiez que l'URN inclut `/<index>/_doc` À l'issue du chemin, la clé d'accès AWS et la clé secrète sont correctes.




## Services d'intégration de recherche avec Elasticsearch sur site

### Configuration Elasticsearch sur site

Cette procédure permet une configuration rapide des données sur site Elasticsearch et Kibana utilisant docker uniquement à des fins de test. Si le serveur Elasticsearch et Kibana existent déjà, passez à l'étape 5.

1. Suivez ceci "Procédure d'installation de Docker" pour installer docker. Nous utilisons le "Procédure d'installation de CentOS Docker" dans cette configuration.

```
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
```

   - Pour démarrer docker après le redémarrage, entrez les informations suivantes :

```
sudo systemctl enable docker
```

   - Réglez le `vm.max_map_count` valeur jusqu'à 262144 :

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

   - Pour conserver le paramètre après le redémarrage, saisissez les informations suivantes :

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

2. Suivez le "Guide de démarrage rapide d'Elasticsearch" Section auto-gérée pour installer et exécuter Elasticsearch et Kibana docker. Dans cet exemple, nous avons installé la version 8.1.

Notez le nom d'utilisateur/mot de passe et le jeton créés par Elasticsearch, vous devez utiliser ces éléments pour démarrer l'interface utilisateur Kibana et l'authentification du terminal de la plateforme StorageGRID.

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.

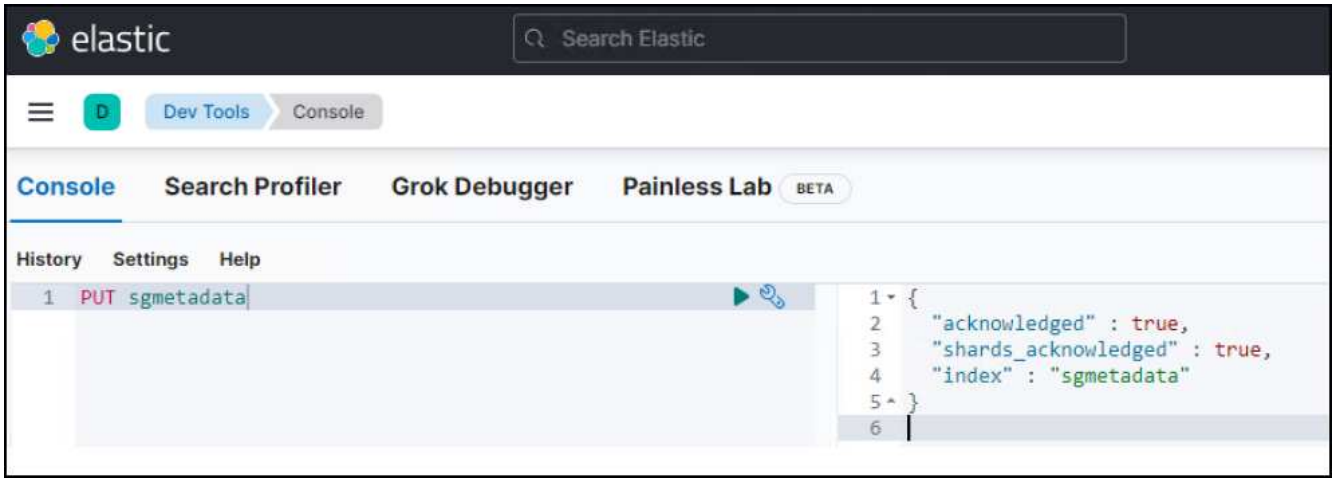
1. In a new terminal session, run:

```
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.

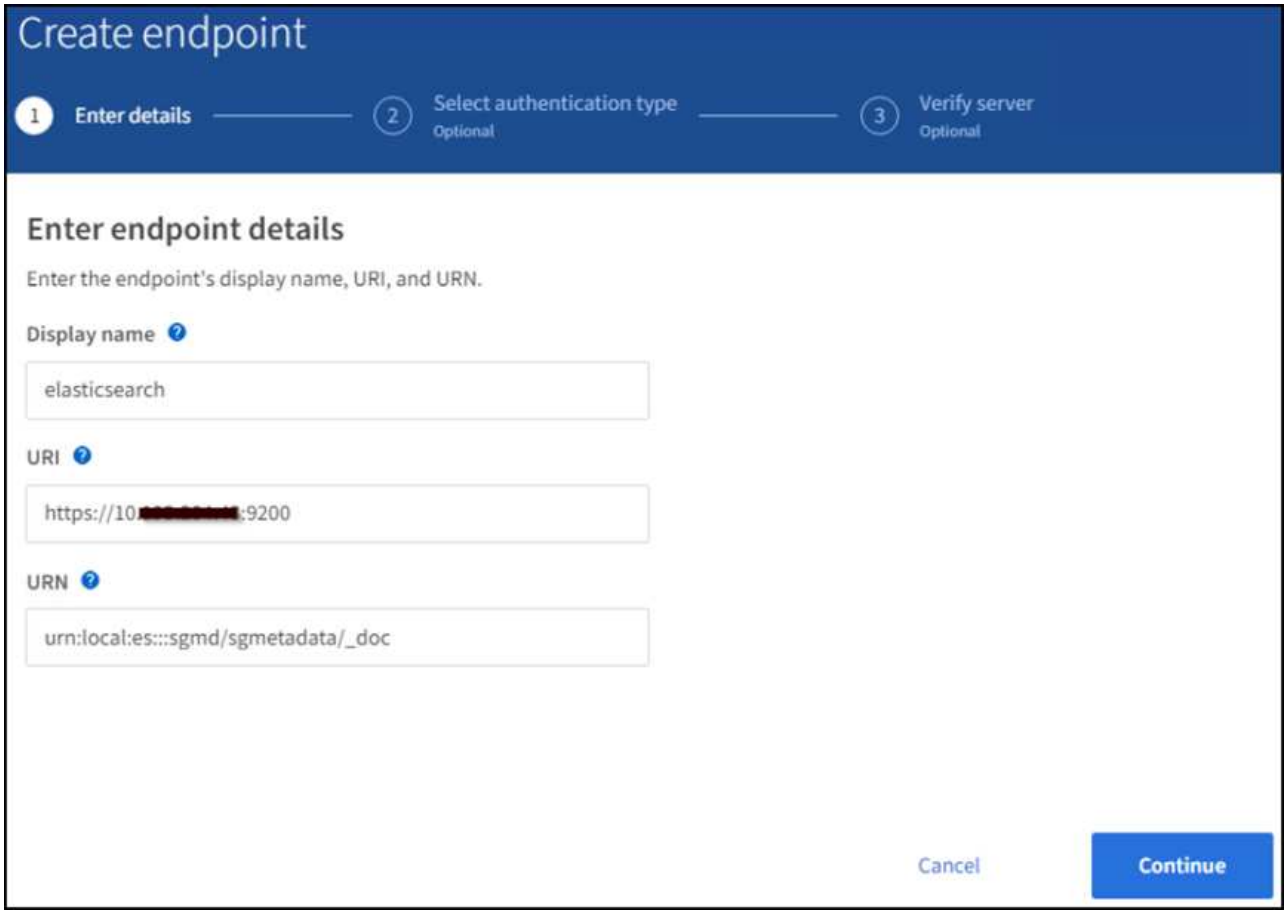
3. Après le démarrage du conteneur kibana docker, le lien URL `https://0.0.0.0:5601` s'affiche dans la console. Remplacez 0.0.0.0 par l'adresse IP du serveur dans l'URL.
4. Connectez-vous à l'interface utilisateur Kibana en utilisant le nom d'utilisateur `elastic` Et le mot de passe généré par Elastic dans l'étape précédente.
5. Pour la première connexion, sur la page d'accueil du tableau de bord, sélectionnez Explorer par vous-même. Dans le menu, sélectionnez gestion > Outils de développement.
6. Sur l'écran Console des outils de développement, entrez `PUT <index>` Où vous utilisez cet index pour stocker les métadonnées des objets StorageGRID. Nous utilisons le nom de l'index `sgmetadata` dans cet exemple. Cliquez sur le petit symbole de triangle pour exécuter la commande PUT. Le résultat attendu s'affiche dans le panneau de droite comme indiqué dans l'exemple d'écran suivant.



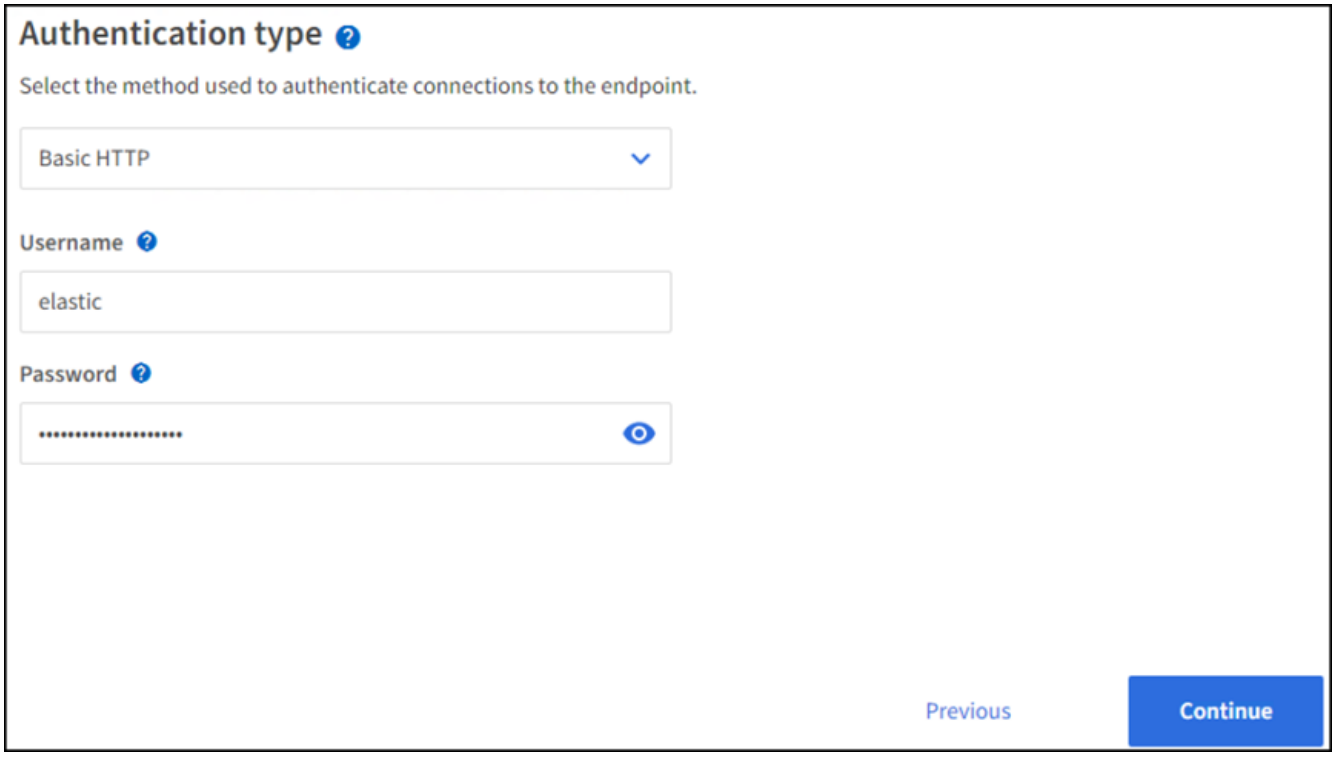

## Configuration du terminal des services de plate-forme

Pour configurer les terminaux pour les services de plate-forme, procédez comme suit :

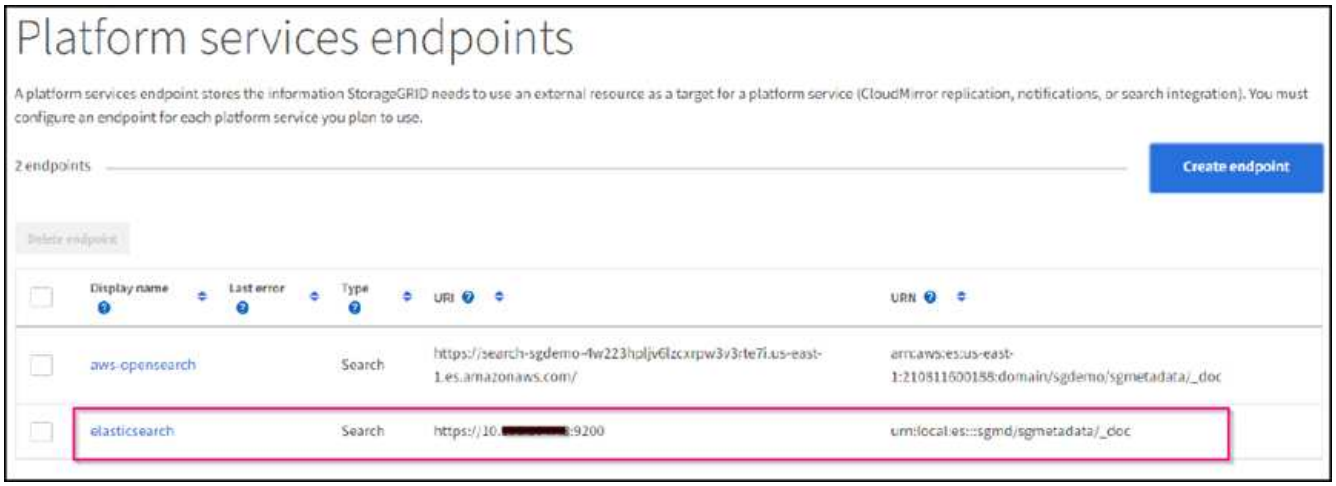
1. Dans tenant Manager, accédez à STORAGE(S3) > terminaux des services de plateforme
2. Cliquez sur Créer un point final, entrez les informations suivantes, puis cliquez sur Continuer :
   - Exemple de nom d'affichage : `elasticsearch`
   - URI : `https://<elasticsearch-server-ip or hostname>:9200`
   - URN : `urn:<something>:es:::<some-unique-text>/<index-name>/_doc` Où l'index-name est le nom que vous avez utilisé sur la console Kibana. Exemple : `urn:local:es:::sgmd/sgmetadata/_doc`




3. Sélectionnez Basic HTTP comme type d'authentification, saisissez le nom d'utilisateur `elastic` Et le mot de passe généré par le processus d'installation Elasticsearch. Pour passer à la page suivante, cliquez sur Continuer.



4. Sélectionnez ne pas vérifier le certificat et le test et Créer un noeud final pour vérifier le noeud final. Si la vérification est réussie, un écran de point final similaire à la capture d'écran suivante s'affiche. Si la vérification échoue, vérifiez que les entrées URN, URI et nom d'utilisateur/mot de passe sont correctes.




## Configuration du service d'intégration de la recherche de compartiments

Une fois le terminal du service de plateforme créé, l'étape suivante consiste à configurer ce service au niveau du compartiment pour envoyer les métadonnées d'objet au terminal défini lors de la création ou de la suppression d'un objet, ou encore lors de la mise à jour de ses métadonnées ou balises.

Vous pouvez configurer l'intégration de la recherche à l'aide du Gestionnaire de locataires afin d'appliquer un code XML de configuration StorageGRID personnalisé à un compartiment comme suit :

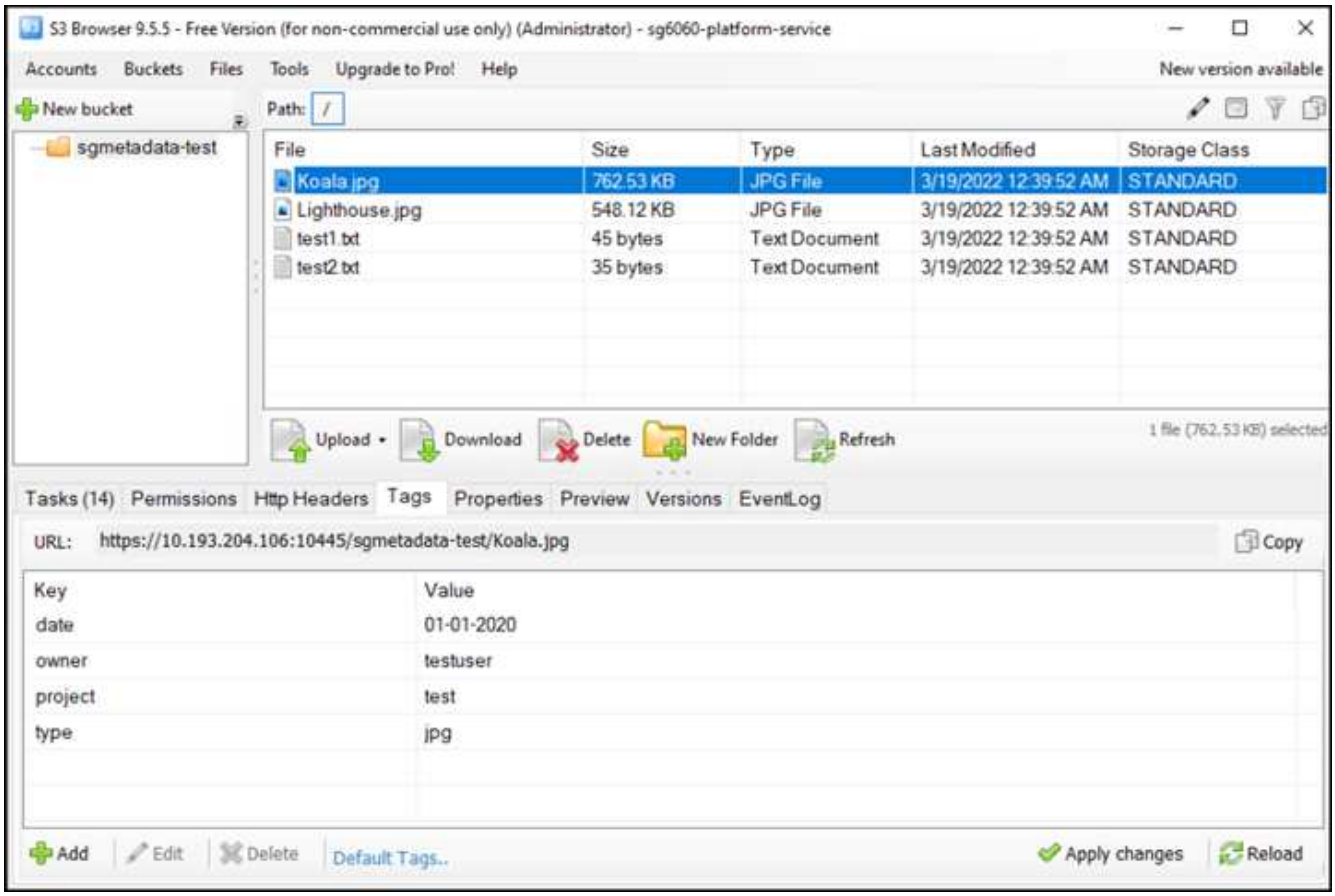
1. Dans le Gestionnaire des locataires, accédez à STORAGE(S3) > compartiments
2. Cliquez sur Créer un compartiment, entrez le nom du compartiment (par exemple, `sgmetadata-test`) et acceptez la valeur par défaut `us-east-1` région.
3. Cliquez sur Continuer > Créer un compartiment.
4. Pour afficher la page de présentation du compartiment, cliquez sur le nom du compartiment, puis sélectionnez Platform Services.
5. Sélectionnez la boîte de dialogue Activer l'intégration de la recherche. Dans la zone XML fournie, entrez le XML de configuration à l'aide de cette syntaxe.

   L'URN mis en surbrillance doit correspondre au terminal des services de plateforme que vous avez défini. Vous pouvez ouvrir un autre onglet du navigateur pour accéder au Gestionnaire de locataires et copier l'URN à partir du noeud final de services de plateforme défini.

   Dans cet exemple, nous n'avons utilisé aucun préfixe, ce qui signifie que les métadonnées de chaque objet de ce compartiment sont envoyées au terminal Elasticsearch précédemment défini.

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

6. Utilisez le navigateur S3 pour vous connecter à StorageGRID avec la clé secrète/d'accès par locataire, et téléchargez les objets de test vers `sgmetadata-test` et ajoutez des balises ou des métadonnées personnalisées aux objets.



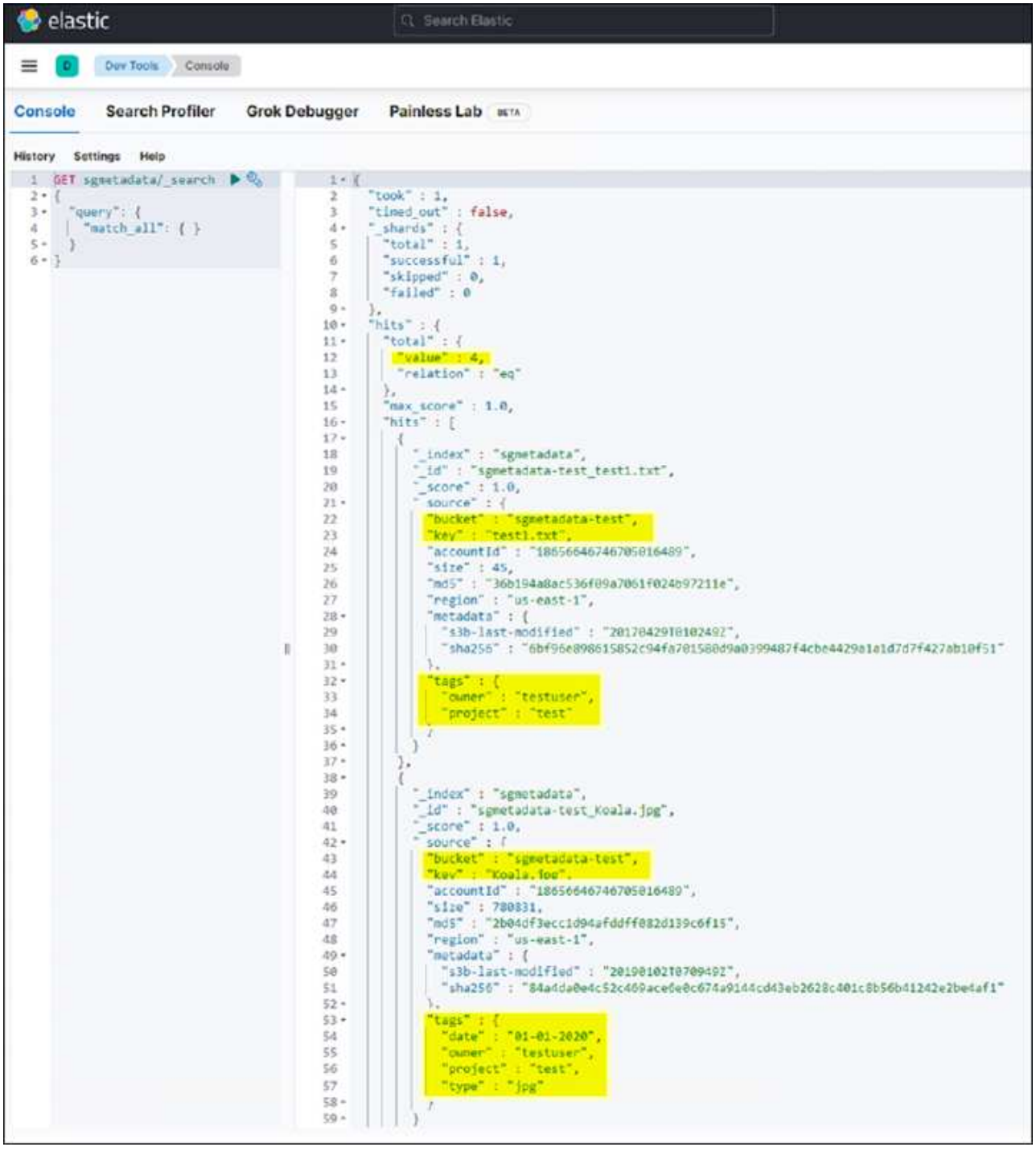

7. Utilisez l'interface utilisateur Kibana pour vérifier que les métadonnées de l'objet ont été chargées dans l'index des métadonnées sgmetadata.

   a. Dans le menu, sélectionnez gestion > Outils de développement.

   b. Collez l'exemple de requête dans le panneau de la console à gauche et cliquez sur le symbole du triangle pour l'exécuter.

      L'exemple de résultat de la requête 1 dans la capture d'écran suivante montre quatre enregistrements. Ceci correspond au nombre d'objets dans le godet.

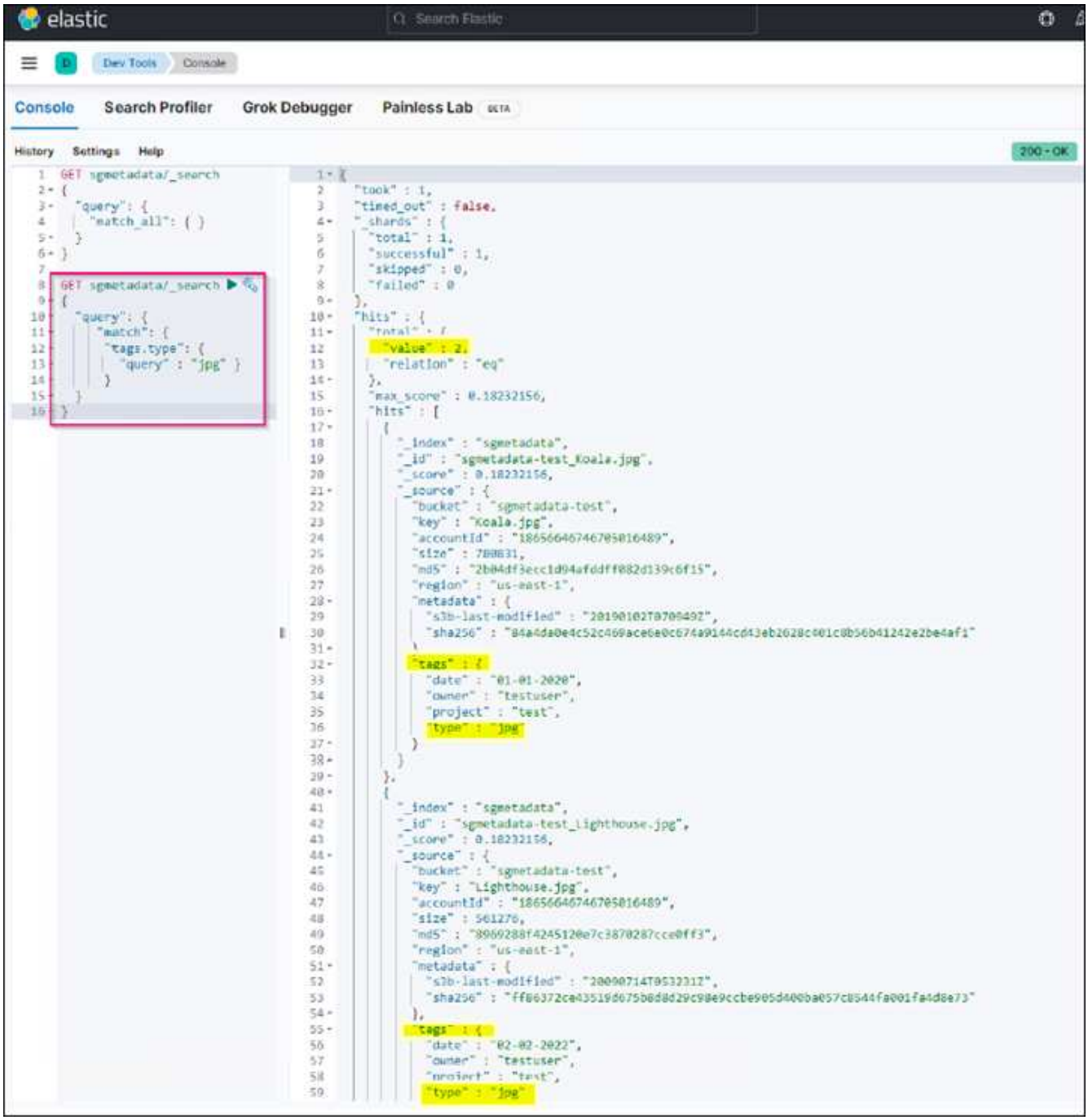
```
GET sgmetadata/_search
{
    "query": {
        "match_all": { }
}
}
```




Le résultat de l'exemple de requête 2 dans la capture d'écran suivante montre deux enregistrements de type de balise jpg.

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

+



### Où trouver des informations complémentaires

Pour en savoir plus sur les informations données dans ce livre blanc, consultez ces documents et/ou sites web :

- "Qu'est-ce que les services de plateforme"
- "Documentation StorageGRID 11.6"

*Par Angela Cheng*

# Clone de nœud

Considérations et performances sur le clonage des nœuds.

### Considérations relatives au clonage de nœuds

Le clone de nœud peut être une méthode plus rapide pour remplacer les nœuds d'appliance existants dans le cadre d'une mise à jour technologique, d'une augmentation de la capacité ou d'une augmentation de la performance du système StorageGRID. Le clone de nœud peut également être utile pour la conversion en chiffrement de nœud avec un KMS ou pour le remplacement d'un nœud de stockage DDP8 par DDP16.

- La capacité utilisée du nœud source n'est pas pertinente pour le temps nécessaire à la fin du processus de clonage. Le clone de nœud est une copie complète du nœud, y compris l'espace libre dans le nœud.
- Les appareils source et cible doivent avoir la même version PGE
- La capacité du nœud de destination doit toujours être supérieure à la source
  - Assurez-vous que la nouvelle appliance de destination possède un lecteur plus grand que la source
  - Si l'appliance de destination possède des lecteurs de même taille et est configurée pour DDP8, vous pouvez configurer la destination pour DDP16. Si la source est déjà configurée pour DDP16, le clone de nœud ne sera pas possible.
  - Lorsque vous utilisez des appliances SG5660 ou SG5760 pour des appliances SG6060, sachez que les SG6060 disposent de 60 disques de capacité lorsque le SG6060 ne présente que 58.
- Le processus de clonage de nœud nécessite que le nœud source soit hors ligne de la grille pendant toute la durée du processus de clonage. Si un nœud supplémentaire se déconnecte pendant ce temps, les services client peuvent être affectés.
- Un nœud de stockage ne peut être hors ligne que pendant 15 jours. Si l'estimation du processus de clonage est proche de 15 jours ou supérieure à 15 jours, utilisez les procédures d'extension et de désaffectation.
- Pour un SG6060 avec tiroirs d'extension, vous devez ajouter la durée nécessaire à la taille de tiroir correcte au moment de l'appliance de base pour obtenir la durée totale du clone.
- Le nombre de volumes d'une appliance de stockage cible doit être supérieur ou égal au nombre de volumes du nœud source. Vous ne pouvez pas cloner un nœud source avec 16 volumes de magasin d'objets (rangedb) vers une appliance de stockage cible avec 12 volumes de magasin d'objets, même si l'appliance cible a une capacité supérieure au nœud source. La plupart des appliances de stockage disposent de 16 volumes de stockage objet, à l'exception de l'appliance SGF6112 qui ne dispose que de 12 volumes de stockage objet. Par exemple, vous ne pouvez pas cloner à partir d'un SG5760 vers un SGF6112.

## Estimations des performances des clones de nœuds

Les tableaux suivants contiennent des estimations calculées pour la durée du clone de nœud. Les conditions varient donc, les entrées dans **BOLD** peuvent risquer de dépasser la limite de 15 jours pour un nœud en panne.

### DDP8

#### SG5612 → tous

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 1 jour | 2 jours | 2.5 jours | 3 jours | 4 jours | 4.5 jours |
| 25 GO | 1 jour | 2 jours | 2.5 jours | 3 jours | 4 jours | 4.5 jours |

#### SG5712 → tout

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 1 jour | 2 jours | 2.5 jours | 3 jours | 4 jours | 4.5 jours |
| 25 GO | 1 jour | 2 jours | 2.5 jours | 3 jours | 4 jours | 4.5 jours |

#### SG5660 → SG5760

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 3 jours | 6 jours | 7 jours | 8.5 jours | 11.5 jours | **13 jours** |
| 25 GO | 3 jours | 6 jours | 7 jours | 8.5 jours | 11.5 jours | **13 jours** |

#### SG5660 → SG6060

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 2.5 jours | 4.5 jours | 5.5 jours | 6.5 jours | 9 jours | 10 jours |
| 25 GO | 2 jours | 4 jours | 5 jours | 6 jours | 8 jours | 9 jours |

**SG5760 → SG5760**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 3 jours | 6 jours | 7 jours | 8.5 jours | 11.5 jours | **13 jours** |
| 25 GO | 3 jours | 6 jours | 7 jours | 8.5 jours | 11.5 jours | **13 jours** |

**SG5760 → SG6060**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 2.5 jours | 4.5 jours | 5.5 jours | 6.5 jours | 9 jours | 10 jours |
| 25 GO | 1.5 jours | 3 jours | 3.5 jours | 4.5 jours | 6 jours | 6.5 jours |

**SG6060 → SG6060**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 2.5 jours | 4.5 jours | 5.5 jours | 6.5 jours | 8.5 jours | 9.5 jours |
| 25 GO | 1.5 jours | 3 jours | 3.5 jours | 4 jours | 5.5 jours | 6 jours |

### DDP16

**SG5760 → SG5760**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 3.5 jours | 6.5 jours | 8 jours | 9.5 jours | 12.5 jours | **14 jours** |
| 25 GO | 3.5 jours | 6.5 jours | 8 jours | 9.5 jours | 12.5 jours | **14 jours** |

**SG5760 → SG6060**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 2.5 jours | 5 jours | 6 jours | 7.5 jours | 10 jours | 11 jours |

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 25 GO | 2 jours | 3.5 jours | 4 jours | 5 jours | 6.5 jours | 7 jours |

**SG6060 → SG6060**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 3.5 jours | 5 jours | 6 jours | 7 jours | 9.5 jours | 10.5 jours |
| 25 GO | 2 jours | 3 jours | 4 jours | 4.5 jours | 6 jours | 7 jours |

**Tiroir d'extension (à ajouter au-dessus des SG6060 pour chaque tiroir de l'appliance source)**

| Vitesse de l'interface réseau | Taille de disque de 4 To | Taille de disque de 8 To | Taille de disque de 10 To | Taille des disques de 12 To | Taille de disque de 16 To | Taille des disques de 18 To |
|---|---|---|---|---|---|---|
| 10 GBIT/S. | 3.5 jours | 5 jours | 6 jours | 7 jours | 9.5 jours | 10.5 jours |
| 25 GO | 2 jours | 3 jours | 4 jours | 4.5 jours | 6 jours | 7 jours |

*Par Aron Klein*

# Comment utiliser le remap de port

Vous devrez peut-être remmapper un port entrant ou sortant pour plusieurs raisons. Vous pouvez passer du service d'équilibrage de la charge CLB existant au point de terminaison actuel de l'équilibreur de charge des services nginx et maintenir le même port pour réduire l'impact sur les clients, utiliser le port 443 pour le client S3 sur un réseau client de nœud d'administration ou pour les restrictions de pare-feu.

## Migration des clients S3 de CLB à NGINX avec le remap du port

Dans les versions antérieures à StorageGRID 11.3, le service Load Balancer inclus sur les nœuds de passerelle est le composant Connection Load Balancer (CLB). Dans StorageGRID 11.3, NetApp présente le service NGINX en tant que solution intégrée riche en fonctionnalités pour l'équilibrage de la charge du trafic HTTP(s). Étant donné que le service CLB reste disponible dans la version actuelle de StorageGRID, vous ne pouvez pas réutiliser le port 8082 dans la nouvelle configuration de noeud final d'équilibreur de charge. Pour contourner ce problème, le port entrant 8082 est remappé sur 10443. Toutes les requêtes HTTPS arrivant sur le port 8082 de la passerelle sont alors redirigées vers le port 10443, en contournant le service CLB et en se connectant au service NGINX. Bien que les instructions suivantes soient pour VMware, LA fonctionnalité PORT_REMAP existe pour toutes les méthodes d'installation et vous pouvez utiliser un processus similaire pour les déploiements et les appliances sans système d'exploitation.

### Déploiement du nœud de passerelle de machine virtuelle VMware

Les étapes suivantes concernent un déploiement StorageGRID dans lequel le ou les nœuds de passerelle sont déployés dans VMware vSphere 7 en tant que machines virtuelles utilisant le format OVF (Open Virtualization format) de StorageGRID. Le processus implique la suppression destructive de la machine virtuelle et le redéploiement de la machine virtuelle avec le même nom et la même configuration. Avant de mettre la machine virtuelle sous tension, modifiez la propriété vApp pour remmapper le port, puis mettez la machine virtuelle sous tension et suivez le processus de restauration du nœud.

#### Prérequis

- Vous exécutez StorageGRID 11.3 ou une version ultérieure
- Vous avez téléchargé les fichiers d'installation VMware de la version StorageGRID installée et y avez accès.
- Vous disposez d'un compte vCenter avec les autorisations d'allumer/d'éteindre les machines virtuelles, de modifier les paramètres des machines virtuelles et des vApps, de supprimer les machines virtuelles de vCenter et de déployer les machines virtuelles via OVF.
- Vous avez créé un terminal d'équilibrage de charge
  - Le port est configuré sur le port de redirection souhaité
  - Le certificat SSL du noeud final est identique à celui installé pour le service CLB dans le certificat de serveur Configuration/certificats de serveur/noeuds finaux du service API de stockage objet ou le client peut accepter une modification du certificat.

> ⓘ
>
> ```
> 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.
> ```

#### Détruisez le premier nœud de passerelle

Pour détruire le premier nœud de passerelle, procédez comme suit :

1. Choisissez le nœud de passerelle à utiliser si la grille en contient plusieurs.
2. Le cas échéant, supprimez les adresses IP de nœud de toutes les entités DNS Round Robin ou de tous les pools d'équilibrage de charge.
3. Attendez que le délai de mise en service (TTL) et les sessions ouvertes expirent.
4. Mettez le nœud VM hors tension.
5. Retirez le nœud VM du disque.

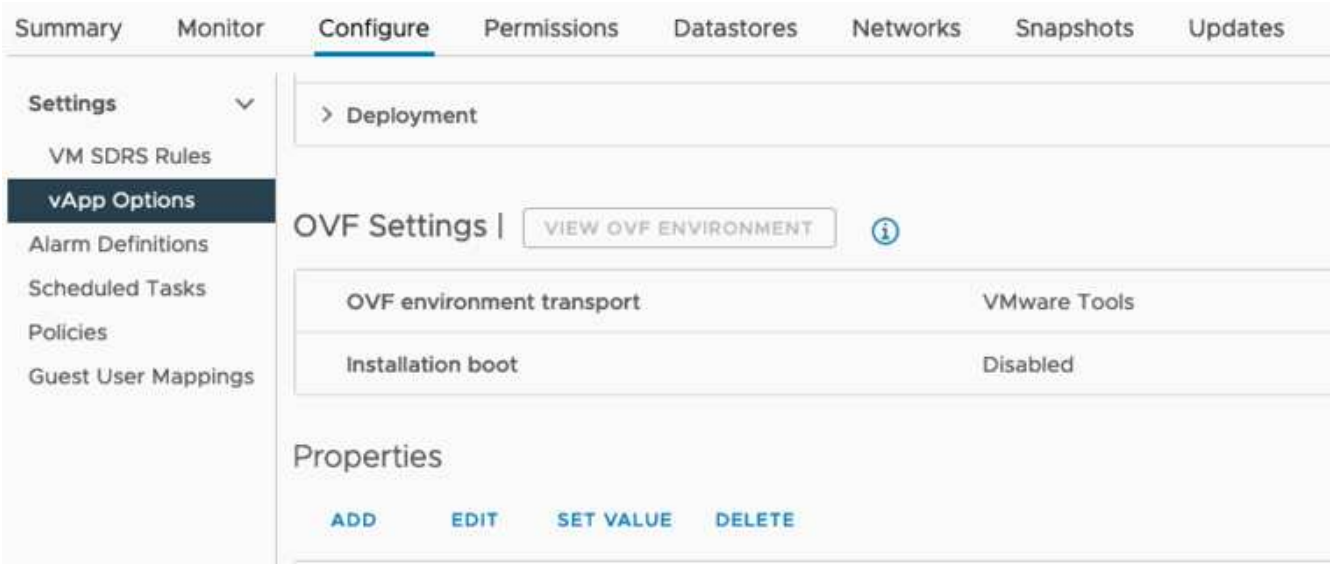
#### Déployez le nœud de passerelle de remplacement

Pour déployer le nœud de passerelle de remplacement, procédez comme suit :

1. Déployer la nouvelle machine virtuelle à partir d'OVF, en sélectionnant les fichiers .ovf, .mf et .vmdk à partir du package d'installation téléchargé à partir du site de support :
   - vsphere-gateway.mf
   - vsphere-gateway.ovf

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

2. Une fois la machine virtuelle déployée, sélectionnez-la dans la liste des machines virtuelles, puis cliquez sur l'onglet configurer Options vApp.



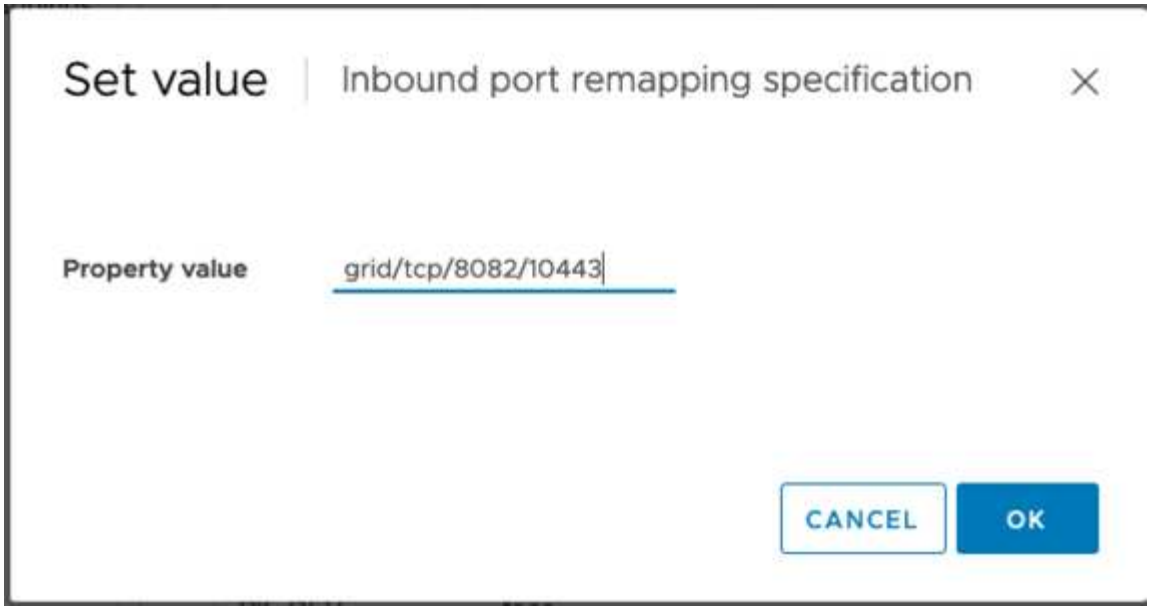

3. Faites défiler jusqu'à la section Propriétés et sélectionnez LA propriété PORT_REMAP_INBOUND


4. Faites défiler jusqu'en haut de la liste Propriétés et cliquez sur Modifier


5. Sélectionnez l'onglet Type, vérifiez que la case configurable par l'utilisateur est cochée, puis cliquez sur Enregistrer.


6. En haut de la liste Propriétés, la propriété "PORT_REMAP_INBOUND" étant toujours sélectionnée, cliquez sur définir la valeur.


7. Dans le champ valeur de la propriété, entrez le réseau (grille, admin ou client), TCP, le port d'origine (8082) et le nouveau port (10443) avec "/" entre chaque valeur, comme illustré ci-dessous.

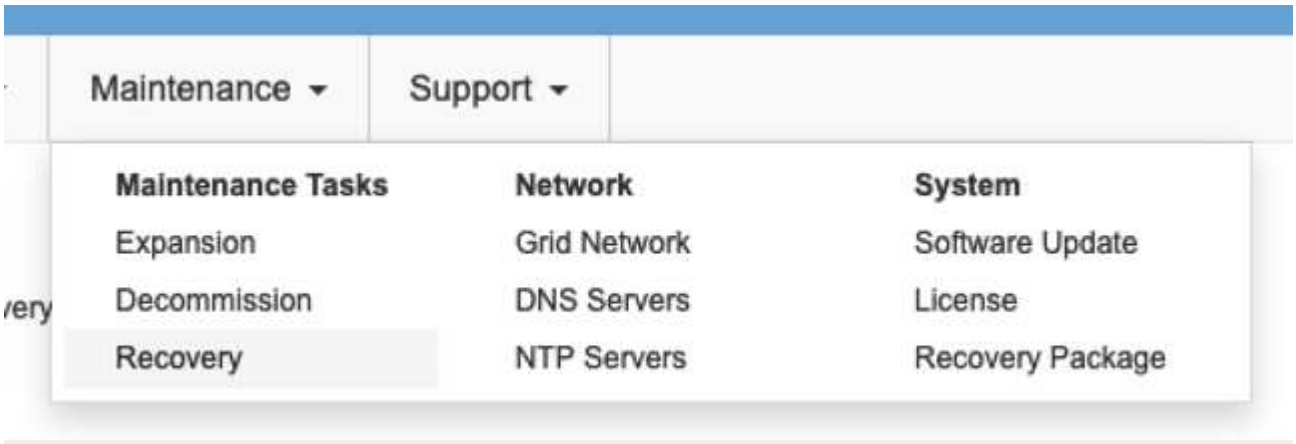



8. Si vous utilisez plusieurs réseaux, utilisez une virgule (,) pour séparer les chaînes réseau, par exemple GRID/tcp/8082/10443,admin/tcp/8082/10443,client/tcp/8082/10443

**Restaurez le nœud de passerelle**

Pour restaurer le nœud de passerelle, procédez comme suit :

1. Accédez à la section Maintenance/récupération de l'interface utilisateur de gestion du grid.




2. Mettez le nœud de la machine virtuelle sous tension et attendez que le nœud apparaisse dans la section Maintenance/Recovery Pending Nodes de l'interface utilisateur Grid Management.


> 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. Une fois le nœud restauré, l'IP peut être incluse dans toutes les entités DNS round-Robin ou dans les pools d'équilibrage de charge, le cas échéant.

   Maintenant, toutes les sessions HTTPS sur le port 8082 sont sur le port 10443

## Remadez le port 443 pour l'accès du client S3 sur un nœud d'administration

La configuration par défaut dans le système StorageGRID d'un nœud d'administration ou d'un groupe haute disponibilité contenant un nœud d'administration permet de réserver les ports 443 et 80 pour l'interface du gestionnaire de locataires et de gestion. Elle ne peut pas être utilisée pour les terminaux d'équilibrage de charge. La solution consiste à utiliser la fonction de remap de port et à rediriger le port entrant 443 vers un nouveau port qui sera configuré comme point final d'équilibrage de charge. Une fois cette opération terminée, le trafic client S3 pourra utiliser le port 443, l'interface de gestion Grid sera uniquement accessible via le port 8443 et l'interface de gestion des locataires sera uniquement accessible sur le port 9443. La fonction de remap port ne peut être configurée qu'au moment de l'installation du nœud. Pour mettre en œuvre un remap de port d'un nœud actif dans la grille, celui-ci doit être réinitialisé à l'état préinstallé. Il s'agit d'une procédure destructive qui inclut une restauration de nœud une fois la modification de configuration effectuée.

### Sauvegarde des journaux et des bases de données

Les nœuds d'administration contiennent des journaux d'audit, des metrics prometheus, ainsi que des informations historiques sur les attributs, les alarmes et les alertes. La présence de plusieurs nœuds d'administration signifie que vous avez plusieurs copies de ces données. Si vous ne disposez pas de plusieurs nœuds d'administration dans votre grid, veillez à conserver ces données à restaurer une fois le nœud restauré à la fin de ce processus. Si vous disposez d'un autre nœud d'administration dans votre grid, vous pouvez copier les données à partir de ce nœud pendant le processus de restauration. Si vous ne disposez pas d'un autre nœud d'administration dans la grille, vous pouvez suivre ces instructions pour copier les données avant de détruire le nœud.

#### Copie des journaux d'audit

1. Connectez-vous au nœud d'administration :
   a. Saisissez la commande suivante : `ssh admin@grid_node_IP`
   b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

c. Entrez la commande suivante pour passer à la racine : `su -`

d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`

f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

2. Créer le répertoire pour copier tous les fichiers journaux d'audit dans un emplacement temporaire sur un nœud de grille distinct, nous allons utiliser *Storage_node_01*:

a. `ssh admin@storage_node_01_IP`

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

3. De retour sur le nœud admin, arrêtez le service AMS pour l'empêcher de créer un nouveau fichier journal : `service ams stop`

4. Renommez le fichier audit.log de sorte qu'il ne remplace pas le fichier existant lorsque vous le copiez sur le nœud d'administration restauré.

a. Renommez audit.log en un nom de fichier numéroté unique tel que aaaa-mm-jj.txt.1. Par exemple, vous pouvez renommer le fichier journal d'audit 2015-10-25.txt.1

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

5. Redémarrez le service AMS : `service ams start`

6. Copier tous les fichiers journaux d'audit : `scp * admin@storage_node_01_IP:/var/local/tmp/saved-audit-logs`

#### Copiez les données Prometheus

La copie de la base de données Prometheus peut prendre une heure ou plus. Certaines fonctionnalités de Grid Manager ne seront pas disponibles tant que les services seront arrêtés sur le nœud d'administration.

1. Créez le répertoire pour copier les données prometheus vers un emplacement temporaire sur un nœud de grille distinct. Là encore, nous allons utiliser *Storage_node_01*:

a. Connectez-vous au nœud de stockage :

i. Saisissez la commande suivante : `ssh admin@storage_node_01_IP`

ii. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

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

2. Connectez-vous au nœud d'administration :

a. Saisissez la commande suivante : `ssh admin@admin_node_IP`

b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

c. Entrez la commande suivante pour passer à la racine : `su -`

d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`

f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

3. Depuis le nœud d'administration, arrêtez le service Prometheus : `service prometheus stop`

   a. Copiez la base de données Prometheus du nœud d'administration source vers le nœud d'emplacement de sauvegarde du nœud de stockage : `/rsync -azh --stats "/var/local/mysql_ibdata/prometheus/data" " storage_node_01_IP:/var/local/tmp/prometheus/"`

4. Redémarrez le service Prometheus sur le nœud d'administration source.`service prometheus start`

#### Sauvegarder les informations historiques

Les informations historiques sont stockées dans une base de données mysql. Pour vider une copie de la base de données, vous aurez besoin de l'utilisateur et du mot de passe de NetApp. Si vous avez un autre nœud d'administration dans la grille, cette étape n'est pas nécessaire et la base de données peut être clonée à partir d'un nœud d'administration restant pendant le processus de restauration.

1. Connectez-vous au nœud d'administration :

   a. Saisissez la commande suivante : `ssh admin@admin_node_IP`

   b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

   c. Entrez la commande suivante pour passer à la racine : `su -`

   d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

   e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`

   f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

2. Arrêtez les services StorageGRID sur le noeud d'administration et démarrez ntp et mysql

   a. Arrêter tous les services : `service servermanager stop`

   b. redémarrez le service ntp : `service ntp start`..restart mysql service: `service mysql start`

3. Vider la base de données mi dans /var/local/tmp

   a. entrez la commande suivante : `mysqldump –u username –p password mi > /var/local/tmp/mysql-mi.sql`

4. Copiez le fichier de vidage mysql sur un autre noeud, nous utiliserons *Storage_node_01: scp /var/local/tmp/mysql-mi.sql _storage_node_01_IP:/var/local/tmp/mysql-mi.sql*

a. Lorsque vous n'avez plus besoin d'un accès sans mot de passe à d'autres serveurs, supprimez la clé privée de l'agent SSH. Entrez : `ssh-add -D`

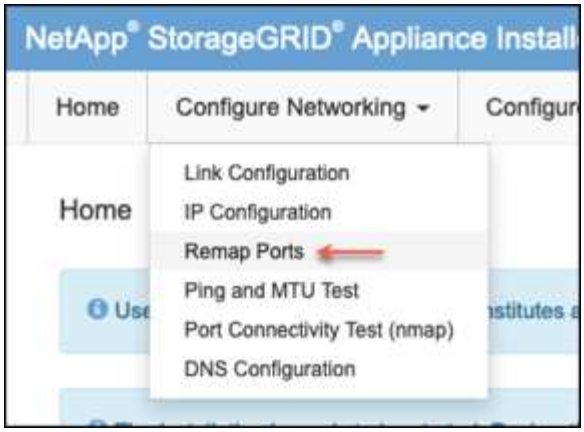
**Reconstruire le nœud d'administration**

Maintenant que vous disposez d'une copie de sauvegarde de toutes les données et journaux souhaités sur un autre nœud d'administration de la grille ou stockées dans un emplacement temporaire, il est temps de réinitialiser l'appliance afin que le remap des ports puisse être configuré.

1. La réinitialisation d'une appliance la ramène à l'état pré-installé, où elle conserve uniquement le nom d'hôte, les adresses IP et les configurations réseau. Toutes les données seront perdues, c'est pourquoi nous nous sommes assurés de disposer d'une sauvegarde de toute information importante.
   a. entrez la commande suivante : `sgareinstall`

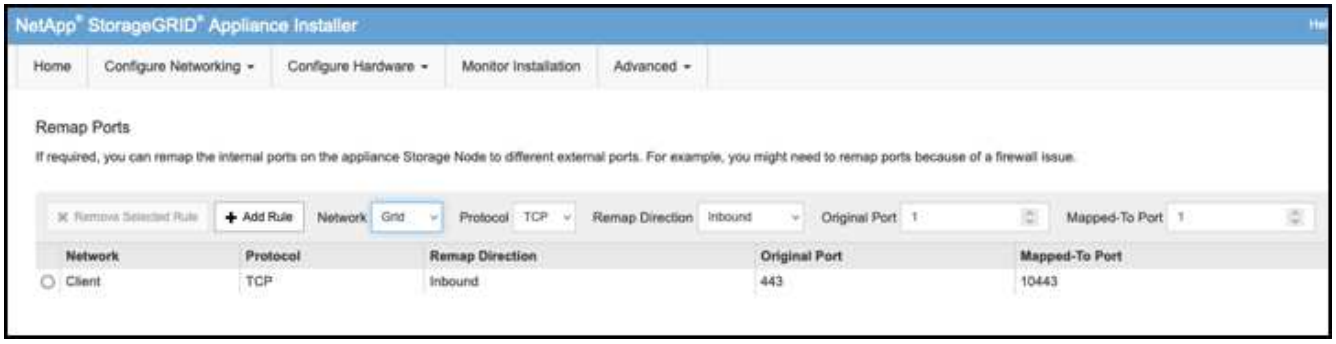
```
root@sg100-01:~ # sgareinstall
WARNING: All StorageGRID Webscale services on this node will be shut
down.
WARNING: Data stored on this node may be lost.
WARNING: You will have to reinstall StorageGRID Webscale to this
node.

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. Après un certain temps, l'appliance redémarre et vous pouvez accéder à l'interface utilisateur PGE du nœud.
3. Accédez à la page configurer la mise en réseau




4. Sélectionnez le réseau, le protocole, la direction et les ports souhaités, puis cliquez sur le bouton Ajouter une règle.

> Le remap du port entrant 443 sur le RÉSEAU DE LA GRILLE interrompt les procédures d'installation et d'extension. Il n'est pas recommandé de remapper le port 443 sur le réseau DE LA GRILLE.




5. L'un des mappages de port souhaités a été ajouté, vous pouvez revenir à l'onglet Home et cliquer sur le bouton Start installation.

Vous pouvez maintenant suivre les procédures de restauration du nœud Admin dans le "documentation produit"

## Restaurer les bases de données et les journaux

Maintenant que le nœud d'administration a été restauré, vous pouvez restaurer les metrics, les journaux et les informations d'historique. Si vous avez un autre nœud d'administration dans la grille, suivez la procédure "documentation produit" en utilisant les scripts *prometheus-clone-db.sh* et *mi-clone-db.sh*. S'il s'agit de votre seul nœud d'administration et que vous avez choisi de sauvegarder ces données, vous pouvez suivre les étapes ci-dessous pour restaurer les informations.

### Copiez à nouveau les journaux d'audit

1. Connectez-vous au nœud d'administration :
   a. Saisissez la commande suivante : `ssh admin@grid_node_IP`
   b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.
   c. Entrez la commande suivante pour passer à la racine : `su -`

d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.

e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`

f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

2. Copiez les fichiers journaux d'audit conservés sur le nœud d'administration restauré : `scp admin@grid_node_IP:/var/local/tmp/saved-audit-logs/YYYY* .`
3. Pour plus de sécurité, supprimez les journaux d'audit du nœud de grille défaillant après avoir vérifié qu'ils ont bien été copiés sur le nœud d'administration restauré.
4. Mettez à jour les paramètres utilisateur et groupe des fichiers journaux d'audit sur le nœud d'administration restauré : `chown ams-user:bycast *`

Vous devez également restaurer tout accès client existant au partage d'audit. Pour plus d'informations, reportez-vous aux instructions d'administration de StorageGRID.

**Restaurez des metrics Prometheus**

> La copie de la base de données Prometheus peut prendre une heure ou plus. Certaines fonctionnalités de Grid Manager ne seront pas disponibles tant que les services seront arrêtés sur le nœud d'administration.

1. Connectez-vous au nœud d'administration :
   a. Saisissez la commande suivante : `ssh admin@grid_node_IP`
   b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.
   c. Entrez la commande suivante pour passer à la racine : `su -`
   d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.
   e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`
   f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

2. Depuis le nœud d'administration, arrêtez le service Prometheus : `service prometheus stop`
   a. Copiez la base de données Prometheus depuis l'emplacement de sauvegarde temporaire vers le nœud d'administration : `/rsync -azh --stats "backup_node:/var/local/tmp/prometheus/" "/var/local/mysql_ibdata/prometheus/"`
   b. vérifiez que les données se trouvent dans le chemin approprié et qu'elles sont complètes `ls /var/local/mysql_ibdata/prometheus/data/`
3. Redémarrez le service Prometheus sur le nœud d'administration source.`service prometheus start`

**Restaurer les informations historiques**

1. Connectez-vous au nœud d'administration :
   a. Saisissez la commande suivante : `ssh admin@grid_node_IP`
   b. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.
   c. Entrez la commande suivante pour passer à la racine : `su -`
   d. Entrez le mot de passe indiqué dans le `Passwords.txt` fichier.
   e. Ajoutez la clé privée SSH à l'agent SSH. Entrez : `ssh-add`
   f. Entrez le mot de passe d'accès SSH répertorié dans le `Passwords.txt` fichier.

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

2. Copiez le fichier de vidage mysql à partir du nœud alternatif : `scp grid_node_IP_:/var/local/tmp/mysql-mi.sql /var/local/tmp/mysql-mi.sql`
3. Arrêtez les services StorageGRID sur le noeud d'administration et démarrez ntp et mysql
   a. Arrêter tous les services : `service servermanager stop`
   b. redémarrez le service ntp : `service ntp start`..restart mysql service: `service mysql start`
4. Supprimez la base de données mi et créez une nouvelle base de données vide : `mysql -u username -p password -A mi -e "drop database mi; create database mi;"`
5. restaurez la base de données mysql à partir du vidage de la base de données : `mysql -u username -p password -A mi < /var/local/tmp/mysql-mi.sql`
6. Redémarrez tous les autres services `service servermanager start`

*Par Aron Klein*

# Procédure de relocalisation du site dans le grid et de modification du réseau à l'échelle du site

Ce guide décrit la préparation et la procédure à suivre pour déplacer un site StorageGRID dans une grille multi-sites. Vous devez avoir une compréhension complète de cette procédure et prévoir à l'avance pour assurer un processus sans heurt et minimiser l'interruption pour les clients.

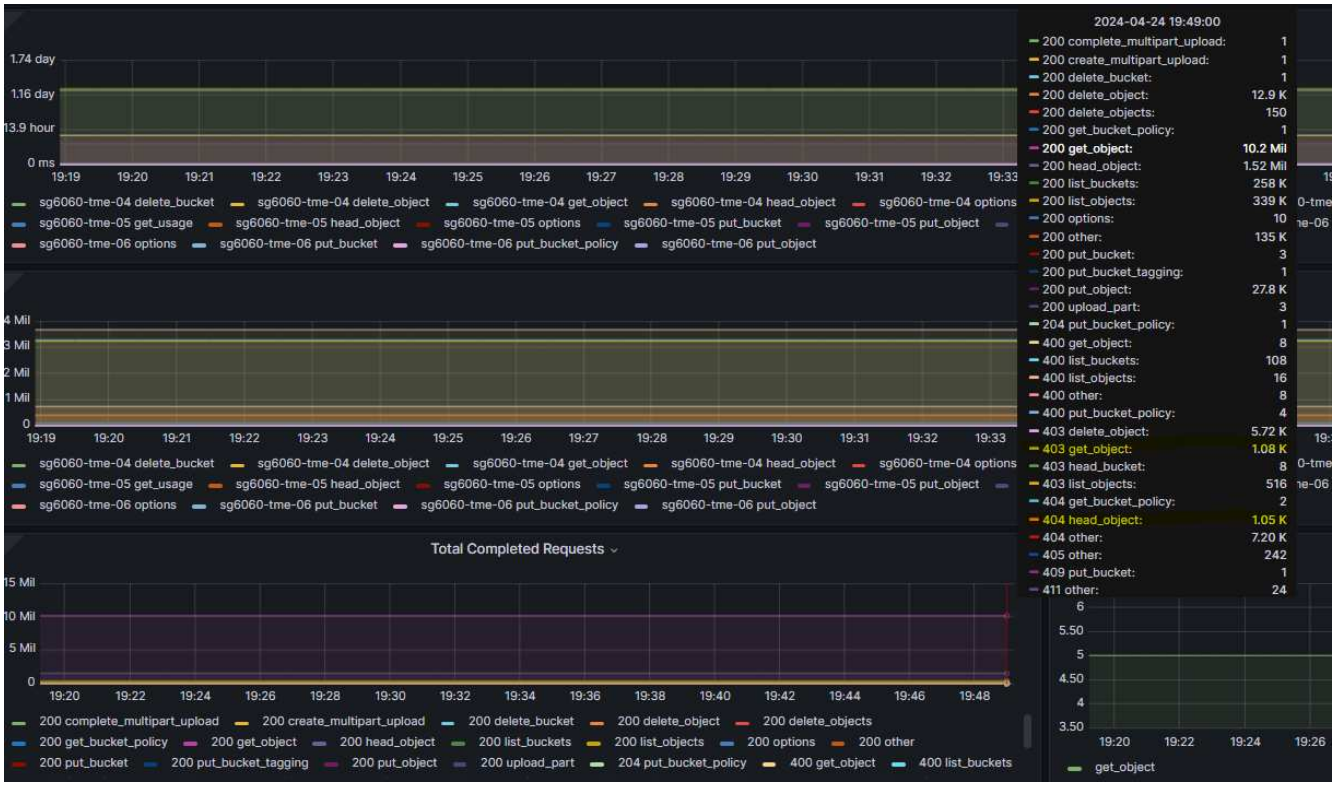
Si vous devez modifier le réseau grille de la grille entière, reportez-vous à la section "Modifiez les adresses IP de tous les nœuds de la grille".

## Considérations avant la relocalisation du site

- Le déplacement du site doit être terminé et tous les nœuds doivent être en ligne dans les 15 jours pour éviter la reconstruction de la base de données Cassandra.
  "Panne d'un nœud de stockage de plus de 15 jours"
- Si une règle ILM de la règle active utilise un comportement d'ingestion strict, envisagez de la modifier en vue de l'équilibrer ou de la double allocation si le client souhaite continuer à PLACER les objets dans la

grille pendant la relocalisation du site.

- Pour les appliances de stockage de 60 disques ou plus, ne déplacez jamais le tiroir avec des disques installés. Étiquetez chaque lecteur de disque et retirez-le du boîtier de stockage avant de le emballer/déplacer.
- Changement d'appliance StorageGRID le réseau local virtuel du réseau de la grille peut être effectué à distance sur le réseau d'administration ou le réseau client. Sinon, prévoyez d'être sur site pour effectuer la modification avant ou après la mutation.
- Vérifiez si l'application client utilise la TÊTE ou si l'objet de non-existence est utilisé avant la MISE. Si oui, remplacez la cohérence du compartiment par site fort pour éviter les erreurs HTTP 500. Si vous n'êtes pas sûr, consultez la présentation S3 graphiques Grafana **Gestionnaire de grille > support > métriques**, placez le curseur de la souris sur le graphique « demande totale terminée ». S'il y a un nombre très élevé de 404 objets GET ou 404 objets Head, une ou plusieurs applications utilisent probablement l'objet Head ou Get nonexistence. Le compte est cumulatif, passez la souris sur différents chronologies pour voir la différence.




### Procédure de modification de l'adresse IP de la grille avant le déplacement du site

### Étapes

1. Si un nouveau sous-réseau de réseau Grid sera utilisé au nouvel emplacement, "Ajoutez le sous-réseau à la liste de sous-réseau du réseau Grid"
2. Connectez-vous au nœud d'administration principal, utilisez change-ip pour effectuer une modification de l'adresse IP de la grille. **Stage** doit être effectué avant d'arrêter le nœud pour le déplacement.
   a. Sélectionnez 2 puis 1 pour modification de l'adresse IP de la grille

```
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. sélectionnez 5 pour afficher les modifications

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

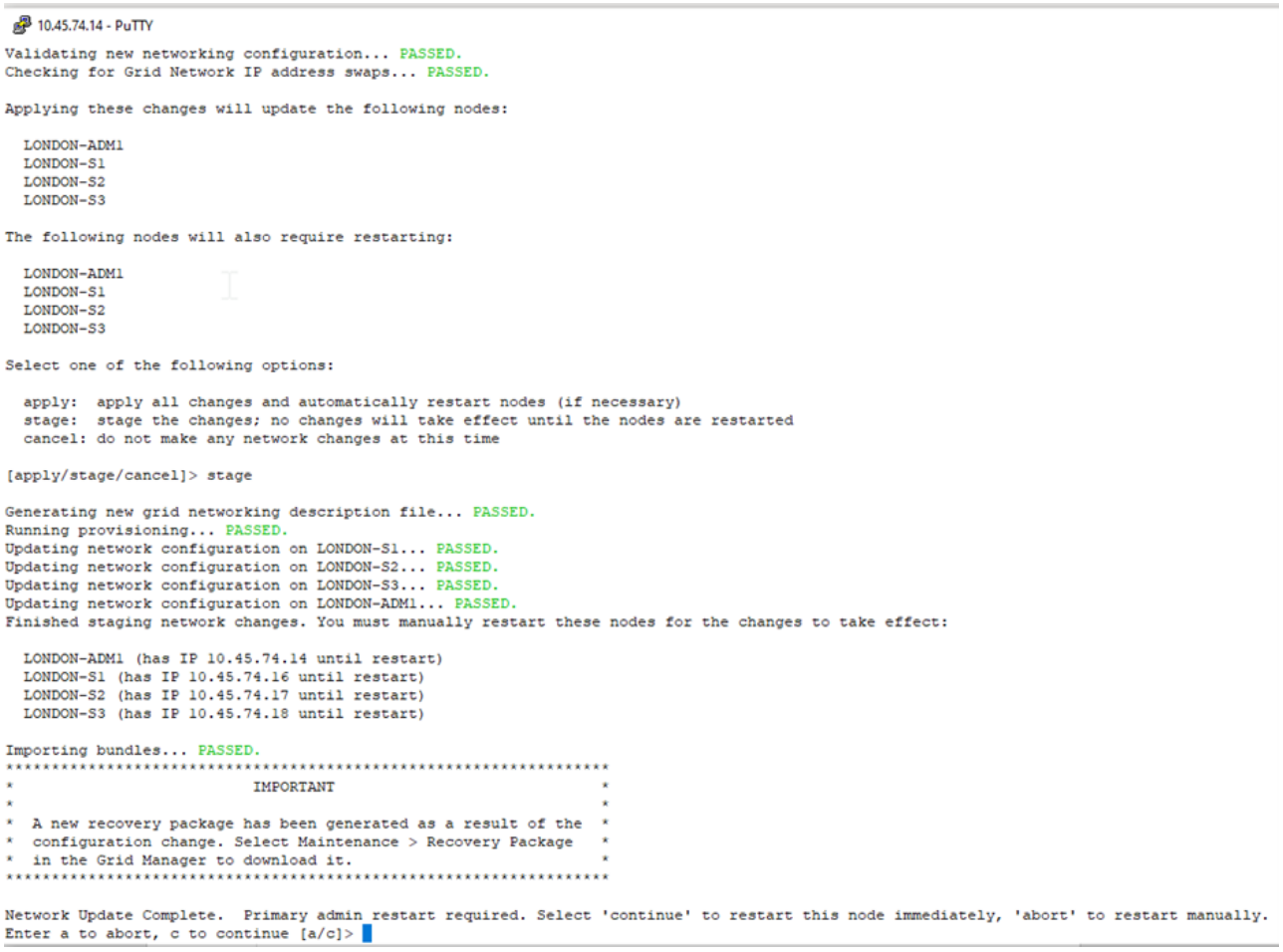
c. sélectionner 10 pour valider et appliquer la modification.

```
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
```

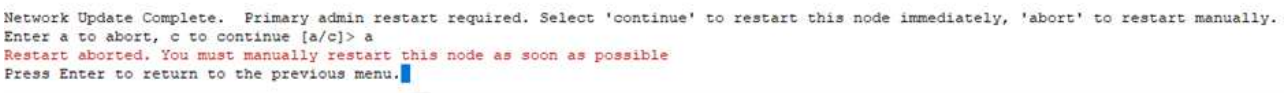
d. Vous devez sélectionner **stage** dans cette étape.

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

Applying these changes will update the following nodes:

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

The following nodes will also require restarting:

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

Select one of the following options:

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

[apply/stage/cancel]> stage
```

e. Si le nœud d'administration principal est inclus dans la modification ci-dessus, entrez **'a' pour redémarrer manuellement le nœud d'administration principal**




f. Appuyez sur ENTER pour revenir au menu précédent et quitter l'interface change-ip.




3. À partir de Grid Manager, téléchargez le nouveau package de récupération. **Grid Manager** > **Maintenance** > **paquet de récupération**
4. Si une modification VLAN est nécessaire sur l'appliance StorageGRID, reportez-vous à la section Modification du VLAN de l'appliance.
5. Arrêtez tous les nœuds et/ou appliances sur le site, étiquetez/retirez les disques si nécessaire, puis démettez, emballez et déplacez-les.
6. Si vous prévoyez de modifier l'adresse ip du réseau d'administration et/ou le VLAN et l'adresse ip du client, vous pouvez effectuer la modification après le déplacement.

### Modification du VLAN de l'appliance

La procédure ci-dessous suppose que vous disposez d'un accès à distance au réseau client ou administrateur de l'appliance StorageGRID pour effectuer la modification à distance.

#### Étapes

1. Avant d'arrêter l'appareil,
"mettez l'appareil en mode de maintenance".

2. Utilisation d'un navigateur pour accéder à l'interface graphique du programme d'installation de l'appliance StorageGRID à l'aide de https://<admin-or-client-network-ip>:8443. Impossible d'utiliser Grid IP car la nouvelle Grid IP est déjà en place une fois que l'appliance est en mode maintenance.
3. Modifiez le VLAN pour le réseau Grid. Si vous accédez à l'appliance sur le réseau client, vous ne pouvez pas modifier le VLAN client pour le moment, vous pouvez le modifier après le déplacement.
4. connectez l'appliance à l'appliance et arrêtez le nœud en utilisant « shutdoown -h now »
5. Une fois les appliances prêtes sur le nouveau site, accédez à l'interface utilisateur graphique du programme d'installation de l'appliance StorageGRID à l'aide de https://<grid-network-ip>:8443. Vérifiez que l'état du stockage est optimal et que la connectivité réseau est assurée par les autres nœuds Grid à l'aide des outils ping/nmap disponibles dans l'interface graphique.
6. Si vous prévoyez de modifier l'adresse IP du réseau client, vous pouvez modifier le VLAN client à ce stade. Le réseau client n'est pas prêt tant que vous n'avez pas mis à jour l'adresse ip du réseau client à l'aide de l'outil change-ip à l'étape suivante.
7. Quittez le mode maintenance. Dans le programme d'installation de l'appliance StorageGRID, sélectionnez **Avancé** > **redémarrer le contrôleur**, puis sélectionnez **redémarrer dans StorageGRID**.
8. Une fois que tous les nœuds sont actifs et que Grid n'indique aucun problème de connectivité, utilisez change-ip pour mettre à jour le réseau d'administration de l'appliance et le réseau client, si nécessaire.

# Guides d’utilisation et d’outils

## Utilisez le connecteur Cloudera Hadoop S3A avec StorageGRID

Hadoop est devenu l’un des préférés des data Scientists depuis un certain temps. Hadoop permet le traitement distribué d’importants jeux de données sur des clusters d’ordinateurs à l’aide d’infrastructures de programmation simples. Hadoop a été conçu pour évoluer verticalement de serveurs uniques à des milliers de machines, chaque machine étant en possession de ressources de calcul et de stockage locales.

### Pourquoi utiliser S3A pour les flux de travail Hadoop ?

Comme le volume de données a augmenté au fil du temps, l’approche qui consiste à ajouter de nouveaux ordinateurs avec leurs propres ressources de calcul et de stockage est devenue inefficace. L’évolutivité linéaire engendre des défis pour utiliser les ressources efficacement et gérer l’infrastructure.

Pour relever ces challenges, le client Hadoop S3A propose des E/S haute performance par rapport au stockage objet S3. L’implémentation d’un workflow Hadoop avec S3A vous permet d’exploiter le stockage objet en tant que référentiel de données et de séparer les ressources de calcul et de stockage. Vous pouvez ainsi faire évoluer indépendamment les ressources de calcul et de stockage. Qui dissocie le calcul et le stockage pour vous permettre de consacrer la quantité de ressources adaptée à vos tâches de calcul, et d’assurer la capacité en fonction de la taille de votre jeu de données. Par conséquent, vous pouvez réduire votre TCO global pour les workflows Hadoop.

### Configurer le connecteur S3A pour utiliser StorageGRID

#### Prérequis

- Une URL de terminal StorageGRID S3, une clé d’accès s3 pour un locataire et une clé secrète pour le test de connexion à Hadoop S3A.
- Un cluster Cloudera ainsi que l’autorisation root ou sudo pour chaque hôte du cluster afin d’installer le package Java.

En avril 2022, Java 11.0.14 avec Cloudera 7.1.7 a été testé contre StorageGRID 11.5 et 11.6. Cependant, le numéro de version de Java peut être différent au moment d’une nouvelle installation.

#### Installez le package Java

1. Vérifier le "Matrice de support Cloudera" Pour la version JDK prise en charge.
2. Téléchargez le "Package Java 11.x" Correspondant au système d’exploitation du cluster Cloudera. Copiez ce package sur chaque hôte du cluster. Dans cet exemple, le progiciel rpm est utilisé pour CentOS.
3. Connectez-vous à chaque hôte en tant que root ou en utilisant un compte avec l’autorisation sudo. Effectuez les étapes suivantes sur chaque hôte :
   a. Installez le package :

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

b. Vérifiez l'emplacement d'installation de Java. Si plusieurs versions sont installées, définissez la nouvelle version installée par défaut :

```
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. Ajoutez cette ligne à la fin de `/etc/profile`. Le chemin doit correspondre au chemin de la sélection ci-dessus :

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

d. Exécutez la commande suivante pour que le profil prenne effet :

```
source /etc/profile
```

**Configuration HDFS S3A de Cloudera**

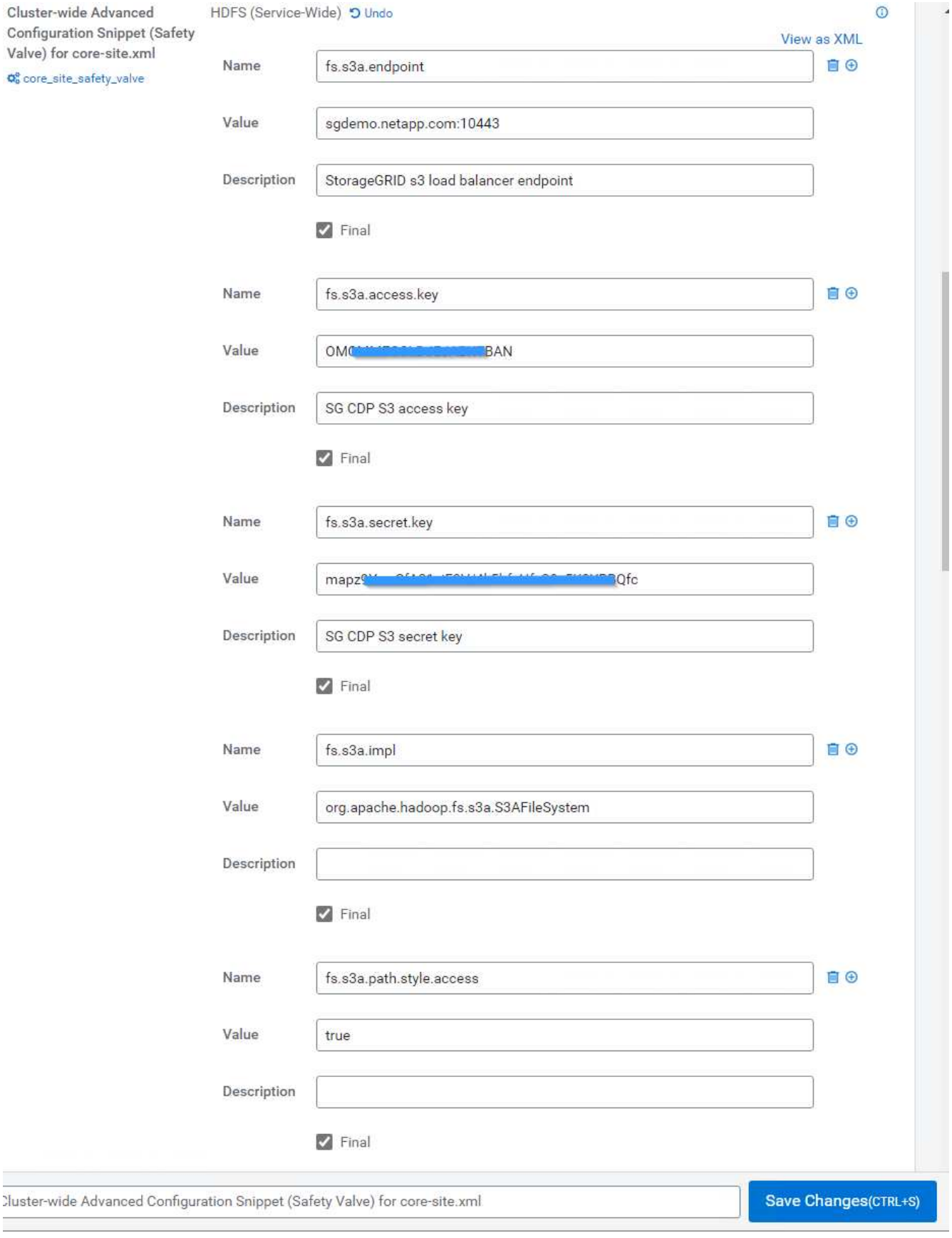
**Étapes**

1. Dans l'interface graphique Cloudera Manager, sélectionnez clusters > HDFS et sélectionnez Configuration.
2. Sous CATÉGORIE, sélectionnez Avancé, puis faites défiler vers le bas pour rechercher `Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml`.
3. Cliquez sur le signe (+) et ajoutez les paires de valeurs suivantes.

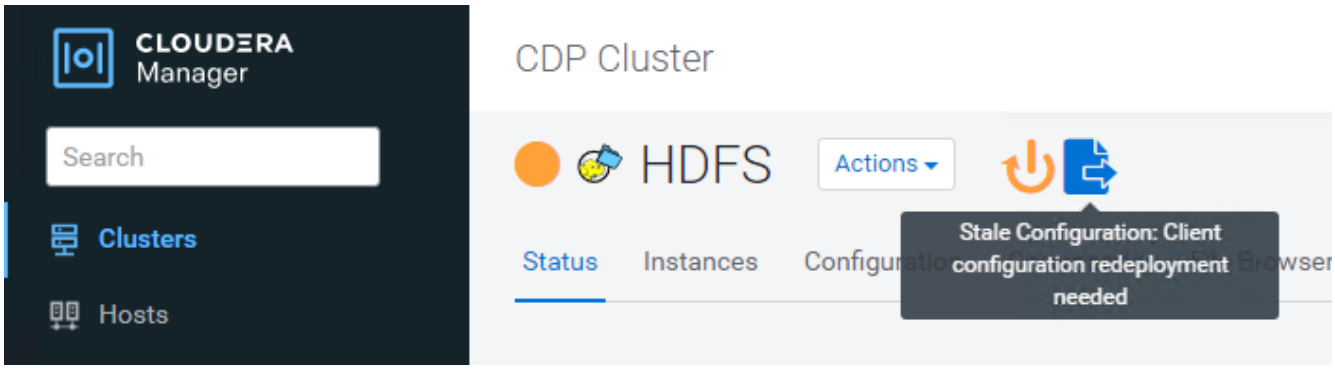
| Nom | Valeur |
|---|---|
| fs.s3a.access.key | *<clé d'accès s3 de StorageGRID>* |
| fs.s3a.secret.key | *<clé secrète S3 du locataire StorageGRID>* |
| fs.s3a.connection.ssl.enabled | [vrai ou faux] (la valeur par défaut est https si cette entrée est manquante) |
| fs.s3a.endpoint | *<noeud final StorageGRID S3:port>* |

| Nom | Valeur |
| --- | --- |
| fs.s3a.impl | ORG.apache.hadoop.fs.s3a.S3AFileSystem |
| fs.s3a.path.style.access | [vrai ou faux] (le style d'hôte virtuel par défaut est défini si cette entrée est manquante) |

**Exemple de capture d'écran**




4. Cliquez sur le bouton Enregistrer les modifications. Sélectionnez l'icône Configuration obsolète dans la barre de menus HDFS, sélectionnez redémarrer les services obsolètes sur la page suivante, puis sélectionnez redémarrer maintenant.




## Tester la connexion S3A à StorageGRID

### Effectuer un test de connexion de base

Connectez-vous à l'un des hôtes du cluster Cloudera, puis entrez `hadoop fs -ls s3a://<bucket-name>/`.

L'exemple suivant utilise le chemin syle avec un compartiment hdfs-test pré-existant et un objet test.

```
[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.
```

### Dépannage

#### Scénario 1

Utilisez une connexion HTTPS à StorageGRID et obtenez un `handshake_failure` erreur après un délai de 15 minutes.

**Raison :** ancienne version JRE/JDK utilisant la suite de chiffrement TLS obsolète ou non prise en charge pour la connexion à StorageGRID.

**Exemple de message d'erreur**

```
[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
```

**Résolution :** Assurez-vous que JDK 11.x ou version ultérieure est installé et défini par défaut la bibliothèque Java. Reportez-vous à la Installez le package Java pour plus d'informations.

**Scénario 2 :**

Impossible de se connecter à StorageGRID avec message d'erreur `Unable to find valid certification path to requested target.`

**Raison:** le certificat du serveur de noeuds finaux StorageGRID S3 n'est pas approuvé par le programme Java.

Exemple de message d'erreur :

```
[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
```

**Resolution:** NetApp recommande d'utiliser un certificat de serveur délivré par une autorité de signature de certificat public connu pour s'assurer que l'authentification est sécurisée. Vous pouvez également ajouter un certificat d'autorité de certification ou de serveur personnalisé au magasin de confiance Java.

Procédez comme suit pour ajouter une autorité de certification ou un certificat de serveur personnalisé StorageGRID au magasin d'approbation Java.

1. Sauvegardez le fichier Java cacerts existant par défaut.

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

2. Importez le certificat de noeud final StorageGRID S3 dans le magasin de confiance Java.

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

**Conseils de dépannage**

1. Augmentez le niveau de journalisation hadoop pour DÉBOGUER.

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

2. Exécutez la commande et dirigez les messages du journal vers error.log.

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

*Par Angela Cheng*

# Utilisez S3cmd pour tester et démontrer l'accès S3 sur StorageGRID

S3cmd est un outil de ligne de commande gratuit et un client pour les opérations S3. Vous pouvez utiliser s3cmd pour tester et démontrer l'accès s3 avec StorageGRID.

## Installez et configurez S3cmd

Pour installer S3cmd sur un poste de travail ou un serveur, téléchargez-le à partir de "Client S3 en ligne de commande". S3cmd est préinstallé sur chaque nœud StorageGRID comme outil pour faciliter le dépannage.

## Étapes de configuration initiale

1. s3cmd --configure
2. Fournissez uniquement Access_Key et secret_key, pour le reste conservez les valeurs par défaut.
3. Tester l'accès avec les informations d'identification fournies ? [O/n] : n (ignorer le test car il échouera)
4. Enregistrer les paramètres ? [o/N] y
   a. Configuration enregistrée dans '/root/.s3cfg'
5. Dans les champs .s3cfg, rendre vide Host_base et Host_bucket après le signe "=" :
   a. host_base =
   b. host_bucket =

> ⓘ Si vous spécifiez Host_base et Host_bucket à l'étape 4, il n'est pas nécessaire de spécifier un noeud final avec --host dans la CLI. Exemple :

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

## Exemples de commandes de base

- **Créer un compartiment :**

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

- **Liste de tous les compartiments:**

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

- **Liste de tous les compartiments et de leur contenu:**

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

- **Liste des objets dans un compartiment spécifique :**

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

- **Supprimer un compartiment :**

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

- **Mettre un objet:**

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

- **Obtenir un objet:**

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

- **Supprimer un objet:**

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

*Par Aron Klein*

# Base de données en mode Vertica Eon utilisant NetApp StorageGRID comme stockage communautaire

Ce guide décrit la procédure de création d'une base de données Vertica Eon mode avec stockage communautaire sur NetApp StorageGRID.

## Introduction

Vertica est un logiciel de gestion de base de données analytique. C'est une plateforme de stockage orientée colonnes conçue pour gérer d'importants volumes de données, permettant ainsi des performances d'interrogation très rapides dans un scénario très intensif. Une base de données Vertica s'exécute dans l'un des deux modes suivants : EON ou Enterprise. Vous pouvez déployer les deux modes sur site ou dans le cloud.

Les modes EON et Enterprise diffèrent principalement lorsqu'ils stockent des données :

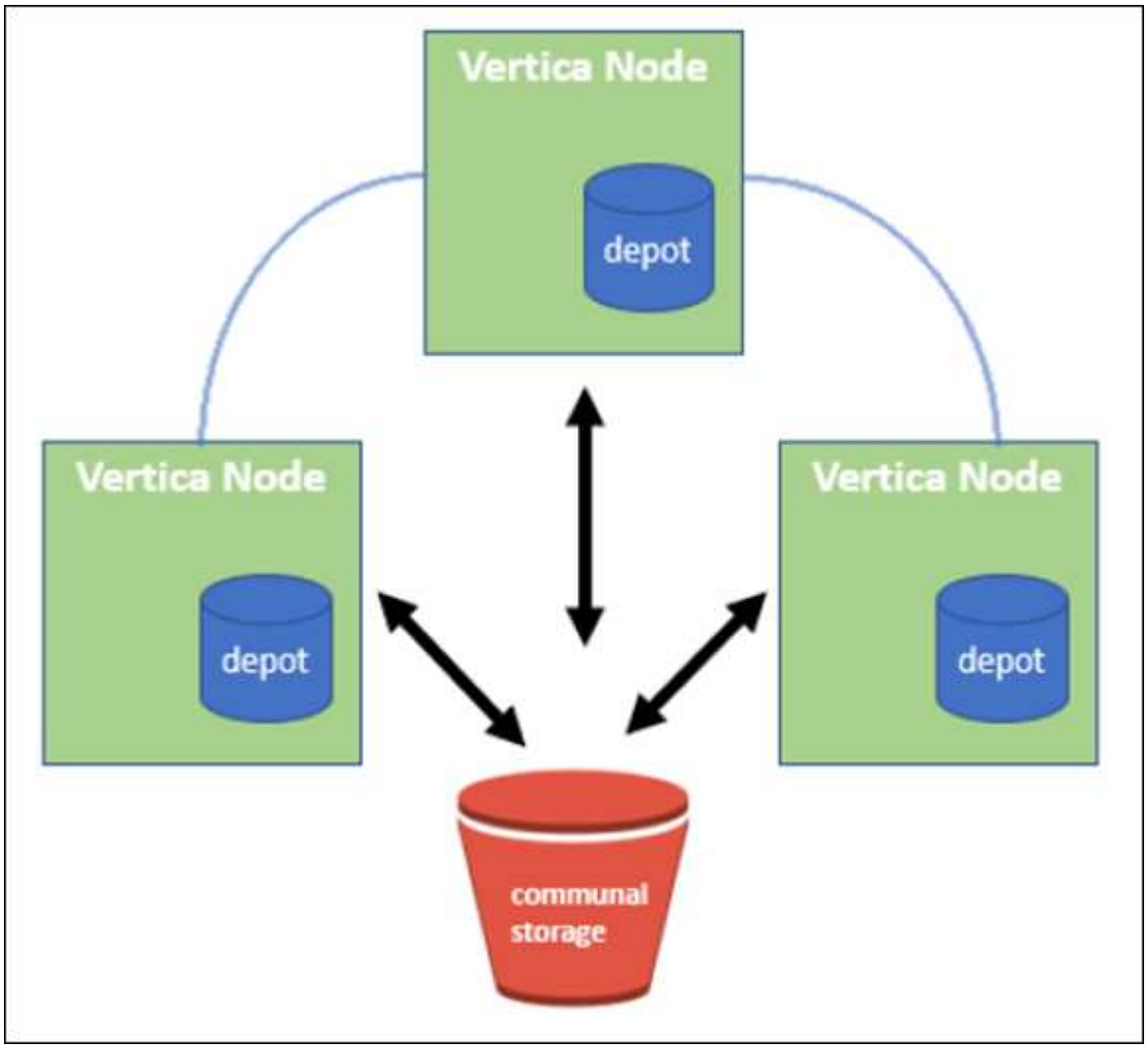
- Les bases de données du mode EON utilisent le stockage communautaire pour leurs données. Ceci est recommandé par Vertica.

- Les bases de données Enterprise mode stockent les données localement dans le système de fichiers des nœuds qui composent la base de données.

### Architecture du mode EON

Le mode EON sépare les ressources de calcul de la couche de stockage communautaire de la base de données, ce qui permet l'évolutivité séparée du calcul et du stockage. Vertica en mode Eon est optimisé pour traiter des charges de travail variables et les isoler les unes des autres à l'aide de ressources de calcul et de stockage distinctes.

EON mode stocke les données dans un magasin d'objets partagés appelé stockage communal : un compartiment S3, hébergé sur site ou sur Amazon S3.




### Stockage communautaire

Au lieu de stocker les données localement, le mode Eon utilise un emplacement de stockage commun unique pour toutes les données et le catalogue (métadonnées). Le stockage communal est l'emplacement de stockage centralisé de la base de données, partagé entre les nœuds de base de données.

Le stockage communal a les propriétés suivantes :

- Le stockage communautaire dans le cloud ou dans un stockage objet sur site est plus résilient et moins vulnérable aux pertes de données dues à des défaillances de stockage que sur un stockage sur disque sur

des machines individuelles.

- Toutes les données peuvent être lues par n'importe quel nœud utilisant le même chemin d'accès.
- La capacité n'est pas limitée par l'espace disque sur les nœuds.
- Les données étant stockées dans la communauté, vous pouvez faire évoluer votre cluster en toute flexibilité pour répondre aux besoins changeants. Si les données étaient stockées localement sur les nœuds, ajouter ou supprimer des nœuds nécessiterait un déplacement de grandes quantités de données entre les nœuds pour les déplacer hors des nœuds supprimés, ou vers les nœuds nouvellement créés.

**Le dépôt**

La vitesse est un inconvénient du stockage commun. L'accès aux données à partir d'un emplacement cloud partagé est plus lent que la lecture à partir d'un disque local. En outre, la connexion au stockage commun peut former un goulot d'étranglement si de nombreux nœuds lisent les données à partir de ce stockage en même temps. Pour améliorer la vitesse d'accès aux données, les nœuds d'une base de données en mode Eon maintiennent un cache de disque local de données appelé dépôt. Lors de l'exécution d'une requête, les nœuds vérifient d'abord si les données dont ils ont besoin se trouvent dans le dépôt. Si c'est le cas, il termine la requête en utilisant la copie locale des données. Si les données ne se trouvent pas dans le dépôt, le nœud extrait les données du stockage commun et enregistre une copie dans le dépôt.

## Recommandations de NetApp StorageGRID

Vertica stocke les données de base de données dans le stockage objet sous la forme de milliers (ou de millions) d'objets compressés (dont la taille observée est de 200 à 500 Mo par objet). Lorsqu'un utilisateur exécute des requêtes de base de données, Vertica récupère la plage de données sélectionnée à partir de ces objets compressés en parallèle à l'aide de l'appel GET de plage d'octets. Chaque PLAGE d'octets GET est d'environ 8 Ko.

Lors du test de requêtes utilisateur externes au dépôt de bases de données de 10 To, 4,000 à 10,000 REQUÊTES GET (OCTET-plage) par seconde ont été envoyées dans la grille. Lors de l'exécution de ce test avec des appliances SG6060, si le taux d'utilisation du processeur par nœud d'appliance est faible (environ 20 à 30 %), 2/3 le temps du processeur est en attente des E/S. Un très faible pourcentage (0 % à 0.5 %) d'attente d'E/S est observé sur le SGF6024.

En raison de la forte demande en IOPS peu élevées avec des latences très faibles (la moyenne doit être inférieure à 0.01 secondes), NetApp recommande l'utilisation du système SFG6024 pour les services de stockage objet. Si le SG6060 est nécessaire pour des bases de données très volumineuses, le client doit travailler avec l'équipe des comptes Vertica sur le dimensionnement du dépôt pour prendre en charge le dataset très interrogé.

Pour le nœud d'administration et le nœud de passerelle d'API, le client peut utiliser le SG100 ou le SG1000. Le choix dépend du nombre de requêtes des utilisateurs en parallèle et de la taille de la base de données. Si le client préfère utiliser un équilibreur de charge tiers, NetApp recommande un équilibreur de charge dédié pour une charge de travail hautes performances. Pour connaître le dimensionnement StorageGRID, consultez l'équipe de gestion de compte NetApp.

D'autres recommandations concernant la configuration de StorageGRID incluent :

- **Topologie de grille**. Ne mélangez pas le SGF6024 avec d'autres modèles d'appliance de stockage sur le même site de réseau. Si vous préférez utiliser le SG6060 pour la protection de l'archivage à long terme, conservez le SGF6024 avec un équilibreur de charge dédié dans son propre site de grid (site physique ou logique) pour une base de données active afin d'améliorer les performances. L'utilisation de différents modèles d'appliance sur le même site réduit les performances globales sur le site.
- **Protection des données**. Utilisez des copies répliquées pour la protection. N'utilisez pas le code

d'effacement pour une base de données active. Le client peut utiliser un code d'effacement pour protéger à long terme les bases de données inactives.

- **N'activez pas la compression de grille**. Vertica compresse les objets avant de les stocker dans le stockage objet. L'activation de la compression grid n'entraîne pas d'économie supplémentaire en matière d'utilisation du stockage et réduit considérablement les performances GET de plage d'octets.
- **Connexion de terminal HTTP et HTTPS S3**. Lors du test de banc d'essai, nous avons observé une amélioration des performances d'environ 5 % lors de l'utilisation d'une connexion HTTP S3 du cluster Vertica vers le point de terminaison de l'équilibreur de charge StorageGRID. Ce choix doit être basé sur les exigences de sécurité du client.

Les recommandations pour une configuration Vertica sont les suivantes :

- **Les paramètres de dépôt par défaut de la base de données Vertica sont activés (valeur = 1) pour les opérations de lecture et d'écriture**. NetApp recommande fortement de maintenir ces paramètres de dépôt activés pour améliorer les performances.
- **Désactiver les limitations de diffusion**. Pour plus de détails sur la configuration, reportez-vous à la section Désactivation des restrictions de diffusion en continu.

## Installation du mode Eon sur site avec stockage communautaire sur StorageGRID

Les sections suivantes décrivent la procédure, dans l'ordre, d'installation du mode Eon sur site avec un stockage communautaire sur StorageGRID. La procédure de configuration du stockage objet compatible S3 (simple Storage Service) sur site est similaire à la procédure décrite dans le guide Vertica, "Installez une base de données en mode Eon sur site".

La configuration suivante a été utilisée pour le test fonctionnel :

- StorageGRID 11.4.0.4
- Vertica 10.1.0
- Trois machines virtuelles (VM) avec CentOS 7.x OS pour les nœuds Vertica afin de former un cluster. Cette configuration est destinée uniquement au test fonctionnel, pas au cluster de base de données de production Vertica.

Ces trois nœuds sont configurés avec une clé Secure Shell (SSH) afin de permettre SSH sans mot de passe entre les nœuds du cluster.

### Informations requises par NetApp StorageGRID

Pour installer Eon mode sur site avec un stockage communautaire sur StorageGRID, vous devez disposer des informations de prérequis suivantes.

- Adresse IP ou nom de domaine complet (FQDN) et numéro de port du terminal StorageGRID S3. Si vous utilisez HTTPS, utilisez un certificat SSL personnalisé (autorité de certification) ou un certificat SSL auto-signé mis en œuvre sur le terminal StorageGRID S3.
- Nom du compartiment. Il doit exister au préalable et être vide.
- L'ID de clé et la clé d'accès secrète avec un accès en lecture et en écriture au compartiment.

### Création d'un fichier d'autorisation pour accéder au terminal S3

Les prérequis suivants s'appliquent lors de la création d'un fichier d'autorisation pour accéder au terminal S3 :

- Vertica est installé.
- Un cluster est configuré, configuré et prêt pour la création de bases de données.

Pour créer un fichier d'autorisation pour accéder au terminal S3, effectuez la procédure suivante :

1. Connectez-vous au nœud Vertica sur lequel vous allez exécuter `admintools` Pour créer la base de données du mode Eon.

   L'utilisateur par défaut est `dbadmin`, Créé lors de l'installation du cluster Vertica.

2. Utilisez un éditeur de texte pour créer un fichier sous le `/home/dbadmin` répertoire. Le nom du fichier peut être tout ce que vous voulez, par exemple, `sg_auth.conf`.
3. Si le terminal S3 utilise un port HTTP standard 80 ou HTTPS 443, ignorez le numéro de port. Pour utiliser HTTPS, définissez les valeurs suivantes :

   ◦ `awsenablehttps = 1`, sinon, définissez la valeur sur `0`.

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

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

   Pour utiliser un certificat SSL personnalisé ou auto-signé pour la connexion HTTPS du noeud final StorageGRID S3, spécifiez le chemin d'accès complet au fichier et le nom du fichier du certificat. Ce fichier doit se trouver au même emplacement sur chaque nœud de la Vertica et avoir des droits d'accès en lecture pour tous les utilisateurs. Ignorez cette étape si le certificat SSL du terminal StorageGRID S3 est signé par une autorité de certification publique.

   - `awscafile = <filepath/filename>`

   Par exemple, consultez le fichier d'exemple suivant :

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

+

> Dans un environnement de production, le client doit implémenter un certificat de serveur signé par une autorité de certification publique sur un terminal d'équilibrage de charge StorageGRID S3.

**Sélection d'un chemin de dépôt sur tous les nœuds de la Vertica**

Choisissez ou créez un répertoire sur chaque nœud pour le chemin de stockage du dépôt. Le répertoire que vous fournissez pour le paramètre chemin de stockage du dépôt doit avoir les éléments suivants :

- Le même chemin sur tous les nœuds du cluster (par exemple, `/home/dbadmin/depot`)
- Être lisible et inscriptible par l'utilisateur dbadmin

- Un stockage suffisant

  Par défaut, Vertica utilise 60 % de l'espace du système de fichiers contenant le répertoire pour le stockage du dépôt. Vous pouvez limiter la taille du dépôt en utilisant le `--depot-size` argument dans le `create_db` commande. Voir "Dimensionnement du cluster Vertica pour une base de données en mode Eon" article pour les directives générales de dimensionnement de la Vertica ou consultez votre gestionnaire de compte Vertica.

  Le `admintools create_db` l'outil tente de créer le chemin de dépôt pour vous si celui-ci n'existe pas.

**Création de la base de données Eon sur site**

Pour créer la base de données Eon sur site, procédez comme suit :

1. Pour créer la base de données, utilisez le `admintools create_db` outil.

   La liste suivante fournit une brève explication des arguments utilisés dans cet exemple. Consultez le document Vertica pour obtenir une explication détaillée de tous les arguments requis et facultatifs.

   - -x <chemin/nom de fichier d'autorisation créé dans « Création d'un fichier d'autorisation pour accéder au noeud final S3 » >.

     Les détails d'autorisation sont stockés dans la base de données après la création. Vous pouvez supprimer ce fichier pour éviter d'exposer la clé secrète S3.

   - --emplacement-communautaire-stockage <s3://storagegrid buckname>
   - -S <liste séparée par des virgules des nœuds de la Vertica à utiliser pour cette base de données>
   - -d <nom de la base de données à créer>
   - -p <mot de passe à définir pour cette nouvelle base de données>. Par exemple, reportez-vous à la commande d'exemple suivante :

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

     La création d'une nouvelle base de données prend plusieurs minutes en fonction du nombre de nœuds de la base de données. Lors de la création de la base de données pour la première fois, vous serez invité à accepter le contrat de licence.

Par exemple, reportez-vous à l'exemple de fichier d'autorisation suivant et `create db` commande :

```
[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.
```

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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/026d63ae9d4a33237bf0e2c2cf2a794a00a000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar |
| 8062464 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar |
| 20024832 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
|---|---|
| 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 |

| Taille de l'objet (octet) | Chemin d'accès complet de la clé de compartiment/objet |
| --- | --- |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |

#### Désactivation des restrictions de diffusion en continu

Cette procédure est basée sur le guide Vertica pour d'autres systèmes de stockage objet sur site et doit s'appliquer à StorageGRID.

1. Après avoir créé la base de données, désactivez le `AWSStreamingConnectionPercentage` paramètre de configuration en le définissant sur `0`. Ce paramètre n'est pas nécessaire pour une installation sur site en mode Eon avec stockage communautaire. Ce paramètre de configuration contrôle le nombre de connexions au magasin d'objets utilisé par Vertica pour les lectures en continu. Dans un environnement cloud, ce paramètre évite que les données en streaming à partir du magasin d'objets utilisent tous les descripteurs de fichier disponibles. Certains poignées de fichiers restent disponibles pour d'autres opérations de stockage d'objets. En raison de la faible latence des magasins d'objets sur site, cette option n'est pas nécessaire.
2. Utiliser un `vsql` instruction permettant de mettre à jour la valeur du paramètre. Le mot de passe est le mot de passe de la base de données que vous avez défini dans la section "création de la base de données Eon sur site". Par exemple, reportez-vous à l'exemple de résultat suivant :

```
[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
```

#### Vérification des paramètres du dépôt

Les paramètres de dépôt par défaut de la base de données Vertica sont activés (valeur = 1) pour les opérations de lecture et d'écriture. NetApp recommande fortement de maintenir ces paramètres de dépôt activés pour améliorer les performances.

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

#### Chargement des données d'échantillon (facultatif)

Si cette base de données est destinée aux tests et sera supprimée, vous pouvez charger des données

d'échantillon dans cette base de données pour les tests. Vertica est fourni avec un exemple de jeu de données, VMart, sous `/opt/vertica/examples/VMart_Schema/` Sur chaque nœud Vertica. Vous trouverez plus d'informations sur cet exemple de jeu de données "ici".

Procédez comme suit pour charger les données d'échantillon :

1. Connectez-vous en tant que dbadmin à l'un des nœuds de la Vertica : cd /opt/vertica/sou/VMart_Schema/
2. Chargez les exemples de données dans la base de données et entrez le mot de passe de la base de données lorsque vous y êtes invité dans les sous-étapes c et d :
   a. `cd /opt/vertica/examples/VMart_Schema`
   b. `./vmart_gen`
   c. `vsql < vmart_define_schema.sql`
   d. `vsql < vmart_load_data.sql`
3. Il existe plusieurs requêtes SQL prédéfinies, vous pouvez les exécuter pour confirmer que les données de test sont chargées correctement dans la base de données. Par exemple : `vsql < vmart_queries1.sql`

### Où trouver des informations complémentaires

Pour en savoir plus sur les informations données dans ce livre blanc, consultez ces documents et/ou sites web :

- "Documentation du produit NetApp StorageGRID 11.7"
- "Fiche technique StorageGRID"
- "Documentation produit de Vertica 10.1"

### Historique des versions

| Version | Date | Historique des versions du document |
|---|---|---|
| Version 1.0 | Septembre 2021 | Version initiale. |

*Par Angela Cheng*

# Analyse des journaux StorageGRID à l'aide de la pile ELK

Avec la fonctionnalité StorageGRID 11.6 syslog Forward, vous pouvez configurer un serveur syslog externe afin de collecter et d'analyser les messages journaux StorageGRID. ELK (Elasticsearch, Logstash, Kibana) est devenu l'une des solutions d'analytique des journaux les plus populaires. Regardez la "Analyse du journal StorageGRID à l'aide de la vidéo ELK" Pour afficher un exemple de configuration ELK et comment elle peut être utilisée pour identifier et dépanner les demandes S3 en échec. Cet article fournit des exemples de fichiers de configuration Logstash, de requêtes Kibana, de graphiques et de tableau de bord, pour vous offrir un démarrage rapide de la gestion des journaux et de l'analytique StorageGRID.

## De formation

- StorageGRID 11.6.0.2 ou version ultérieure
- ELK (Elasticsearch, Logstash et Kibana) 7.1x ou plus installé et en fonctionnement

## Exemples de fichiers

- "Téléchargez le paquet Logstash 7.x." + **md5 checksum** 148c23d0021d9a4bb4a6c0287464deab + **sha256 checksum** f51ec9e2e3f842d5a781566b167a561b4373038b4e7bb3c8b5d52f2f2d2f2f2f2f2d2f2f2f2f2f2f2f2f6f6f
- "Téléchargez le paquet Logstash 8.x." + **md5 checksum** e11bae3a662f87c310ef363d0fe06835 + **total de contrôle sha256** 5c670755742cfdfd5aa723a596ba087e0153a65bcaef3934afdb682f61cd278d

## Hypothèse

Les lecteurs connaissent la terminologie et les opérations de StorageGRID et d'ELK.

## Instructions

Deux exemples de versions sont fournis en raison des différences de noms définies par des motifs grk. + par exemple, le modèle SYSLOGBASE grok dans le fichier de configuration Logstash définit les noms de champs différemment en fonction de la version Logstash installée.

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

**Logstash 7.17 échantillon**

| 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 échantillon**

Table JSON

Search field names

| Actions | Field | Value |
| --- | --- | --- |
| | _id | yuhOiIEBVP6KX4EwqcyU |
| | _index | sglog-2022.06.21 |
| | # _score | - |
| | @timestamp | Jun 21, 2022 @ 18:07:45.444 |
| | t event.original | <28>Jun 21 22:07:45 SITE2-S3 ADE: syslog messages being dropped |
| | t host.hostname | SITE2-S3 |
| | t msg-details | syslog messages being dropped |
| | t process.name | ADE |
| | t syslog_pri | 28 |
| | t timestamp | Jun 21 22:07:45 |

**Étapes**

1. Décompressez l'échantillon fourni en fonction de la version ELK installée. + l'exemple de dossier inclut deux exemples de configuration de Logstash : + **sglog-2-file.conf:** ce fichier de configuration envoie des messages de journal StorageGRID vers un fichier sur Logstash sans transformation de données. Vous pouvez l'utiliser pour confirmer que Logstash reçoit des messages StorageGRID ou pour vous aider à comprendre les modèles de journaux StorageGRID. + **sglog-2-es.conf:** ce fichier de configuration transforme les messages du journal StorageGRID en utilisant divers modèles et filtres. Il comprend des exemples d'instructions de DROP, qui sont basées sur des motifs ou des filtres. Le résultat est envoyé à Elasticsearch pour l'indexation. + Personnalisez le fichier de configuration sélectionné en fonction de l'instruction dans le fichier.
2. Testez le fichier de configuration personnalisé :

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

   Si la dernière ligne renvoyée est similaire à la ligne ci-dessous, le fichier de configuration n'a pas d'erreurs de syntaxe :

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

3. Copiez le fichier conf personnalisé dans la configuration du serveur Logstash : /etc/logstash/conf.d + si vous n'avez pas activé config.reload.automatic dans /etc/logstash/logstash.yml, redémarrez le service Logstash. Dans le cas contraire, attendez que l'intervalle de rechargement de la configuration s'écoule.

```
 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. Vérifiez /var/log/logstash/logstash-plain.log et assurez-vous qu'il n'y a pas d'erreur lors du démarrage de Logstash avec le nouveau fichier de configuration.
5. Vérifiez que le port TCP est démarré et que vous écoutez. + dans cet exemple, le port TCP 5000 est utilisé.

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

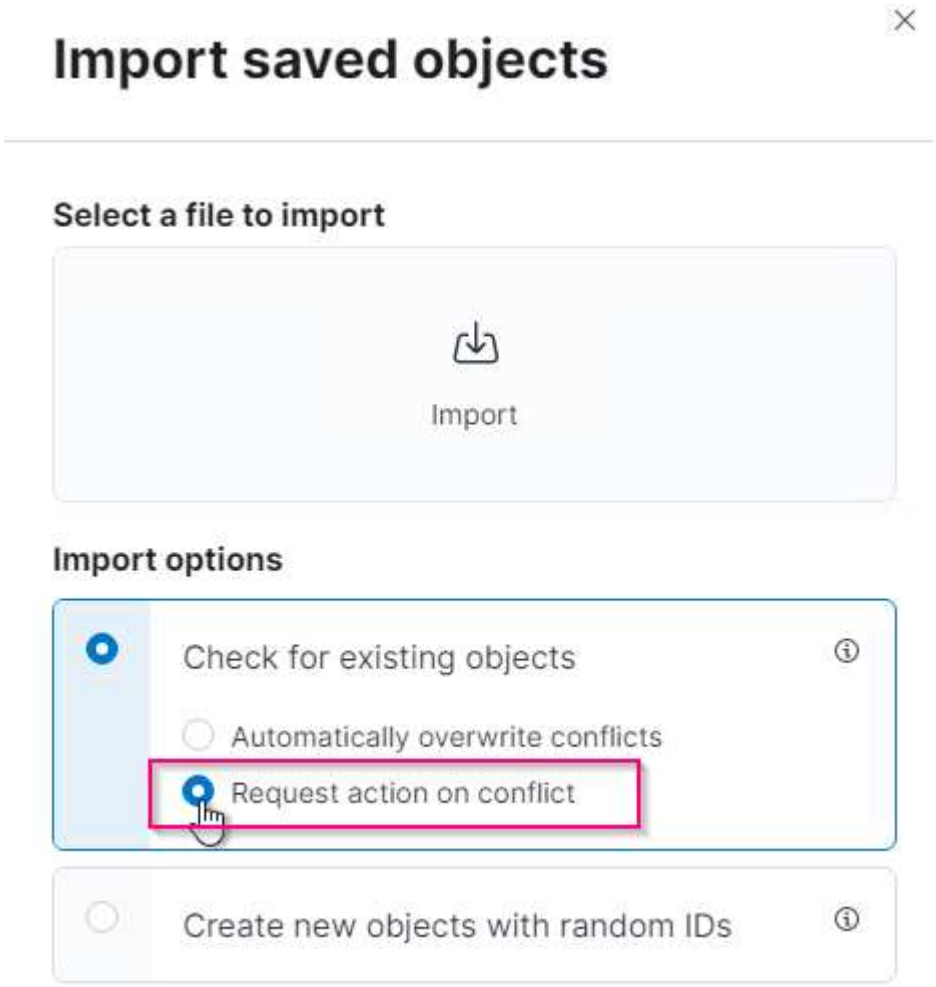
6. À partir de l'interface graphique du gestionnaire StorageGRID, configurez le serveur syslog externe pour envoyer des messages de journal à Logstash. Reportez-vous à la "vidéo de démonstration" pour plus d'informations.
7. Vous devez configurer ou désactiver le pare-feu sur le serveur Logstash pour autoriser la connexion des

nœuds StorageGRID au port TCP défini.

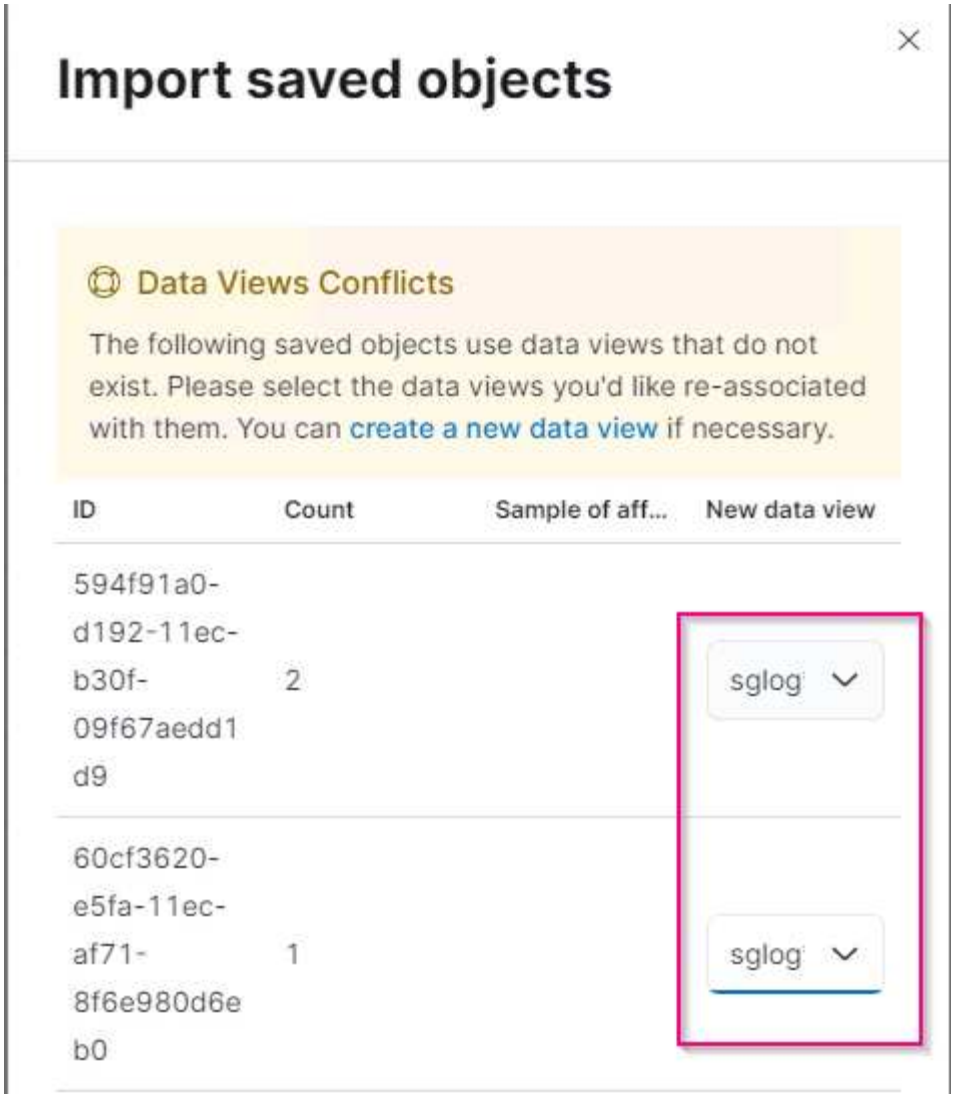
8. Dans l'interface graphique Kibana, sélectionnez Management → Dev Tools. Sur la page Console, exécutez cette commande OBTENIR pour confirmer la création de nouveaux index sur Elasticsearch.

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

9. A partir de l'interface graphique Kibana, créez un motif d'index (ELK 7.x) ou une vue de données (ELK 8.x).
10. Dans l'interface utilisateur graphique de Kibana, entrez « objets lavés » dans la zone de recherche située en haut au centre. + sur la page objets enregistrés, sélectionnez Importer. Sous Options d'importation, sélectionnez « demander une action en cas de conflit »




Importez elk<version>-query-chart-sample.ndjson. + lorsque vous êtes invité à résoudre le conflit, sélectionnez le modèle d'index ou la vue de données que vous avez créé à l'étape 8.




Les objets Kibana suivants sont importés : + **Query** + * audit-msg-s3rq-orlm + * journal de distribution messages liés s3 + * avertissement de niveau de journalisation ou supérieur + * échec de sécurité + **Chart** + * nombre de requêtes s3 basé sur bycast.log + * code d'état HTTP + * décomposition du msg d'audit par type + * réponse moyenne s3 Temps + **Tableau de bord** + * Tableau de bord de demande S3 à l'aide des tableaux ci-dessus.

Vous êtes maintenant prêt à effectuer une analyse des journaux StorageGRID à l'aide de Kibana.

## Ressources supplémentaires

- "syslog101"
- "Qu'est-ce que la pile ELK"
- "Liste des répétitions Grok"
- "Guide débutant de Logstash: Grok"
- "Guide pratique de Logstash : plongée en profondeur syslog"
- "Guide Kibana - Explorez le document"
- "Référence des messages du journal d'audit StorageGRID"

*Par Angela Cheng*

# Grâce à Prometheus et Grafana, vous pouvez renforcer la conservation des metrics

Ce rapport technique fournit des instructions détaillées sur la configuration de NetApp StorageGRID 11.6 avec des services Prometheus et Grafana externes.

## Introduction

StorageGRID stocke les metrics à l'aide de Prometheus et fournit des visualisations de ces metrics via des tableaux de bord intégrés. Vous pouvez accéder en toute sécurité aux metrics Prometheus depuis StorageGRID en configurant des certificats d'accès client et en activant l'accès prometheus pour le client spécifié. Aujourd'hui, la conservation de ces données de mesure est limitée par la capacité de stockage du nœud d'administration. Pour gagner plus de temps et pouvoir créer des visualisations personnalisées de ces metrics, nous déploierons un nouveau serveur Prometheus et Grafana, configurerons notre nouveau serveur afin de grattez les metrics à partir de l'instance IDS, et nous concevons un tableau de bord avec les mesures importantes. Vous pouvez obtenir plus d'informations sur les metrics Prometheus collectés dans la "Documentation StorageGRID".

## Fédérer Prometheus

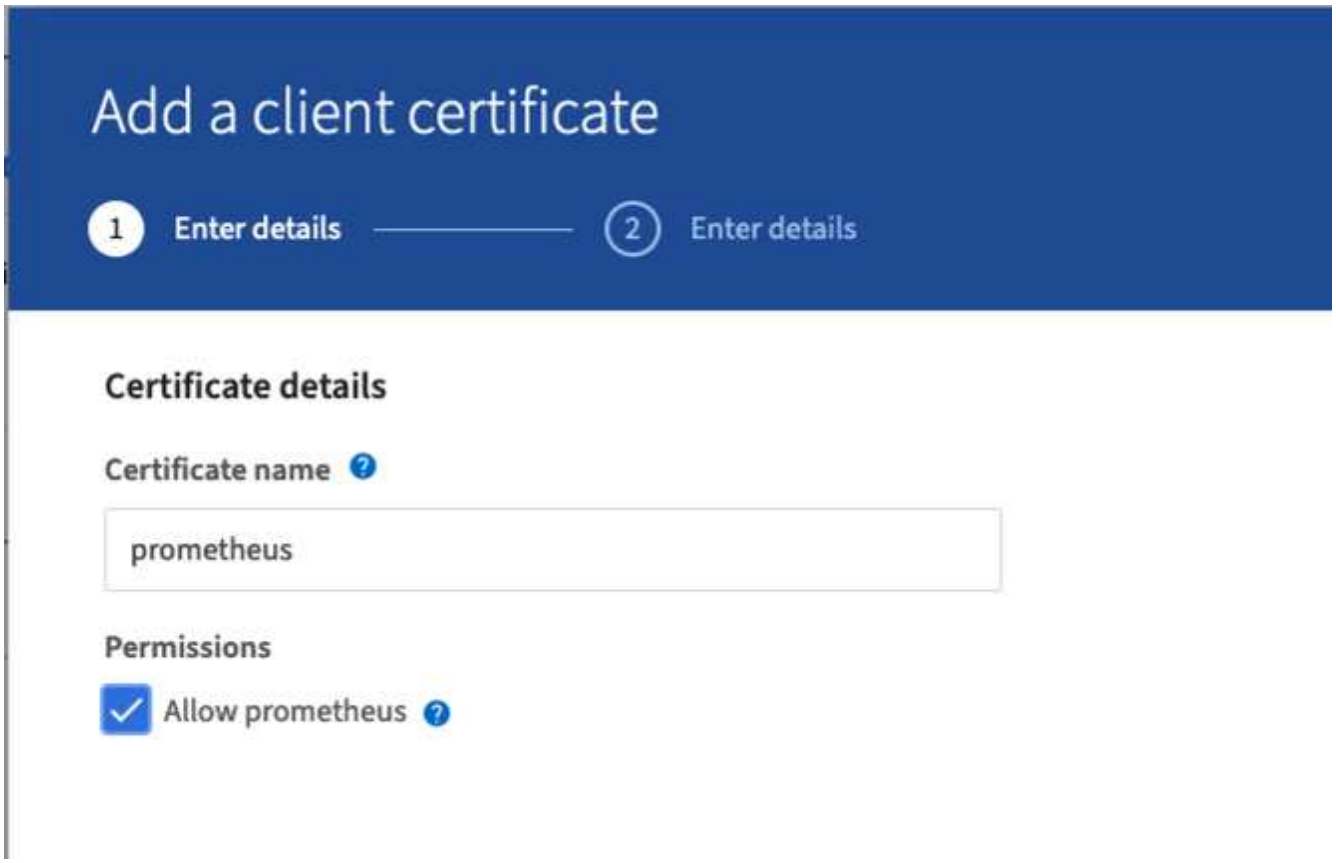
### Détails de laboratoire

Pour les besoins de cet exemple, j'utiliserai toutes les machines virtuelles pour les nœuds StorageGRID 11.6 et un serveur Debian 11. L'interface de gestion StorageGRID est configurée avec un certificat d'autorité de certification public approuvé. Cet exemple ne passera pas par l'installation et la configuration du système StorageGRID ou de l'installation de Debian linux. Vous pouvez utiliser toutes les versions Linux que vous souhaitez prendre en charge par Prometheus et Grafana. Prometheus et Grafana peuvent être installés en tant que conteneurs docker, qu'ils soient issus de la source ou binaires précompilés. Dans cet exemple, je vais installer les binaires Prometheus et Grafana directement sur le même serveur Debian. Téléchargez et suivez les instructions d'installation de base sur https://prometheus.io et https://grafana.com/grafana/ respectivement.

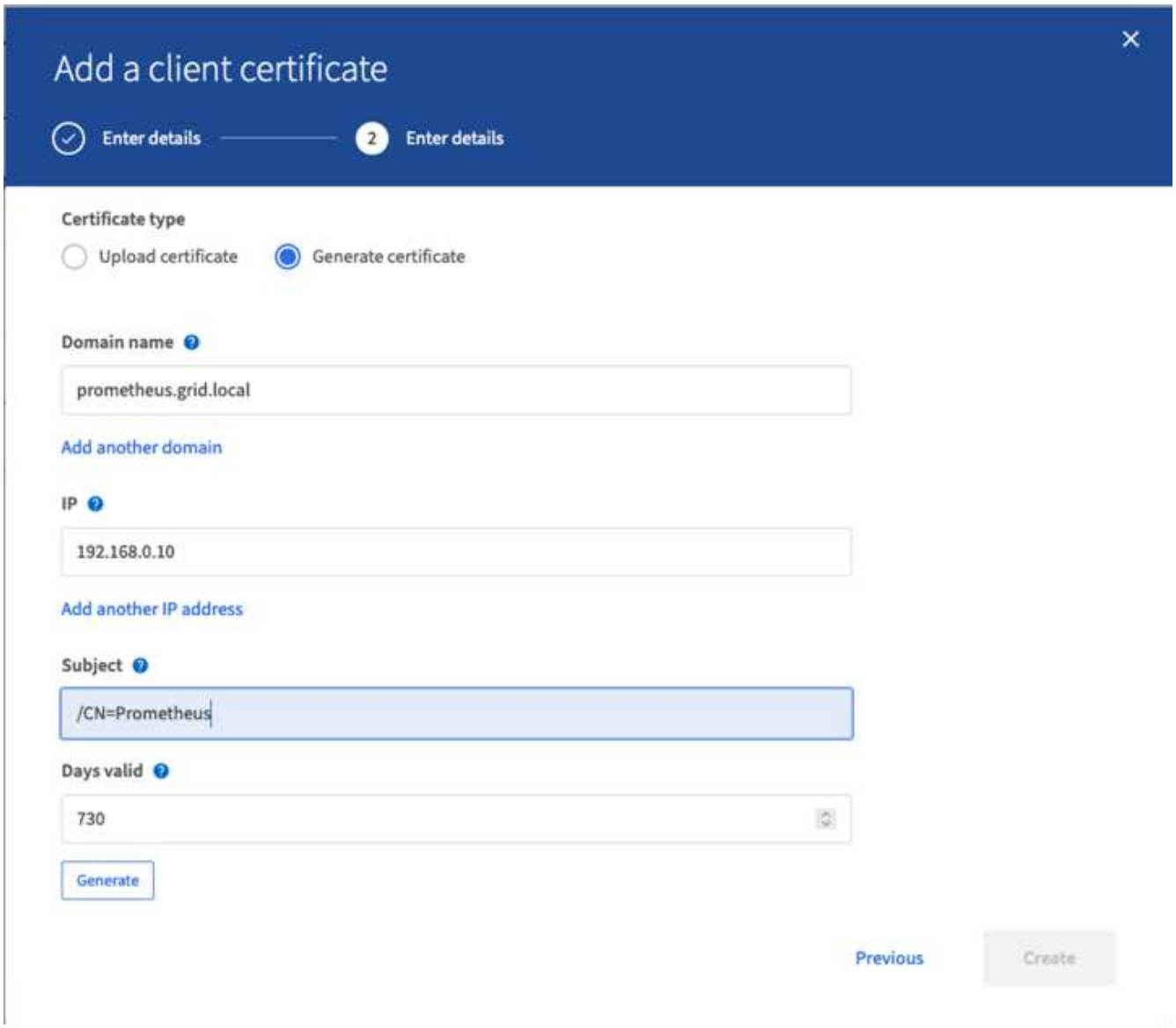
### Configurez StorageGRID pour l'accès client Prometheus

Afin d'accéder aux identifiants de paramètres de la mémoire de la solution, vous devez générer ou télécharger un certificat client avec une clé privée et activer l'autorisation pour le client. L'interface de gestion StorageGRID doit posséder un certificat SSL. Ce certificat doit être approuvé par le serveur prometheus soit par une autorité de certification approuvée, soit manuellement approuvé s'il est auto-signé. Pour en savoir plus, consultez le "Documentation StorageGRID".

1. Dans l'interface de gestion StorageGRID, sélectionnez « CONFIGURATION » en bas à gauche, puis dans la deuxième colonne sous « sécurité », cliquez sur certificats.
2. Sur la page certificats, sélectionnez l'onglet « client » et cliquez sur le bouton « Ajouter ».
3. Indiquez un nom pour le client auquel l'accès sera accordé et utilisez ce certificat. Cliquez sur la case sous "permissions", devant "Autoriser Prometheus" et cliquez sur le bouton Continuer.

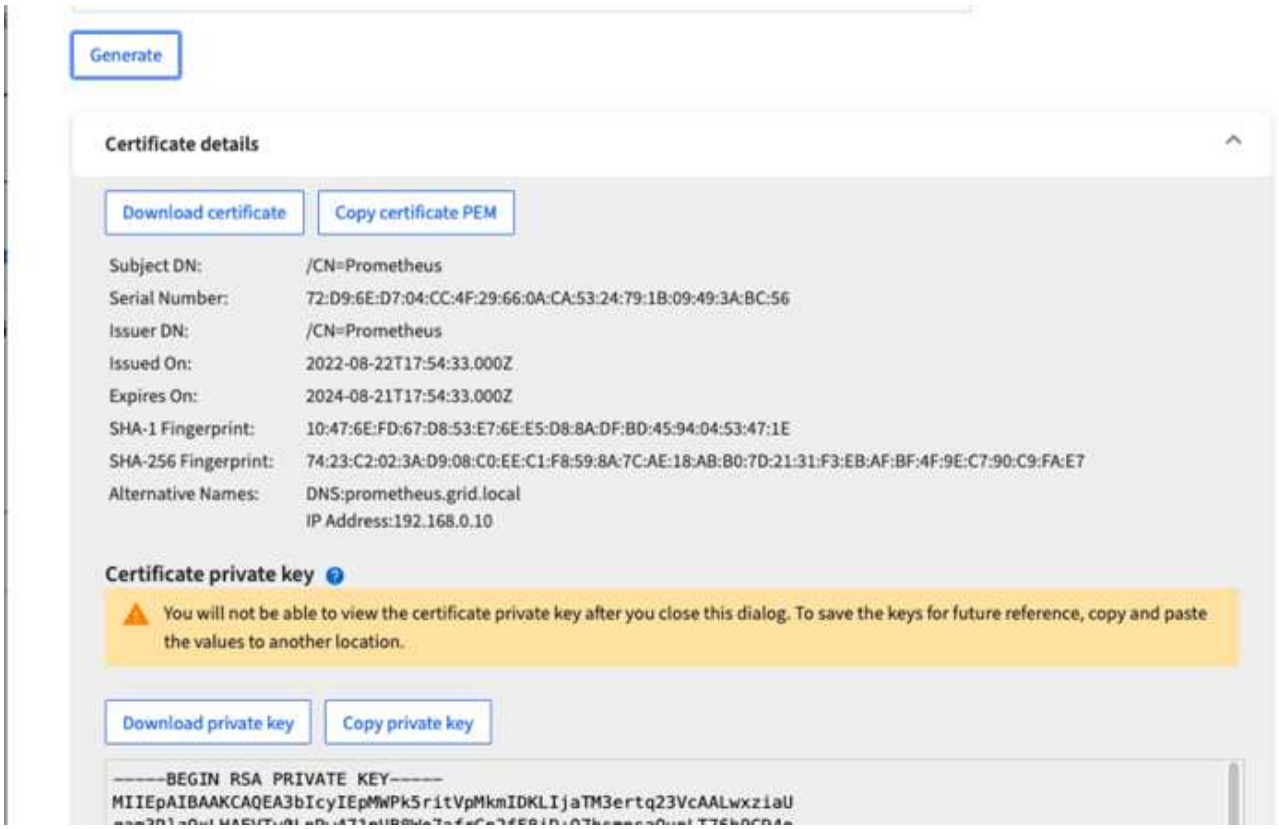



4. Si vous disposez d'un certificat signé par l'autorité de certification, vous pouvez sélectionner le bouton radio « Télécharger le certificat », mais dans notre cas, nous allons permettre à StorageGRID de générer le certificat client en sélectionnant le bouton radio « générer le certificat ». Les champs obligatoires s'affichent pour être renseignés. Saisissez le FQDN du serveur client, l'adresse IP du serveur, l'objet et les jours valides. Cliquez ensuite sur le bouton « générer ».




> 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. Téléchargez le fichier pem de certificat et le fichier pem de clé privée.




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

**Préparez le serveur Linux pour l'installation de Prometheus**

Avant d'installer Prometheus, je souhaite préparer mon environnement avec un utilisateur Prometheus, la structure de répertoires et configurer la capacité pour l'emplacement de stockage des metrics.

1. Créez l'utilisateur Prometheus.

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

2. Créez les répertoires pour les données Prometheus, les certificats client et les metrics.

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

3. J'ai formaté le disque que j'utilise pour la rétention des metrics avec un système de fichiers ext4.

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

4. Je ai ensuite monté le système de fichiers dans le répertoire des metrics de Prometheus.

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

5. Obtenez l'UUID du disque que vous utilisez pour les données de metrics.

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

6. Ajout d'une entrée dans /etc/fstab/ pour que le montage persiste entre les redémarrages à l'aide de l'UUID de /dev/sdb.

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

**Installez et configurez Prometheus**

Lorsque le serveur est prêt, je peux commencer l'installation de Prometheus et configurer le service.

1. Extraire le pack d'installation Prometheus

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

2. Copiez les binaires dans /usr/local/bin et modifiez la propriété de l'utilisateur prometheus créé précédemment

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

3. Copiez les consoles et les bibliothèques dans /etc/prometheus

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

4. Copiez le certificat client et les fichiers pem de clé privée téléchargés précédemment de StorageGRID vers /etc/prometheus/certs
5. Créez le fichier yaml de configuration prometheus

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

6. Insérez la configuration suivante. Le nom du travail peut être tout ce que vous souhaitez. Remplacez les « cibles : ['] » par le FQDN du nœud admin, et si vous avez modifié les noms des certificats et des fichiers de clé privée, mettez à jour la section tls_config pour qu'elle corresponde. enregistrez ensuite le fichier. Si votre interface de gestion de grille utilise un certificat auto-signé, téléchargez le certificat et placez-le avec un nom unique, et dans la section tls_config, ajoutez ca_file: /Etc/prometheus/cert/UIcert.pem

   a. Dans cet exemple, je collecterai tous les metrics commençant par alertManager, cassandra, nœud et StorageGRID. Vous trouverez plus d'informations sur les metrics Prometheus dans la "Documentation StorageGRID".

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

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

Si votre interface de gestion du grid utilise un certificat auto-signé, téléchargez le certificat et placez-le avec le certificat client portant un nom unique. Dans la section tls_config, ajoutez le certificat au-dessus du certificat client et des lignes de clé privée

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

1. Modifiez la propriété de tous les fichiers et répertoires dans /etc/prometheus et /var/lib/prometheus pour l'utilisateur prometheus

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

2. Créez un fichier de service prometheus dans /etc/systemd/system

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

3. Insérez les lignes suivantes, notez le --Storage.tsdb.rerétention=1A qui définit la conservation des données de mesure sur 1 an. Vous pouvez également utiliser --Storage.tsdb.Retention.size=300 Gio pour la conservation sur les limites de stockage. C'est le seul emplacement pour définir la conservation des métriques.

```
[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. Rechargez le service systemd pour enregistrer le nouveau service prometheus. démarrez et activez ensuite le service prometheus.

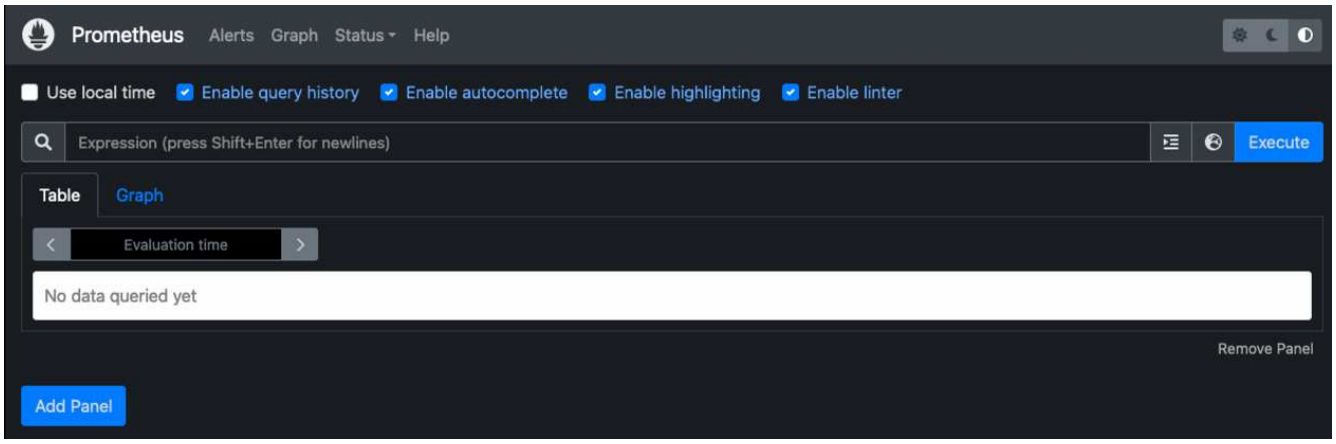
```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Vérifiez que l'entretien fonctionne correctement

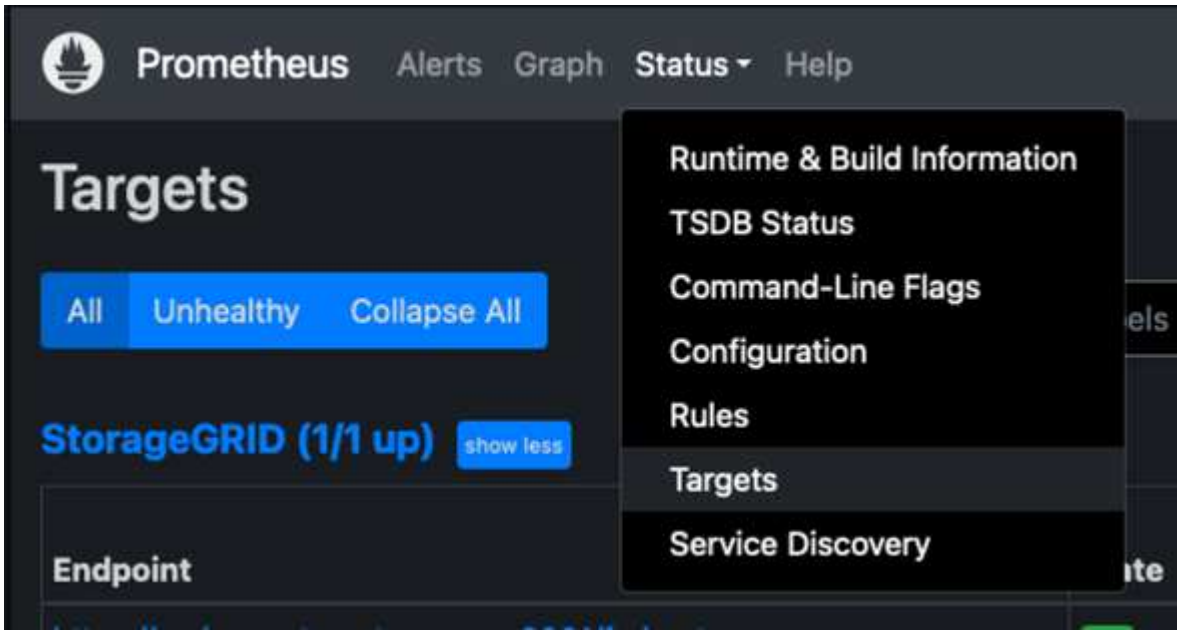
```
sudo systemctl status prometheus
```

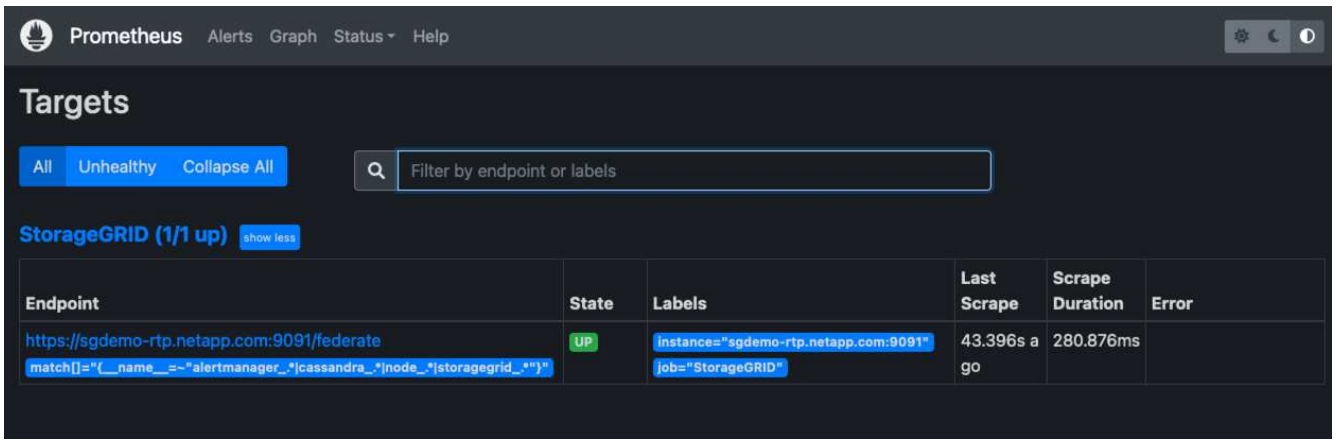
```
● prometheus.service - Prometheus Time Series Collection and Processing
Server
     Loaded: loaded (/etc/systemd/system/prometheus.service; enabled;
vendor preset: enabled)
     Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago
   Main PID: 6498 (prometheus)
      Tasks: 13 (limit: 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. Vous devez maintenant pouvoir naviguer vers l'interface du serveur prometheus http://Prometheus-server:9090 Et voir l'interface utilisateur

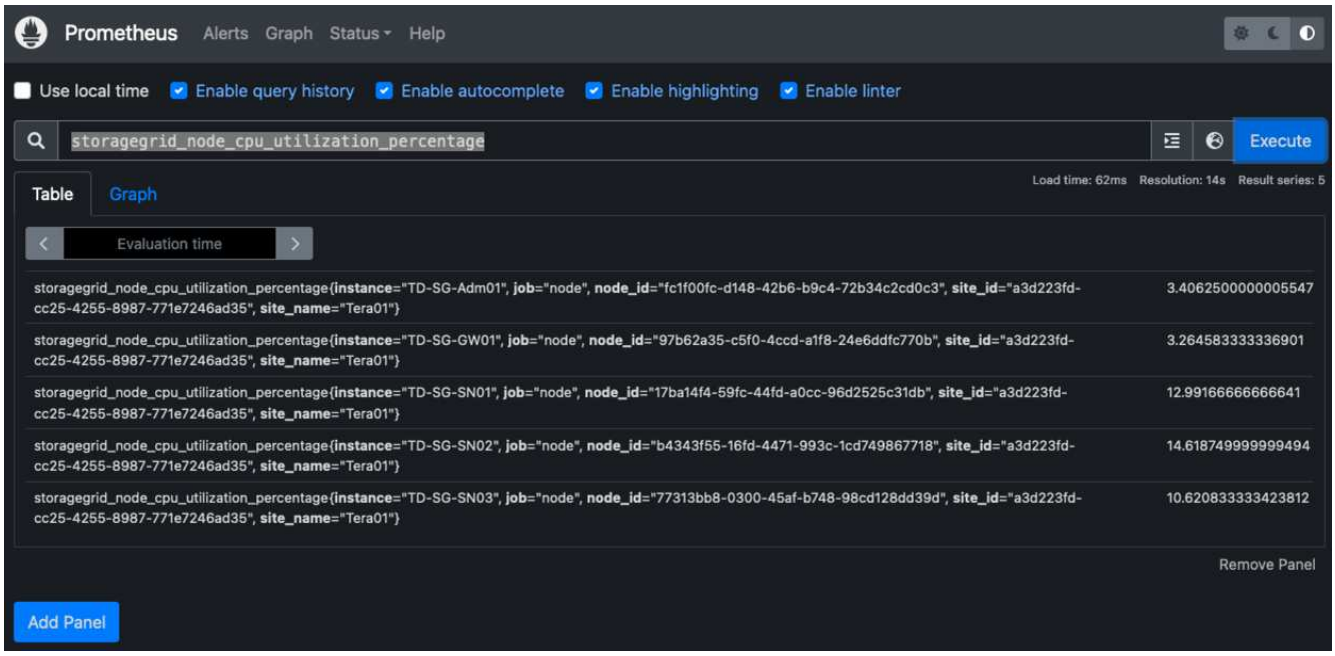



7. Sous cibles « Status », vous pouvez consulter le statut du noeud final StorageGRID configuré dans prometheus.yml




8. Sur la page graphique, vous pouvez exécuter une requête de test et vérifier que les données sont scrapées avec succès. Par exemple, entrez « storagegrid_node_cpu_usage_percent » dans la barre de requêtes et cliquez sur le bouton Exécuter.




## Installer et configurer Grafana

Vous pouvez désormais installer Grafana et configurer un tableau de bord

### Grafana Installation

1. Installez la dernière édition Enterprise de Grafana

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

2. Ajouter ce référentiel pour les versions stables :

```
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. Après avoir ajouté le référentiel.

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

4. Rechargez le service systemd pour enregistrer le nouveau service grafana. Démarrez et activez ensuite le service Grafana.

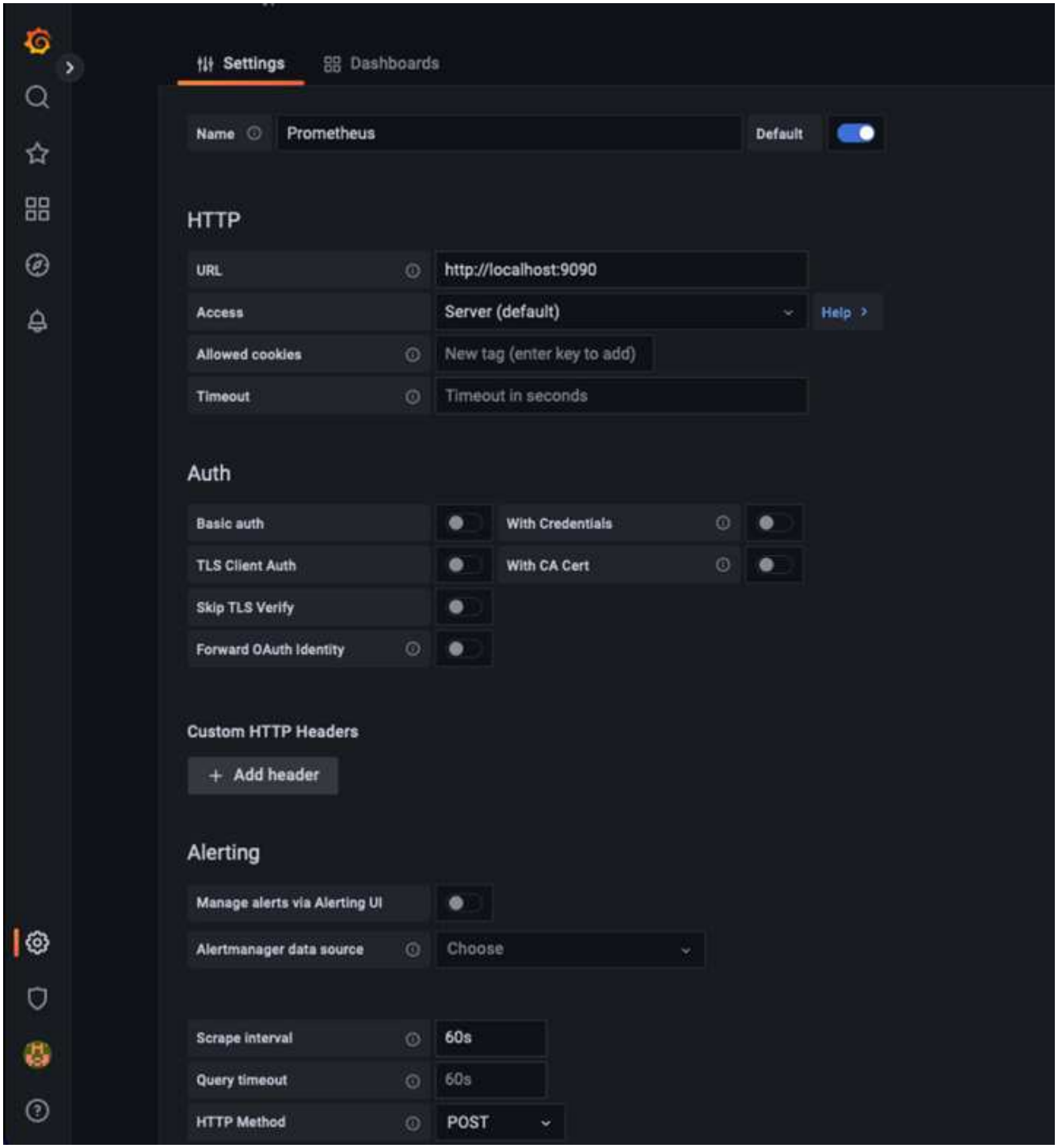
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. Grafana est désormais installé et exécuté. Lorsque vous ouvrez un navigateur vers HTTP://Prometheus-Server:3000, vous êtes accueilli par la page de connexion de Grafana.
6. Les informations d'identification par défaut sont admin/admin et vous devez définir un nouveau mot de passe à mesure qu'il vous invite à.

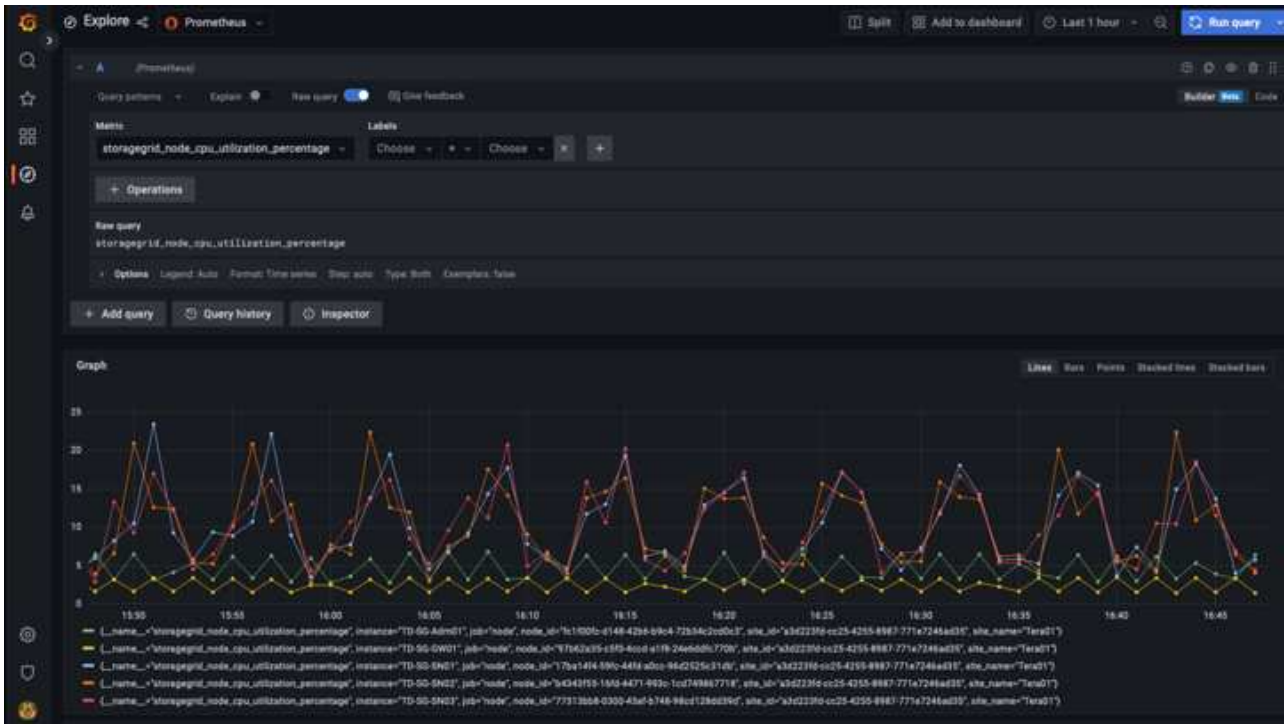
**Créez un tableau de bord Grafana pour StorageGRID**

Lorsque vous installez et exécutez Grafana et Prometheus, vous pouvez désormais vous connecter en créant une source de données et en créant un tableau de bord

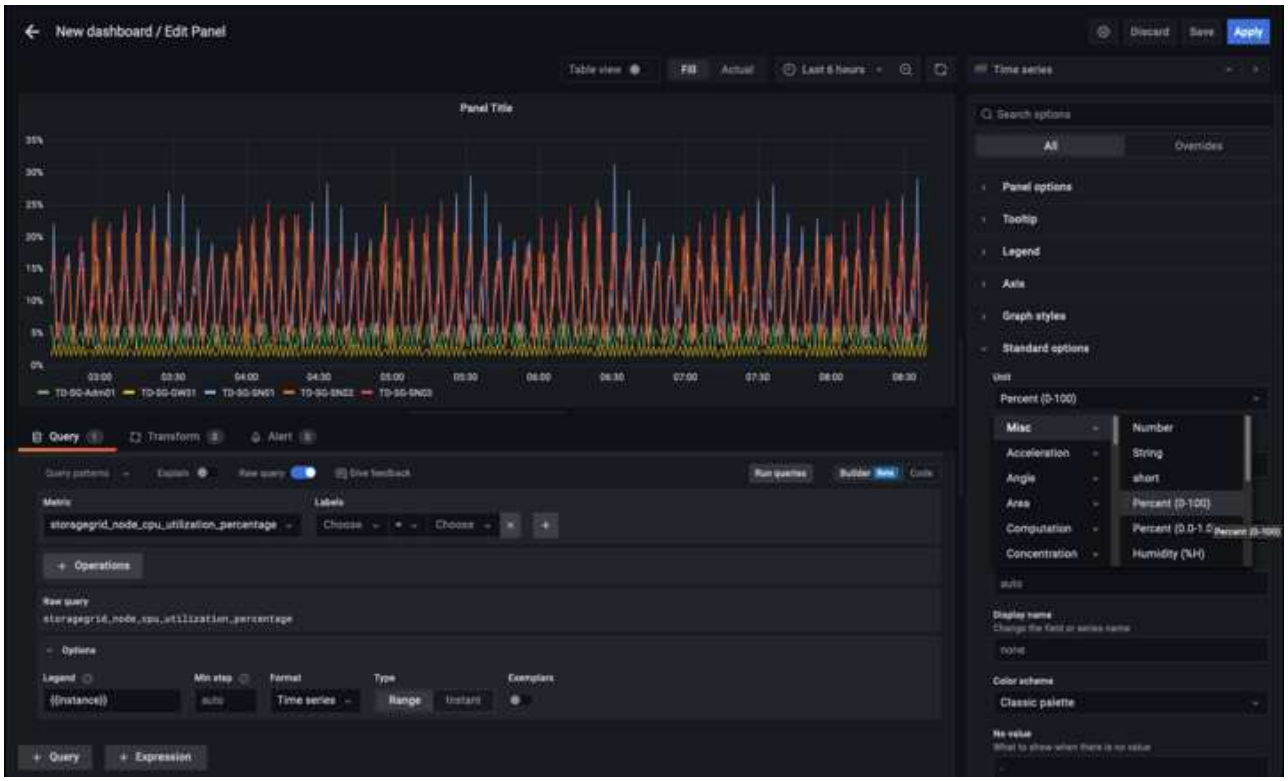
1. Dans le volet de gauche, développez « Configuration » et sélectionnez « sources de données », puis cliquez sur le bouton « Ajouter une source de données »
2. Prometheus est une des principales sources de données. Si ce n'est pas le cas, utilisez la barre de recherche pour trouver Prometheus
3. Configurez la source Prometheus en entrant l'URL de l'instance prometheus et l'intervalle de récupération en fonction de l'intervalle Prometheus. J'ai également désactivé la section d'alertes car je n'ai pas configuré le gestionnaire d'alertes sur prometheus.




4. Une fois les paramètres souhaités saisis, faites défiler l'écran vers le bas et cliquez sur « Enregistrer et tester ».
5. Une fois le test de configuration réussi, cliquez sur le bouton Explorer.
   a. Dans la fenêtre d'exploration, vous pouvez utiliser la même mesure que Prometheus testée avec « storagegrid_node_cpu_use_percent », puis cliquez sur le bouton Run Query



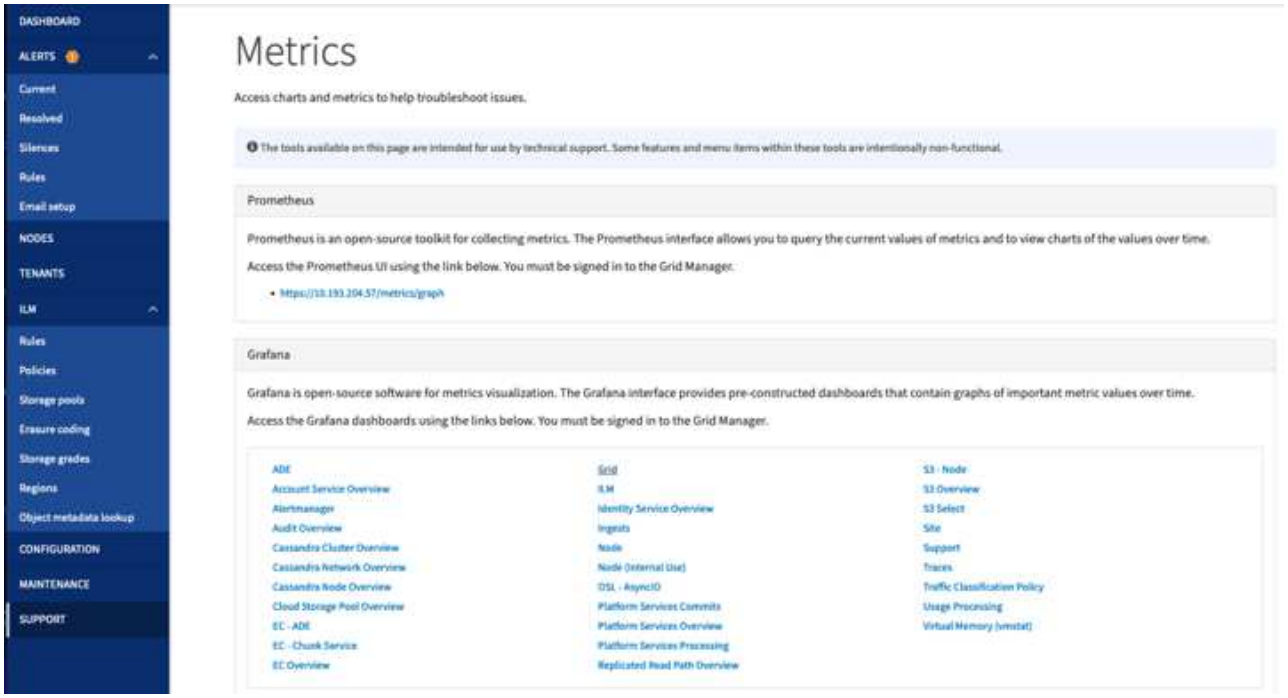

6. Comme la source de données est configurée, nous pouvons créer un tableau de bord.
   a. Dans le volet de gauche, développez « tableaux de bord » et sélectionnez « + nouveau tableau de bord ».
   b. Sélectionnez « Ajouter un nouveau panneau »
   c. Configurez le nouveau panneau en sélectionnant une mesure, puis j'utiliserai à nouveau « storagegrid_node_cpu_use_percentage », saisissez un titre pour le panneau, développez « Options » en bas et pour changer de légende en personnalisé et entrez «{{instance} » pour définir les noms de nœud, et à droite sous « Options standard » définissez « unité » sur « 100 % ». Cliquez ensuite sur « appliquer » pour enregistrer le panneau dans le tableau de bord.




7. Nous pouvons continuer à concevoir notre tableau de bord de ce type pour chaque metric souhaité, mais heureusement que StorageGRID dispose déjà de tableaux de bord avec des panneaux que nous pouvons copier dans nos tableaux de bord personnalisés.

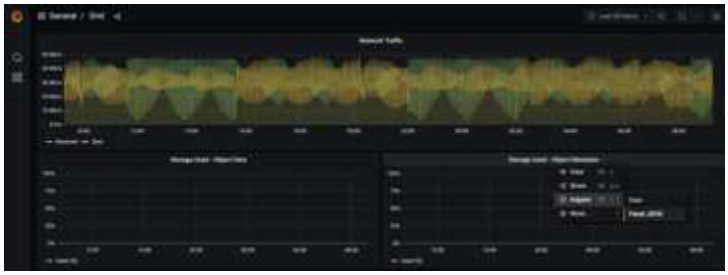
   a. Dans le volet gauche de l'interface de gestion StorageGRID, sélectionnez « support », et en bas de la colonne « Outils », cliquez sur métriques.

   b. Dans les mesures, je vais sélectionner le lien « grille » en haut de la colonne centrale.



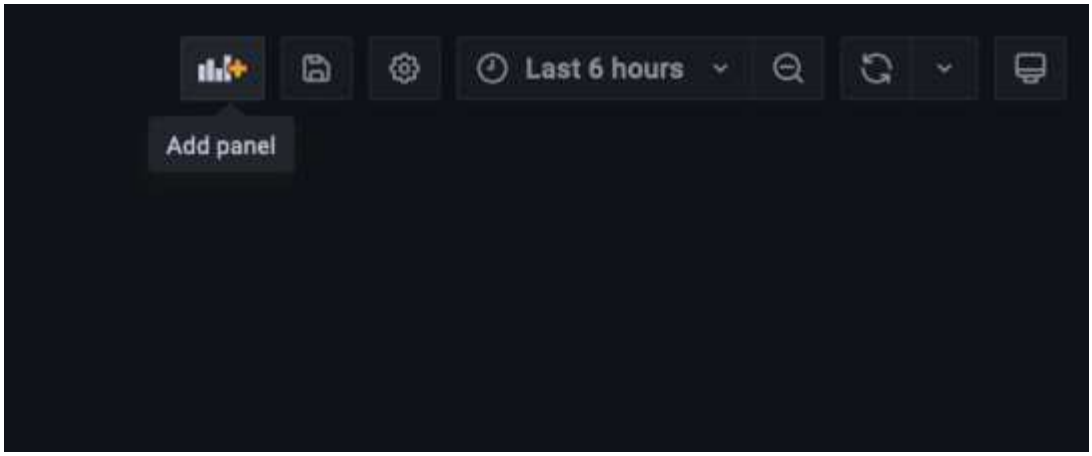

   c. Dans le tableau de bord Grid, sélectionnez le panneau « stockage utilisé - métadonnées de l'objet ».

Cliquez sur la petite flèche vers le bas et sur la fin du titre du panneau pour faire descendre un menu. Dans ce menu, sélectionnez « inspection » et « panneau JSON ».



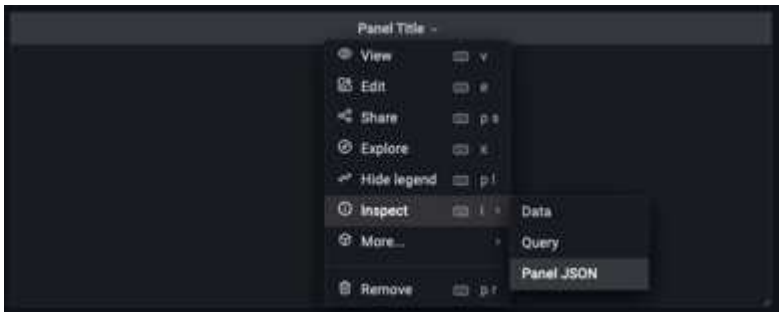
d. Copiez le code JSON et fermez la fenêtre.


e. Dans notre nouveau tableau de bord, cliquez sur l’icône pour ajouter un nouveau panneau.

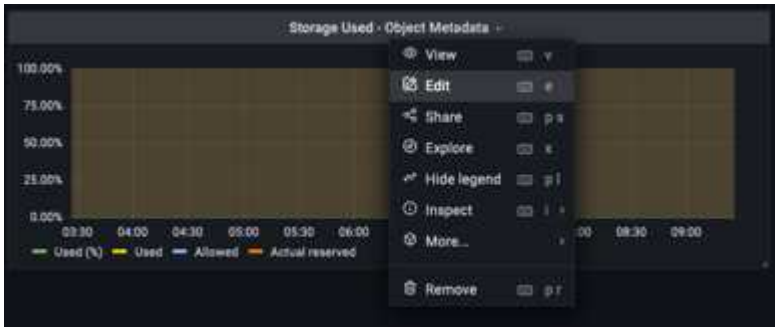


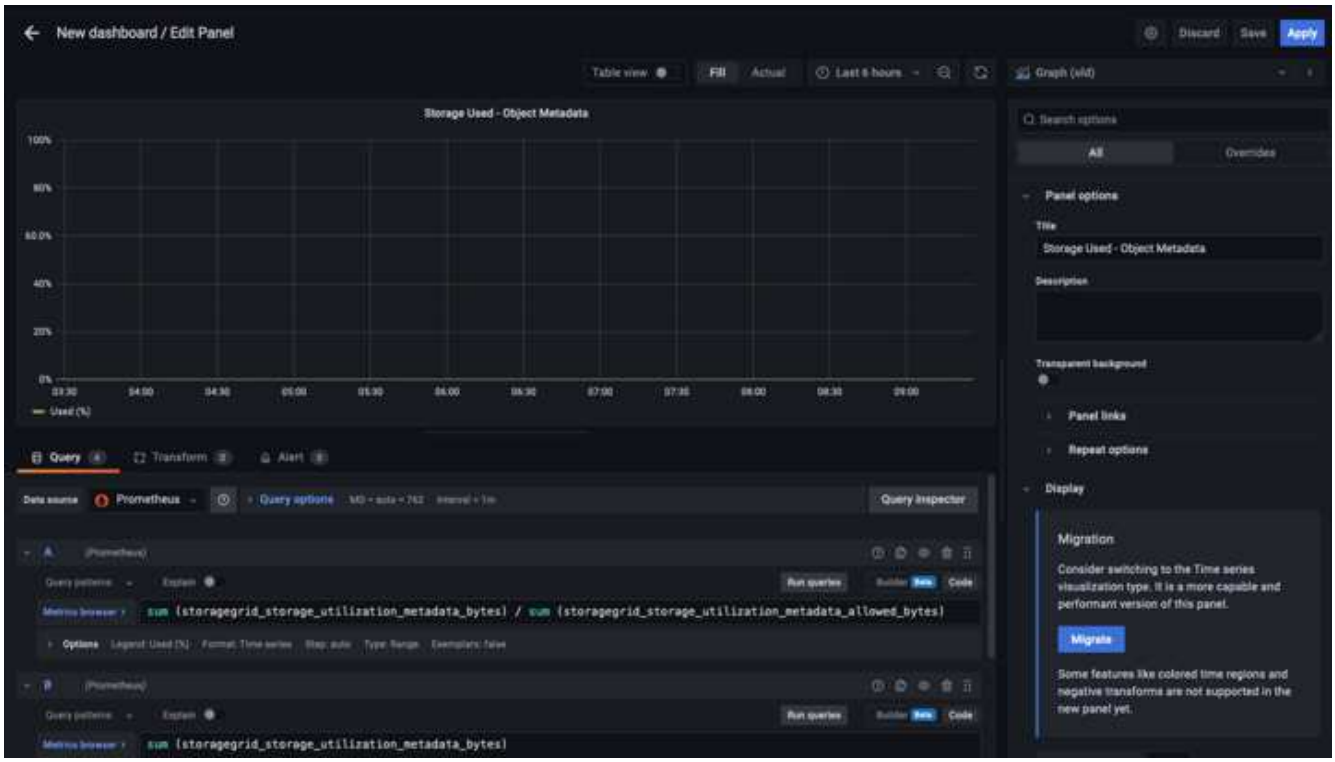

f. Appliquez le nouveau panneau sans apporter de modifications

g. Inspecter le fichier JSON, et tout comme dans le panneau StorageGRID. Supprimez tout code JSON et remplacez-le par le code copié du panneau StorageGRID.




h. Modifiez le nouveau panneau et, à droite, un message migration s'affiche avec un bouton « migrer ». Cliquez sur le bouton, puis sur le bouton « appliquer ».






8. Une fois tous les panneaux en place et configurés comme vous le souhaitez. Enregistrez le tableau de bord en cliquant sur l'icône du disque dans le coin supérieur droit et donnez un nom à votre tableau de bord.

#### Conclusion

Nous disposons désormais d'un serveur Prometheus avec une capacité de stockage et de conservation des données personnalisables. Grâce à cela, nous pouvons continuer à élaborer nos propres tableaux de bord avec les mesures les plus pertinentes pour nos opérations. Vous pouvez obtenir plus d'informations sur les metrics Prometheus collectés dans la "Documentation StorageGRID".

*Par Aron Klein*

# Configuration SNMP Datalog

Configurez Datalog pour collecter les mesures snmp et les traps StorageGRID.

## Configurer Datalog

Datalog est une solution de surveillance qui fournit des mesures, des visualisations et des alertes. La configuration suivante a été implémentée avec l'agent linux version 7.43.1 sur un hôte Ubuntu 22.04.1 déployé localement sur le système StorageGRID.

#### Fichiers de profil Datadog et de déroutement générés à partir du fichier MIB StorageGRID

Datadog fournit une méthode de conversion des fichiers MIB de produit en fichiers de référence Datadog requis pour mapper les messages SNMP.

Ce fichier yaml StorageGRID pour le mappage de résolution des interruptions Datadog généré suivant l'instruction trouvée "ici". + placez ce fichier dans /etc/datadog-agent/conf.d/snmp.d/traps_db/ +

- "Téléchargez le fichier yaml d'interruption" +
  - **somme de contrôle md5** 42e27e4210719945a46172b98c379517 +
  - **sha256 checksum** d0fe5c8e6ca3c902d054f85f8554b70a85f928cba8b7c76391d356f05d2cf73b6887 +

Ce fichier yaml de profil StorageGRID pour le mappage de metrics Datadog généré suivant l'instruction trouvée "ici". + placez ce fichier dans /etc/datadog-agent/conf.d/snmp.d/profiles/ +

- "Téléchargez le fichier yaml de profil" +
  - **somme de contrôle md5** 72bb7784f4801adda4e0c3ea77df19aa +
  - **sha256 checksum** b6b7fadd33063422a8bb8e39b3ead8ab38349ee0229926eadc85f0087b8cee +

### Configuration du datalog SNMP pour les métriques

La configuration de SNMP pour les mesures peut être gérée de deux manières. Vous pouvez configurer la détection automatique en fournissant une plage d'adresses réseau contenant le(s) système(s) StorageGRID ou en définissant les adresses IP des périphériques individuels. L'emplacement de la configuration est différent en fonction de la décision prise. La découverte automatique est définie dans le fichier yaml de l'agent de données. Les définitions explicites de périphériques sont configurées dans le fichier yaml de configuration snmp. Vous trouverez ci-dessous des exemples de chacun d'eux pour le même système StorageGRID.

#### Découverte automatique

configuration située dans /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
```

#### Périphériques individuels

/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'
```

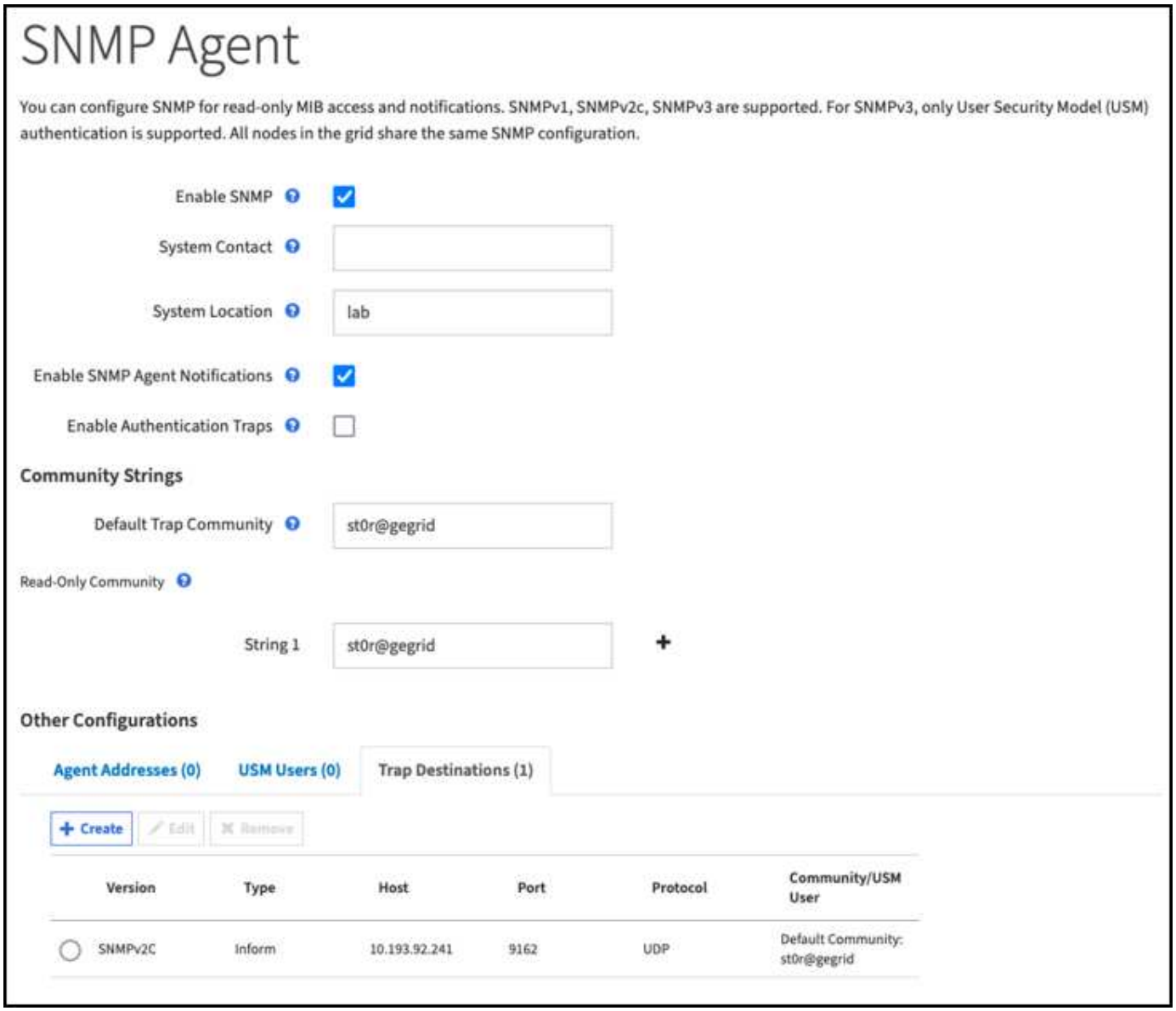
#### Configuration SNMP pour les interruptions

La configuration des traps SNMP est définie dans le fichier de configuration de datadog yaml /etc/datadog-agent/datadog.yaml

```
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
```

#### Exemple de configuration SNMP StorageGRID

L'agent SNMP de votre système StorageGRID se trouve sous l'onglet de configuration, colonne surveillance. Activez SNMP et entrez les informations souhaitées. Si vous souhaitez configurer des interruptions, sélectionnez « destinations des interruptions » et créez une destination pour l'hôte de l'agent Datadog contenant la configuration des interruptions.




*Par Aron Klein*

# Utilisez rclone pour migrer, DÉPLACER et SUPPRIMER des objets sur StorageGRID

Rclone est un outil de ligne de commande et un client gratuits pour les opérations S3. Vous pouvez utiliser rclone pour migrer, copier et supprimer des données d'objet sur StorageGRID. rclone permet de supprimer des compartiments même s'ils ne sont pas vides, grâce à la fonction de « purge » comme illustré ci-dessous.

## Installer et configurer rclone

Pour installer rclone sur un poste de travail ou un serveur, téléchargez-le depuis "rclone.org".

**Étapes de configuration initiale**

1. Créez le fichier de configuration rclone en exécutant le script de configuration ou en créant manuellement le fichier.
2. Dans cet exemple, j'utilise sgdemo pour le nom du terminal StorageGRID S3 distant dans la configuration rclone.
   a. Créez le fichier de configuration ~/.config/rclone/rclone.conf

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

   b. Exécutez la configuration rclone

# 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
```

## Exemples de commandes de base

- **Créer un compartiment :**

  `rclone mkdir remote:bucket`

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

  Utilisez --no-check-certificate si vous devez ignorer les certificats SSL.

- **Liste de tous les compartiments:**

```
rclone lsd remote:
```

```
# rclone lsd sgdemo :
```

- **Liste des objets dans un compartiment spécifique :**

```
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
```

- **Supprimer un compartiment :**

```
rclone rmdir remote:bucket
```

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

- **Mettre un objet:**

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

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

- **Obtenir un objet:**

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

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

- **Supprimer un objet:**

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

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

- **Migrer des objets dans un compartiment**

```
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
```

Utilisez --Progress ou -P pour afficher la progression de la tâche. Sinon, il n'y a pas de sortie.

- **Supprimer un compartiment et tout le contenu de l'objet**

```
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
```

*Par Siegfried Hepp et Aron Klein*

# Bonnes pratiques de déploiement de StorageGRID avec Veeam Backup and Replication

Ce guide se concentre sur la configuration de NetApp StorageGRID et en partie de Veeam Backup and Replication. Ce livre blanc s'adresse aux administrateurs du stockage et du réseau qui connaissent bien les systèmes Linux et sont chargés de la maintenance ou de l'implémentation d'un système NetApp StorageGRID en association avec Veeam Backup and Replication.

## Présentation

Les administrateurs du stockage cherchent à gérer la croissance de leurs données grâce à des solutions qui répondent à leurs besoins en termes de disponibilité, de restauration rapide, d'évolutivité et d'automatisation des règles de conservation des données à long terme. Ces solutions doivent également offrir une protection contre les pertes et les attaques malveillantes. Ensemble, Veeam et NetApp ont créé une solution de protection des données combinant Veeam Backup & Recovery avec NetApp StorageGRID pour le stockage objet sur site.

Veeam et NetApp StorageGRID proposent une solution simple d'utilisation qui s'associent pour répondre aux exigences liées à la croissance rapide des données et à l'augmentation des réglementations à travers le monde. Le stockage objet basé dans le cloud est réputé pour sa résilience, son évolutivité, ses fonctionnalités opérationnelles et sa rentabilité, qui en font un choix naturel comme cible pour vos sauvegardes. Ce document fournit des conseils et des recommandations pour la configuration de votre solution de sauvegarde Veeam et de votre système StorageGRID.

La charge de travail d'objets de Veeam crée un grand nombre d'opérations simultanées de PUT, DELETE et LIST pour les petits objets. L'activation de l'immuabilité ajoute au nombre de demandes dans le magasin d'objets pour définir la conservation et répertorier les versions. Le processus d'une tâche de sauvegarde comprend l'écriture d'objets pour la modification quotidienne. Une fois les nouvelles écritures terminées, la tâche supprime tous les objets basés sur la stratégie de rétention de la sauvegarde. La planification des tâches de sauvegarde se chevauchera presque toujours. Ce chevauchement entraînera une grande partie de la fenêtre de sauvegarde comprenant une charge de travail PUT/DELETE 50/50 sur le magasin d'objets. Ajuster dans Veeam le nombre d'opérations simultanées avec le paramètre de slot de tâche, augmenter la

taille de l'objet en augmentant la taille du bloc de tâche de sauvegarde, réduire le nombre d'objets dans les demandes de suppression multi-objets, et le choix de la fenêtre de temps maximum pour les tâches à effectuer optimisera la solution en termes de performances et de coûts.

Assurez-vous de lire la documentation du produit pour "Sauvegarde et réplication Veeam" et "StorageGRID" avant de commencer. Veeam propose des calculateurs de compréhension du dimensionnement de l'infrastructure Veeam et des exigences de capacité à utiliser avant de dimensionner votre solution StorageGRID. Veuillez toujours consulter les configurations Veeam-NetApp validées sur le site web du programme Veeam Ready pour "Objets compatibles Veeam, immuabilité d'objet et référentiel".

## Configuration Veeam

### Version recommandée

Il est toujours recommandé de rester à jour et d'appliquer les derniers correctifs pour votre système Veeam Backup & Replication 12. Nous recommandons actuellement d'installer au moins le correctif Veeam P20230718.

### Configuration du référentiel S3

Un référentiel de sauvegarde scale-out (SOBR) est le Tier de capacité du stockage objet S3. Le Tier de capacité est une extension du référentiel principal, qui permet de prolonger les périodes de conservation des données et de réduire le coût de la solution de stockage. Veeam a la possibilité d'immuabilité avec l'API S3 Object Lock. Veeam 12 peut utiliser plusieurs compartiments dans un référentiel scale-out. StorageGRID n'a pas de limite pour le nombre d'objets ou la capacité d'un compartiment unique. L'utilisation de plusieurs compartiments peut améliorer les performances lors de la sauvegarde de datasets très volumineux où les données de sauvegarde peuvent atteindre plusieurs pétaoctets dans des objets.

La limitation des tâches simultanées peut être nécessaire en fonction du dimensionnement de la solution et des besoins spécifiques. Les paramètres par défaut spécifient un emplacement de tâche de référentiel pour chaque cœur de processeur et pour chaque emplacement de tâche une limite d'emplacement de tâche simultanée de 64. Par exemple, si votre serveur dispose de 2 cœurs de processeur, 128 threads simultanés au total seront utilisés pour le magasin d'objets. Cela inclut les COMMANDES PUT, GET et batch Delete. Il est recommandé de sélectionner une limite conservatrice pour les créneaux de tâches à commencer par et d'ajuster cette valeur une fois que les sauvegardes Veeam ont atteint l'état stable de nouvelles sauvegardes et que les données de sauvegarde expirent. Veuillez vous adresser à votre équipe de gestion de compte NetApp pour dimensionner le système StorageGRID en fonction des délais et des performances souhaités. Il peut être nécessaire de régler le nombre d'emplacements de tâches et la limite des tâches par emplacement pour obtenir la solution optimale.

### Configuration de la procédure de sauvegarde

Les tâches de sauvegarde Veeam peuvent être configurées avec plusieurs options de taille de bloc qui doivent être prises en compte avec attention. La taille de bloc par défaut est de 1 Mo. Grâce à l'efficacité du stockage, Veeam assure la compression et la déduplication, ce qui permet de créer des tailles d'objet d'environ 500 Ko pour la sauvegarde complète initiale et des objets de 100 à 200 Ko pour les tâches incrémentielles. Nous pouvons considérablement améliorer les performances et réduire les besoins en matière de magasin d'objets en choisissant une taille de bloc de sauvegarde plus importante. Si la taille de bloc supérieure améliore considérablement les performances du magasin d'objets, elle implique toutefois une augmentation potentielle des besoins en capacité de stockage primaire en raison de la réduction des performances du stockage. Il est recommandé de configurer les tâches de sauvegarde avec une taille de bloc de 4 Mo, ce qui crée des objets d'environ 2 Mo pour les sauvegardes complètes et des objets de 700 Ko à 1 Mo pour les sauvegardes incrémentielles. Les clients peuvent même envisager de configurer des tâches de sauvegarde à l'aide d'une taille de bloc de 8 Mo, qui peut être activée avec l'aide du support Veeam.

La mise en œuvre des sauvegardes immuables utilise le verrouillage objet S3 dans le magasin d'objets. L'option immuabilité génère un nombre accru de requêtes auprès du magasin d'objets pour obtenir des mises à jour de listes et de conservation des objets.

Lorsque les rétentions de sauvegarde expirent, les procédures de sauvegarde traitent la suppression des objets. Veeam envoie les demandes de suppression au magasin d'objets dans le cadre de requêtes de suppression de plusieurs objets de 1000 objets par demande. Pour les petites solutions, il peut être nécessaire de l'ajuster afin de réduire le nombre d'objets par demande. En outre, si cette valeur est moindre, les demandes de suppression seront réparties de manière plus homogène entre les nœuds du système StorageGRID. Il est recommandé d'utiliser les valeurs du tableau ci-dessous comme point de départ pour la configuration de la limite de suppression de plusieurs objets. Multipliez la valeur du tableau par le nombre de nœuds pour le type d'appliance choisi pour obtenir la valeur du paramètre dans Veeam. Si cette valeur est égale ou supérieure à 1000, il n'est pas nécessaire d'ajuster la valeur par défaut. Si cette valeur doit être ajustée, contactez le support Veeam pour effectuer cette modification.

| Modèle de type appliance | S3MultiObjectDeleteLimit par nœud |
|---|---|
| SG5712 | 34 |
| SG5760 | 75 |
| SG6060 | 200 |

Pour en savoir plus sur la configuration recommandée en fonction de vos besoins, contactez l'équipe NetApp en charge de votre compte. Les recommandations concernant les paramètres de configuration Veeam incluent :

- Taille du bloc de la tâche de sauvegarde = 4 Mo
- Limite d'emplacement de tâche SOBR = 2-16
- Limite de suppression de plusieurs objets = 34-1000

## Configuration StorageGRID

### Version recommandée

NetApp StorageGRID 11.6 ou 11.7 avec le dernier correctif est la version recommandée pour les déploiements Veeam. De nombreuses fonctionnalités d'optimisation ont été introduites dans StorageGRID 11.6.0.11 et 11.7.0.4, qui seront bénéfiques pour les charges de travail Veeam. Il est toujours recommandé de rester à jour et d'appliquer les derniers correctifs pour votre système StorageGRID.

### Configuration de l'équilibreur de charge et du terminal S3

Dans Veeam, le terminal doit être connecté via HTTPS uniquement. Veeam ne prend pas en charge les connexions non chiffrées. Le certificat SSL peut être un certificat auto-signé, une autorité de certification privée de confiance ou une autorité de certification publique de confiance. Pour assurer un accès continu au référentiel S3, il est recommandé d'utiliser au moins deux équilibreurs de charge dans une configuration haute disponibilité. Les équilibreurs de charge peuvent être un service d'équilibrage de charge intégré fourni par StorageGRID, situé sur chaque nœud d'administration et nœud de passerelle ou sur une solution tierce telle que F5, Kemp, HASProxy, Loadbalanacer.org, etc L'utilisation d'un équilibreur de charge StorageGRID permet de définir des classificateurs du trafic (règles de QoS) capables de hiérarchiser le workload Veeam ou de limiter Veeam à ne pas affecter les workloads prioritaires sur le système StorageGRID.

### Compartiment S3

StorageGRID est un système de stockage mutualisé sécurisé. Il est recommandé de créer un locataire dédié à la charge de travail Veeam. Un quota de stockage peut être attribué en option. Comme bonne pratique, activez « utiliser son propre référentiel d'identité ». Sécurisez l'utilisateur root management du locataire avec un mot de passe approprié. Veeam Backup 12 nécessite une cohérence renforcée pour les compartiments S3. StorageGRID propose plusieurs options de cohérence configurées au niveau du compartiment. Pour les déploiements multi-sites avec Veeam accédant aux données depuis plusieurs sites, sélectionnez « strong-global ». Si les sauvegardes et les restaurations Veeam ont lieu sur un seul site, le niveau de cohérence doit être défini sur « site à forte intensité ». Pour plus d'informations sur les niveaux de cohérence des compartiments, consultez le "documentation". Pour utiliser les sauvegardes StorageGRID contre les immuabilité, S3 Object Lock doit être activé globalement et configuré sur le compartiment lors de la création du compartiment.

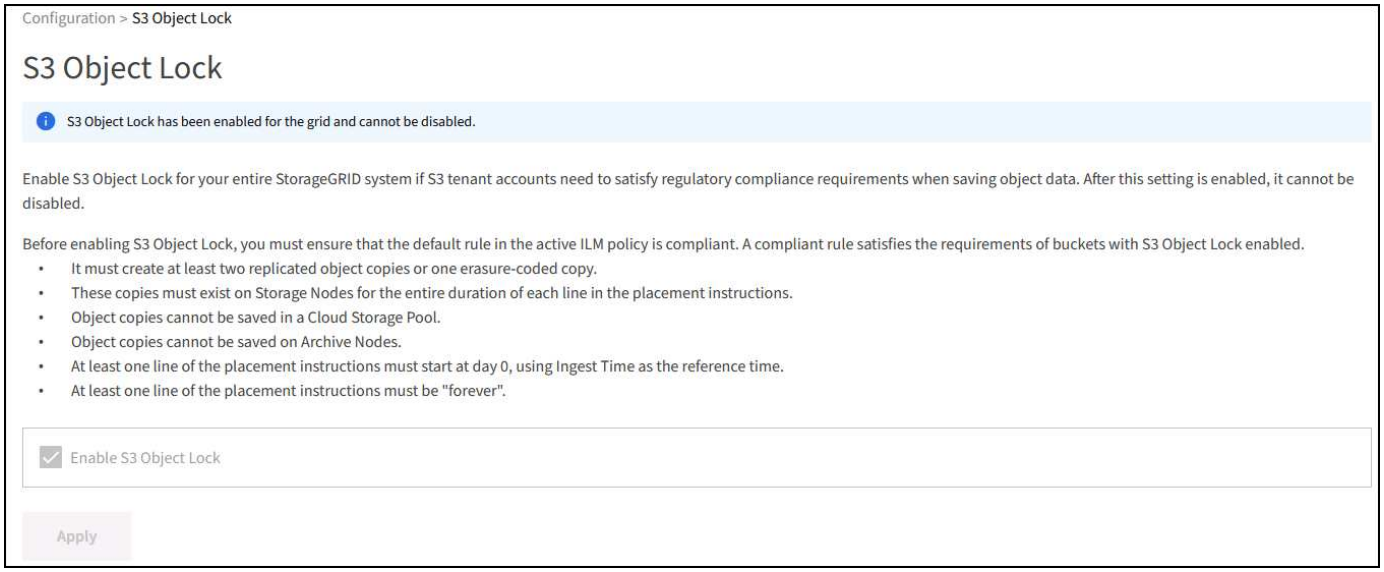
### Gestion du cycle de vie

StorageGRID prend en charge la réplication et le code d'effacement pour la protection au niveau objet sur l'ensemble des nœuds et sites StorageGRID. Le codage d'effacement requiert une taille d'objet d'au moins 200 Ko. La taille de bloc par défaut de Veeam de 1 Mo produit des tailles d'objet qui peuvent souvent être inférieures à cette taille minimale recommandée de 200 Ko après les fonctionnalités d'efficacité du stockage de Veeam. Pour les performances de la solution, il est déconseillé d'utiliser un profil de code d'effacement sur plusieurs sites, sauf si la connectivité entre les sites suffit pour ne pas augmenter la latence ou restreindre la bande passante du système StorageGRID. Dans un système StorageGRID multisite, la règle ILM peut être configurée pour stocker une copie unique sur chaque site. Pour une durabilité ultime, une règle pourrait être configurée de manière à stocker une copie codée en effacement sur chaque site. L'implémentation la plus recommandée pour cette charge de travail est l'utilisation de deux copies en local sur les serveurs Veeam Backup.

## Points clés de la mise en œuvre

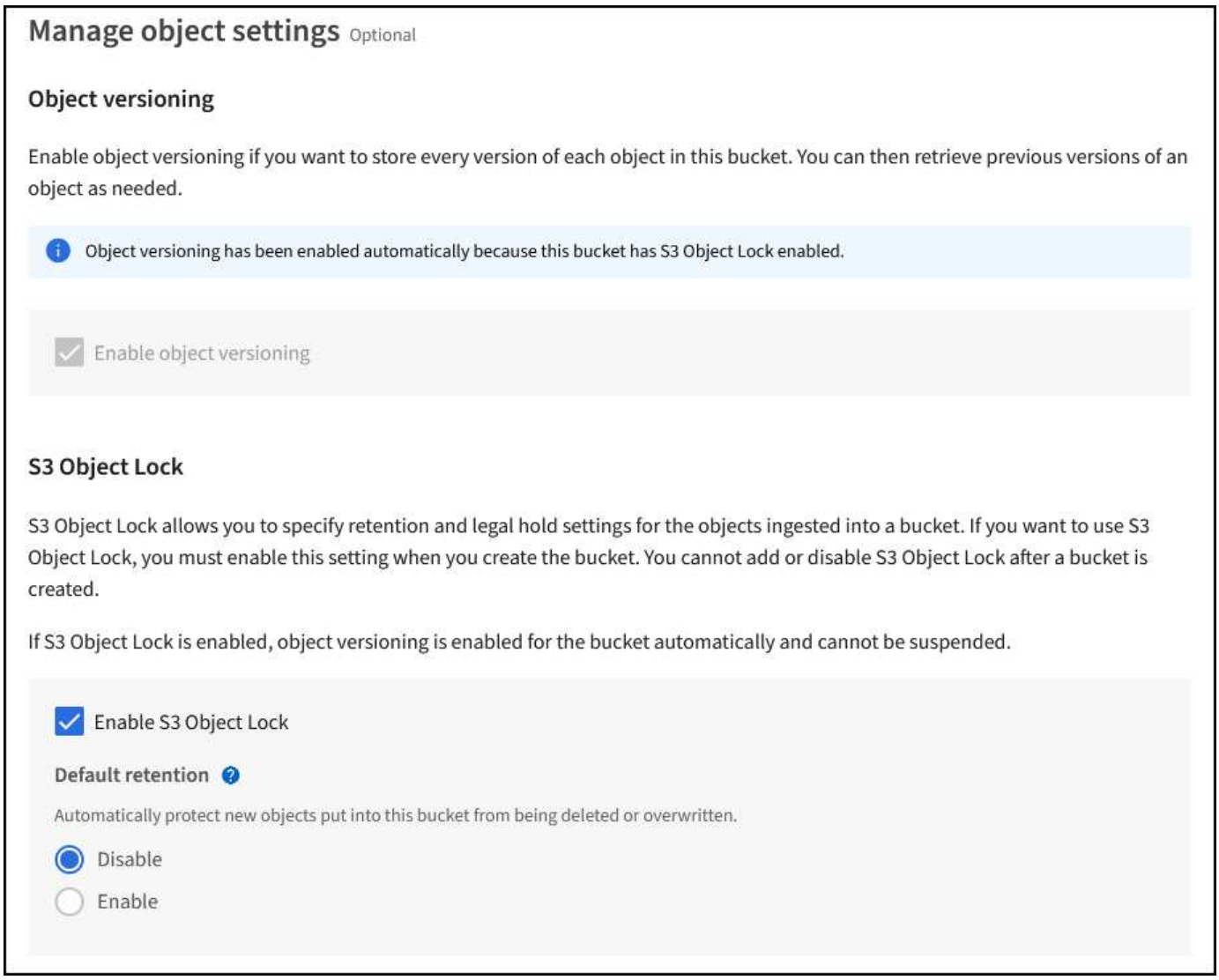
### StorageGRID

Assurez-vous que le verrouillage des objets est activé sur le système StorageGRID si l'immuabilité est requise. Recherchez l'option dans l'interface de gestion sous Configuration/S3 Object Lock.

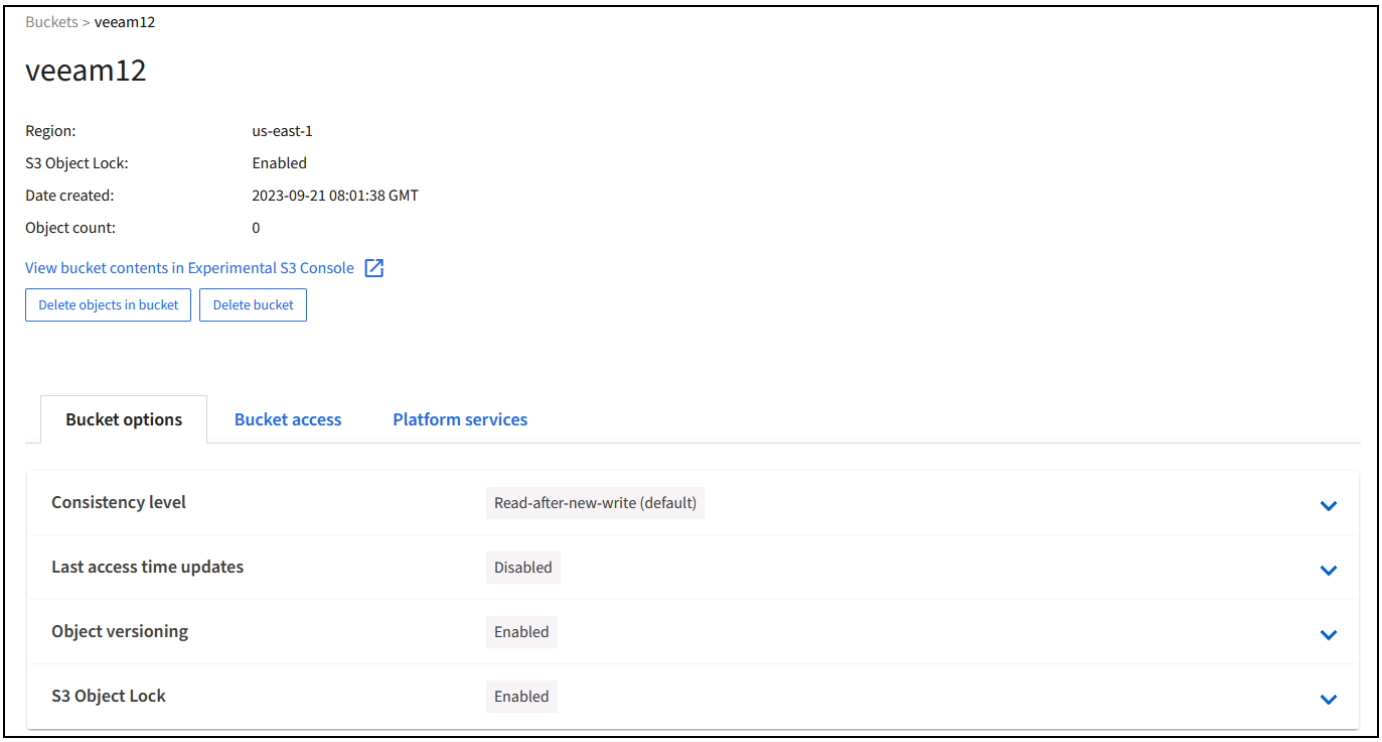



Lors de la création du compartiment, sélectionnez Activer le verrouillage des objets S3 si ce compartiment doit être utilisé pour les sauvegardes sans altération. La gestion des versions de compartiment est alors

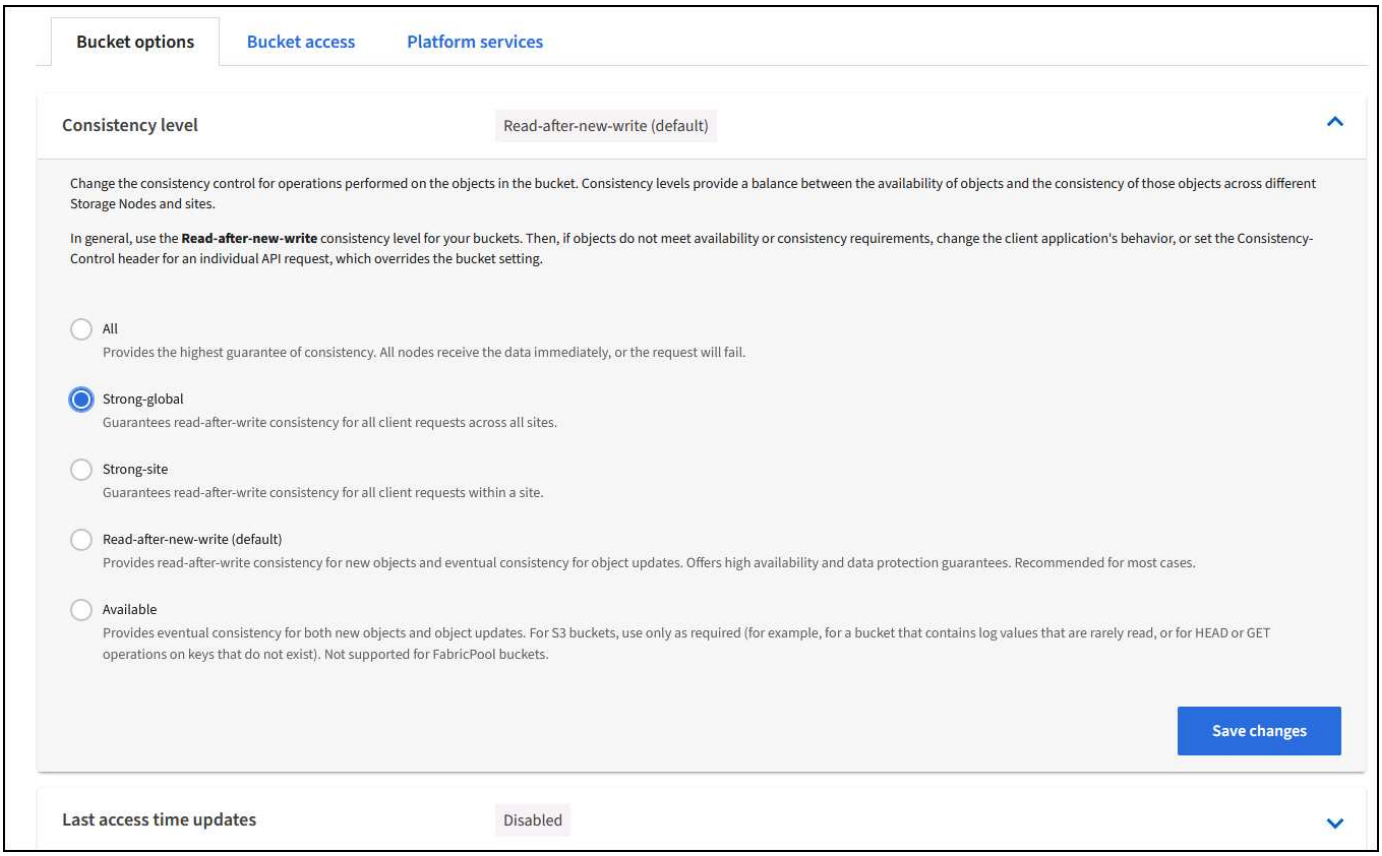
automatiquement activée. Laissez la conservation par défaut désactivée, car Veeam définit la conservation d'objet de manière explicite. La gestion des versions et le verrouillage objet S3 ne doivent pas être sélectionnés si Veeam ne crée pas de sauvegardes immuables.




Une fois le compartiment créé, accédez à la page de détails du compartiment créé. Sélectionnez le niveau de cohérence.



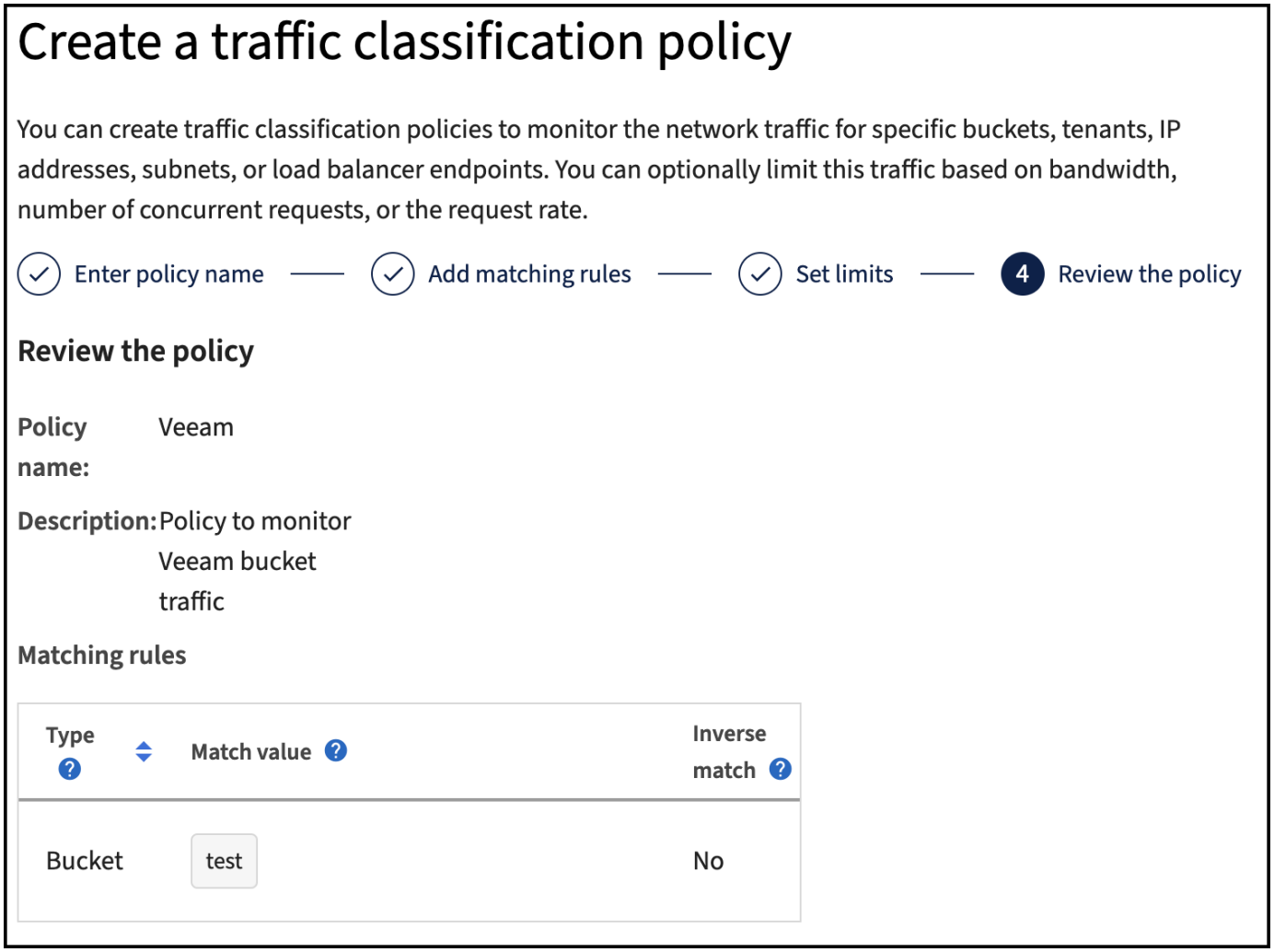

Veeam requiert une cohérence renforcée pour les compartiments S3. Pour les déploiements multi-sites avec Veeam qui accèdent aux données depuis plusieurs sites, sélectionnez « strong-global ». Si les sauvegardes et les restaurations Veeam ont lieu sur un seul site, le niveau de cohérence doit être défini sur « site à forte intensité ». Enregistrez les modifications.




StorageGRID propose un service d’équilibrage de la charge intégré sur chaque nœud d’administration et sur

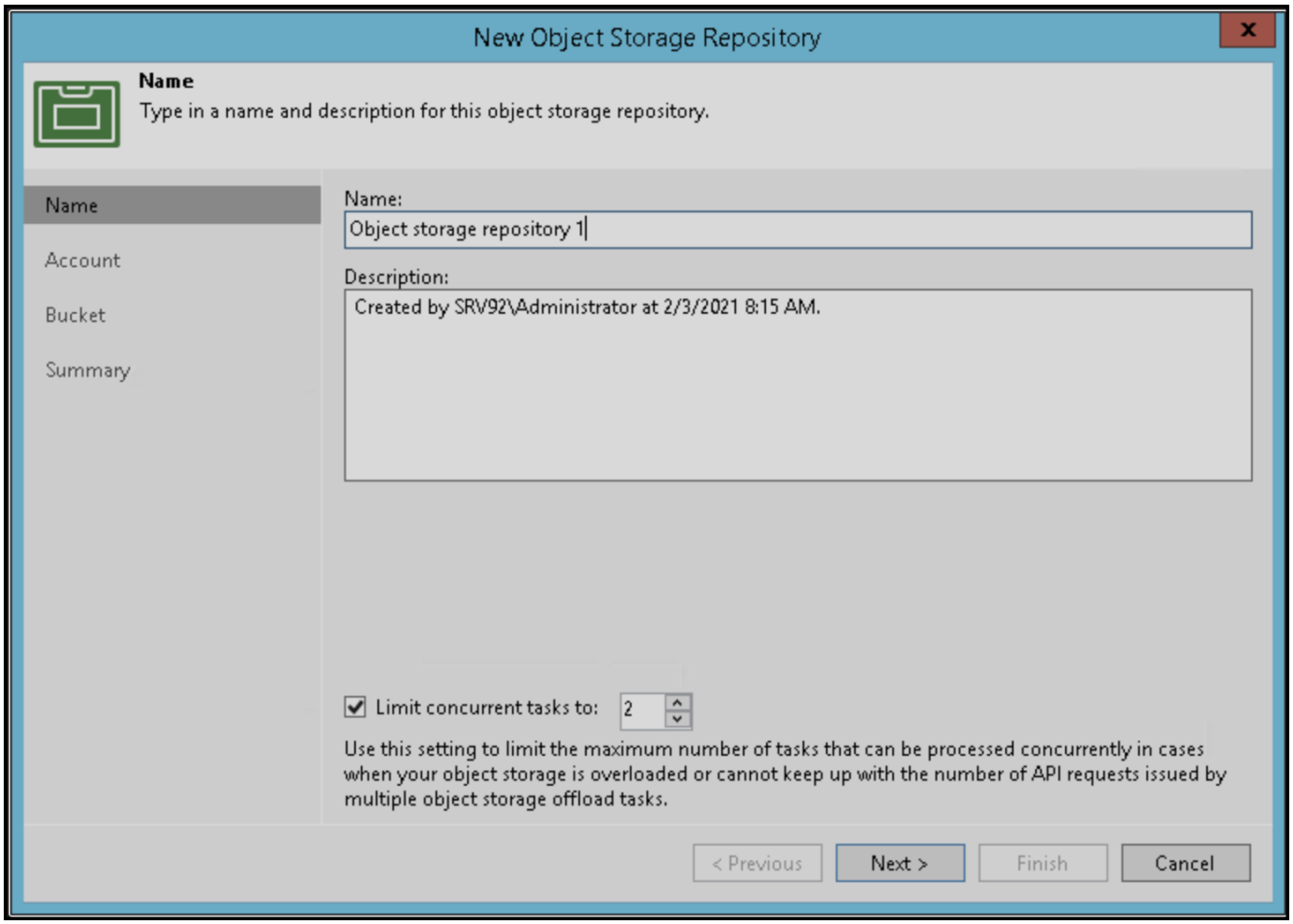
tous les nœuds de passerelle dédiés. L'un des nombreux avantages de l'utilisation de cet équilibreur de charge est la possibilité de configurer des règles de classification du trafic (QoS). Bien qu'elles soient principalement utilisées pour limiter l'impact des applications sur les autres charges de travail client ou pour hiérarchiser une charge de travail sur d'autres, elles fournissent également un bonus de collecte de metrics supplémentaires pour faciliter le contrôle.

Dans l'onglet de configuration, sélectionnez "classification du trafic" et créez une nouvelle stratégie. Attribuez un nom à la règle et sélectionnez le ou les compartiments ou le tenant comme type. Entrez le(s) nom(s) du ou des compartiments ou du tenant. Si la qualité de service est requise, définissez une limite, mais pour la plupart des implémentations, il convient d'ajouter les avantages en termes de surveillance, afin de ne pas fixer de limite.

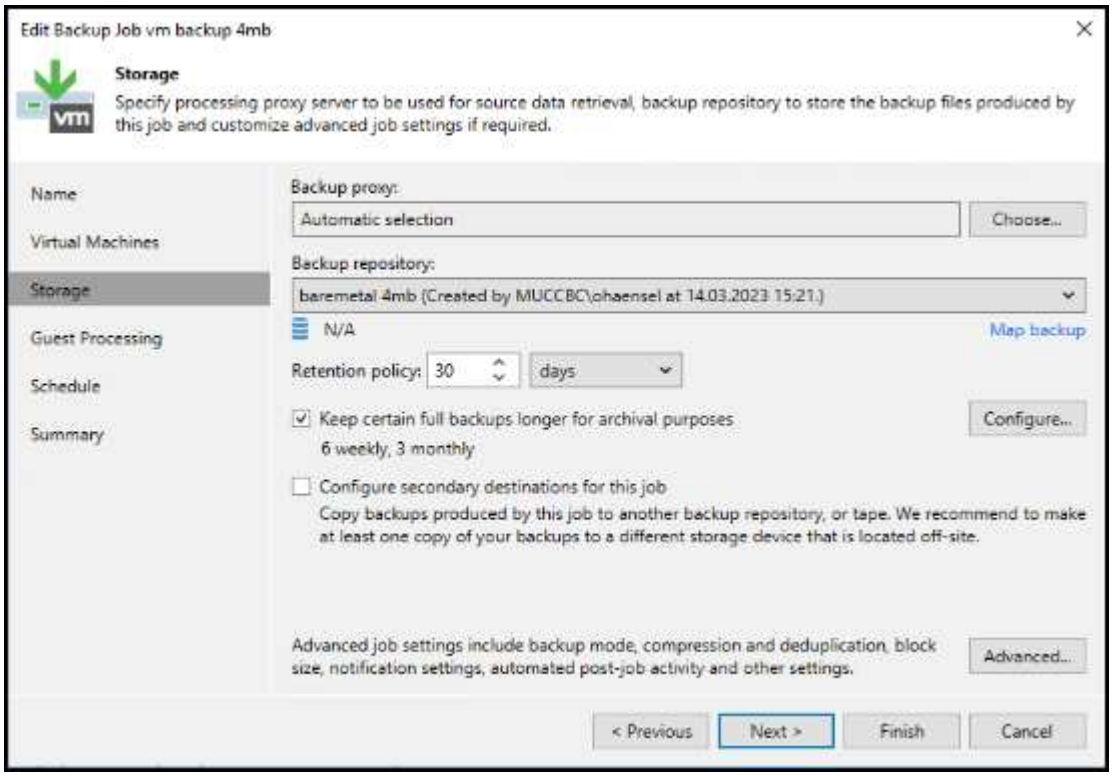



### Veeam

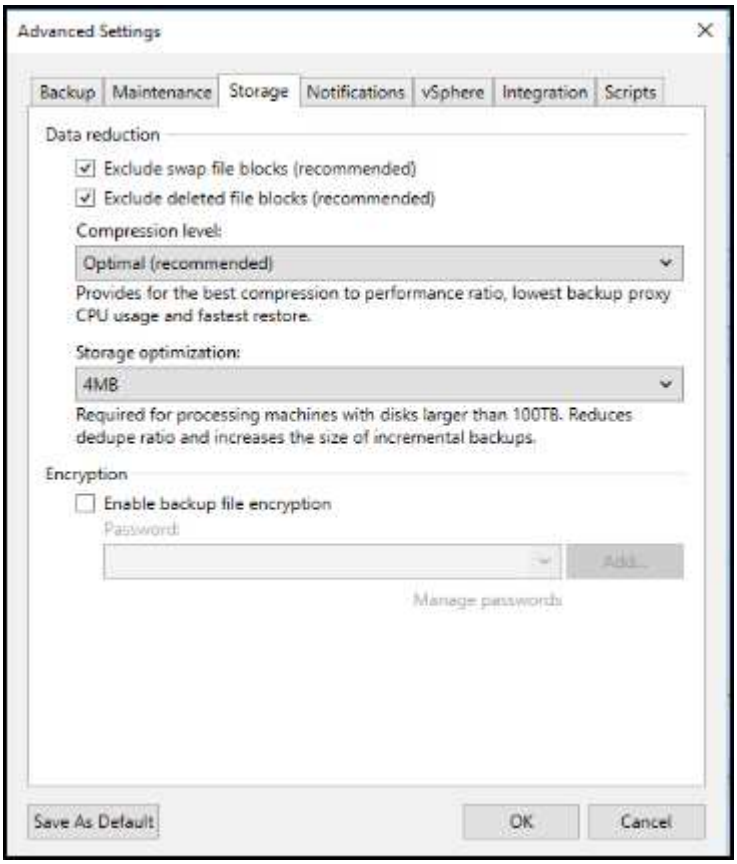
Selon le modèle et la quantité d'appliances StorageGRID, il peut être nécessaire de sélectionner et de configurer une limite au nombre d'opérations simultanées sur le compartiment.




Pour démarrer l'assistant, suivez la documentation Veeam sur la configuration des tâches de sauvegarde dans la console Veeam. Après avoir ajouté des machines virtuelles, sélectionnez le référentiel SOBR.



Cliquez sur Paramètres avancés et définissez les paramètres d'optimisation du stockage sur 4 Mo ou plus. La compression et la déduplication doivent être activées. Modifiez les paramètres invités en fonction de vos besoins et configurez la planification des tâches de sauvegarde.




## Surveillance StorageGRID

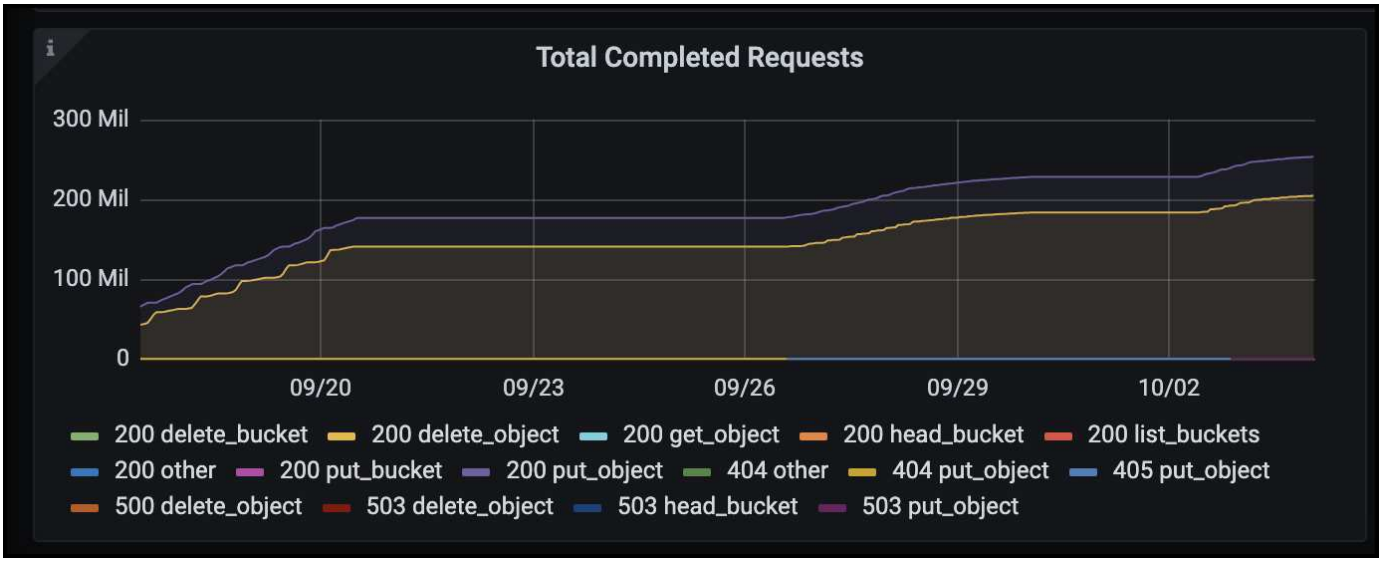
Pour obtenir une vue d'ensemble des performances de Veeam et StorageGRID, vous devez attendre l'expiration du délai de conservation des premières sauvegardes. Jusqu'à présent, la charge de travail Veeam se compose principalement d'opérations PUT et aucune suppression n'a eu lieu. Une fois que les données de sauvegarde arrivent à expiration et que les nettoyages sont en cours, vous pouvez voir l'utilisation cohérente complète du magasin d'objets et ajuster les paramètres dans Veeam, si nécessaire.

StorageGRID fournit des graphiques pratiques pour contrôler le fonctionnement du système, disponibles dans l'onglet support, page Metrics. Les principaux tableaux de bord à examiner seront la vue d'ensemble S3, ILM et la règle de classification du trafic si une règle a été créée. Vous trouverez dans le tableau de bord S3 des informations sur les taux d'opération S3, les latences et les réponses aux demandes.

Les taux S3 et les requêtes actives vous permettent de voir la charge que chaque nœud gère et le nombre total de requêtes par type.


Le graphique durée moyenne indique la durée moyenne de chaque nœud pour chaque type de demande. Il s'agit de la latence moyenne de la demande et peut être un bon indicateur qu'un réglage supplémentaire peut être nécessaire ou que le système StorageGRID peut prendre plus de charge.


Dans le tableau nombre total de demandes terminées, vous pouvez voir les demandes par type et par code de réponse. Si vous voyez des réponses autres que 200 (OK), cela peut indiquer un problème comme le système StorageGRID est fortement chargé et envoie 503 réponses (ralentissement) et un réglage supplémentaire peut être nécessaire, ou le temps est venu d'étendre le système pour augmenter la charge.




Le tableau de bord ILM vous permet de contrôler les performances de suppression de votre système StorageGRID. StorageGRID combine les suppressions synchrones et asynchrones sur chaque nœud afin d'essayer d'optimiser la performance globale de toutes les requêtes.


Dans le cadre d'une règle de classification du trafic, nous pouvons afficher des metrics sur le débit de la demande d'équilibrage de charge, les taux, la durée, ainsi que la taille des objets envoyés et reçus par Veeam.


## Où trouver des informations complémentaires

Pour en savoir plus sur les informations données dans ce livre blanc, consultez ces documents et/ou sites web :

- "Documentation du produit NetApp StorageGRID 11.7"
- "Sauvegarde et réplication Veeam"

*Par Oliver Haensel et Aron Klein*

# Configurez la source de données Dremio avec StorageGRID

Dremio prend en charge la rareté des sources de données, y compris le stockage objet dans le cloud ou sur site. Vous pouvez configurer Dremio pour qu'il utilise StorageGRID comme source de données de stockage objet.

## Configurer la source de données Dremio

### Prérequis

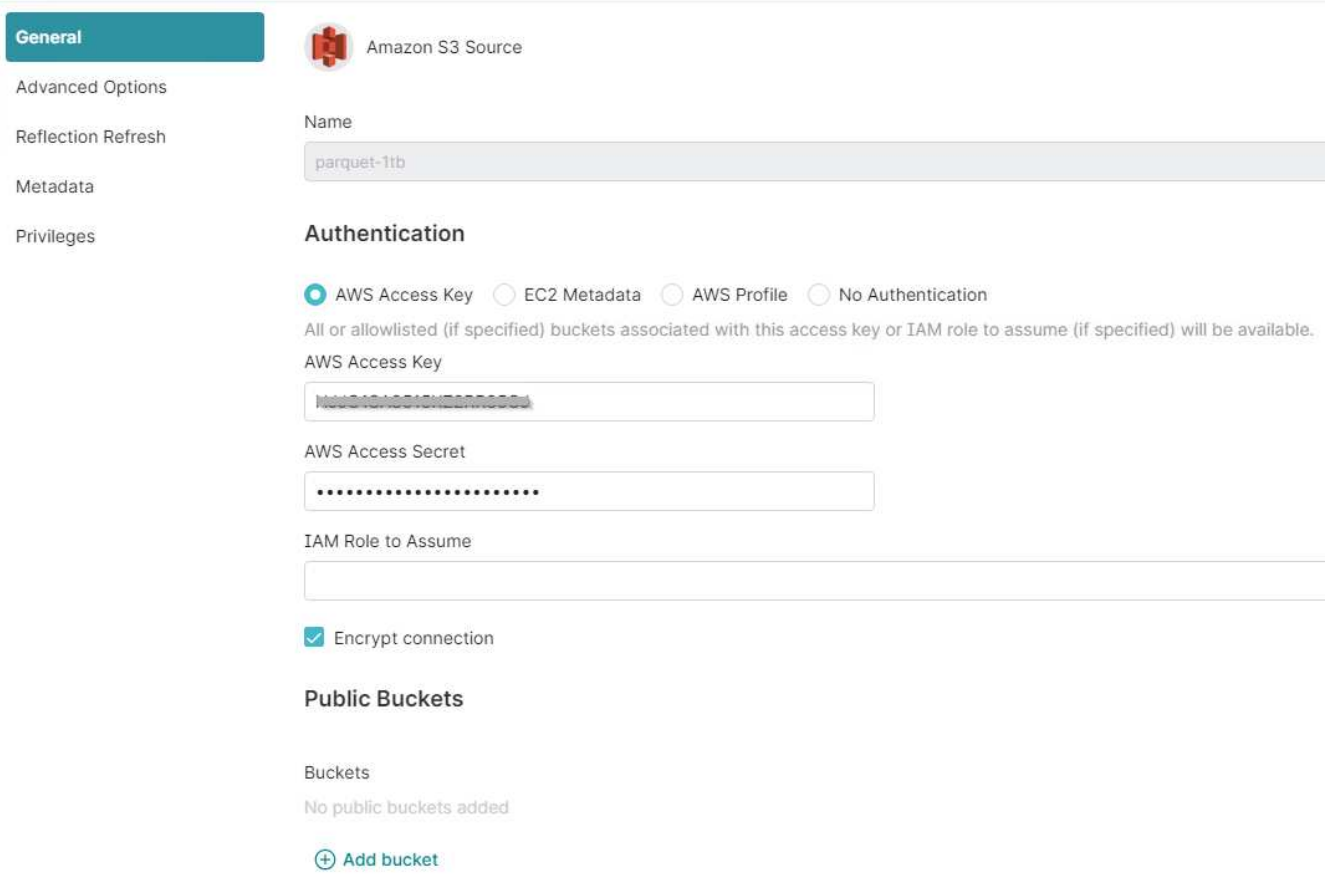
- URL de terminal StorageGRID S3, ID de clé d'accès s3 du locataire et clé d'accès secrète.
- Recommandation de configuration StorageGRID : désactivez la compression (désactivée par défaut). Dremio utilise la plage d'octets GET pour extraire simultanément différentes plages d'octets à partir du même objet pendant la requête. La taille type des demandes de plage d'octets est de 1 Mo. Les objets compressés dégradent les performances GET au niveau de la plage d'octets.

### Guide Dremio

"Connexion à Amazon S3 : Configuration du stockage compatible avec S3".

## Instructions

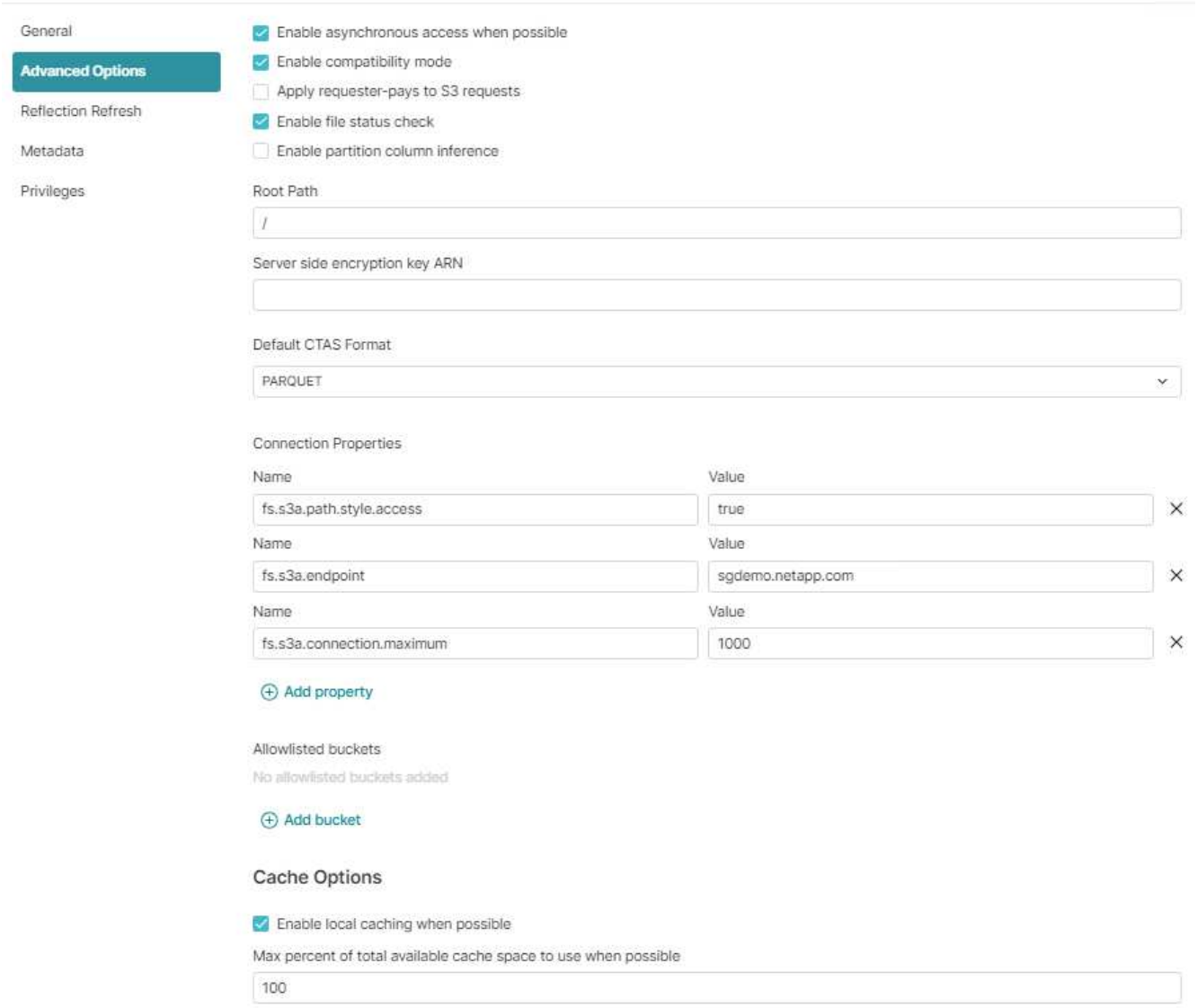
1. Sur la page Datasets Dremio, cliquez sur le signe + pour ajouter une source, sélectionnez 'Amazon S3'.
2. Entrez le nom de cette nouvelle source de données : ID de clé d'accès du locataire StorageGRID S3 et clé d'accès secrète.
3. Cochez la case « crypter la connexion » si vous utilisez https pour la connexion au terminal StorageGRID S3.
   Si vous utilisez un certificat CA auto-signé pour ce noeud final s3, suivez l'instructions du guide Dremio pour ajouter ce certificat CA dans <JAVA_HOME>/jre/lib/Security + du serveur Dremio
   **Exemple de capture d'écran**



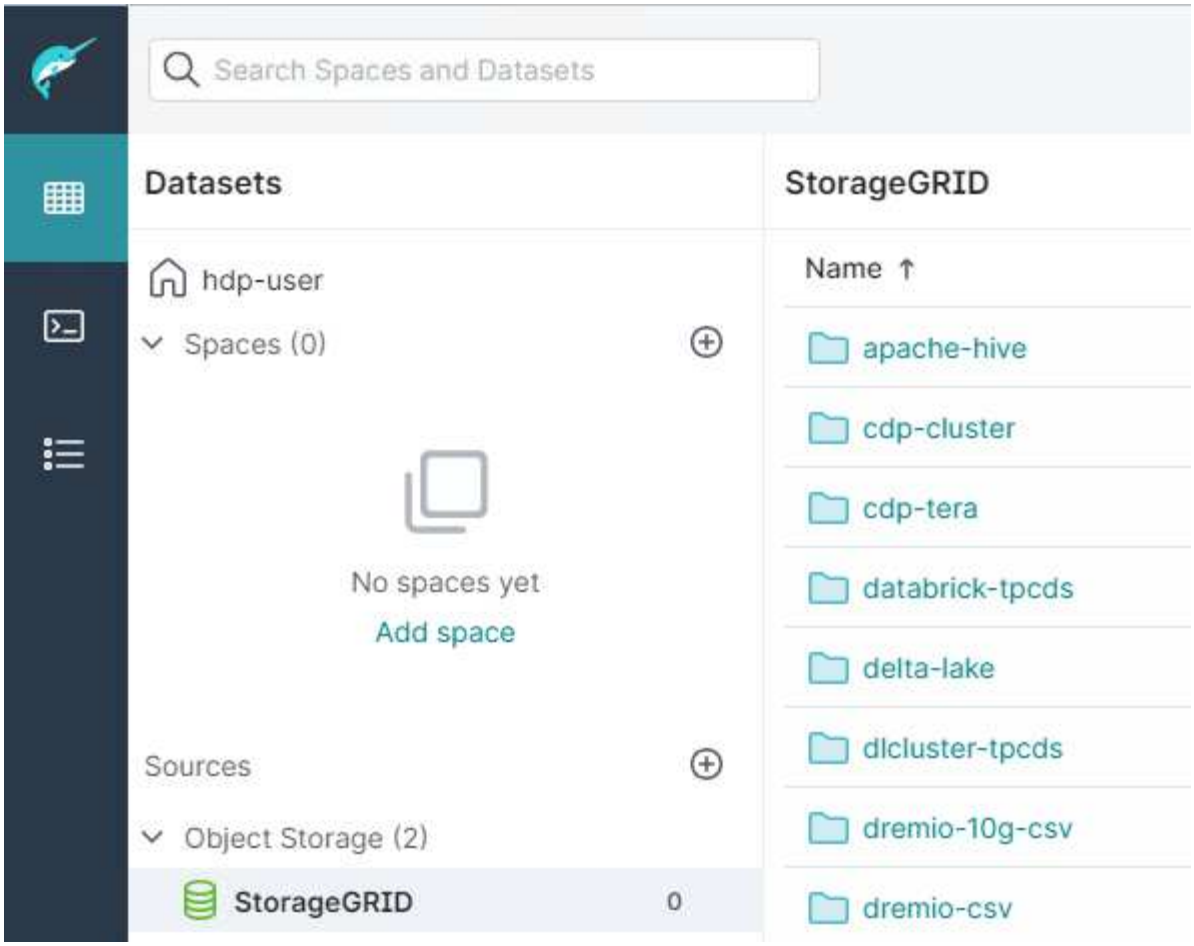

4. Cliquez sur « Options avancées », cochez « Activer le mode de compatibilité ».
5. Sous Propriétés de connexion, cliquez sur + Ajouter des propriétés et ajoutez ces propriétés s3a.
6. fs.s3a.connection.la valeur par défaut maximale est 100. Si vos datasets s3 incluent des fichiers de parquet volumineux comportant au moins 100 colonnes, vous devez entrer une valeur supérieure à 100. Reportez-vous au guide Dremio pour ce réglage.

| Nom | Valeur |
| --- | --- |
| fs.s3a.endpoint | *<noeud final StorageGRID S3:port>* |
| fs.s3a.path.style.access | vrai |
| fs.s3a.connexion.maximum | *<une valeur supérieure à 100>* |

**Exemple de capture d'écran**




7. Configurez les autres options de Dremio en fonction des besoins de votre organisation ou de vos applications.
8. Cliquez sur le bouton Enregistrer pour créer cette nouvelle source de données.
9. Une fois la source de données StorageGRID ajoutée, une liste de rubriques s'affiche dans le panneau de gauche.
   **Exemple de capture d'écran**




*Par Angela Cheng*

# NetApp StorageGRID avec GitLab

NetApp a testé StorageGRID avec GitLab. Voir l'exemple de configuration GitLab ci-dessous. Reportez-vous à la section "Guide de configuration du stockage objet GitLab" pour plus d'informations.

## Exemple de connexion de stockage objet

Pour les installations de package Linux, voici un exemple de `connection` configuration dans le formulaire consolidé. Modifier `/etc/gitlab/gitlab.rb` et ajoutez les lignes suivantes en remplaçant les valeurs souhaitées :

```
# 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'
```

# Procédures et exemples d'API

## Tester et démontrer les options de cryptage S3 sur StorageGRID

StorageGRID et l'API S3 proposent plusieurs façons de chiffrer vos données au repos. Pour en savoir plus, voir "Étudiez les méthodes de cryptage StorageGRID".

Ce guide présente les méthodes de chiffrement de l'API S3.

### Chiffrement côté serveur (SSE)

SSE permet au client de stocker un objet et de le chiffrer à l'aide d'une clé unique gérée par StorageGRID. Lorsque l'objet est demandé, il est décrypté par la clé stockée dans StorageGRID.

#### Exemple SSE

- PLACER un objet avec SSE

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

- DIRIGEZ l'objet pour vérifier le chiffrement

```
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": {}
}
```

- OBTENIR l'objet

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

## Chiffrement côté serveur avec clés fournies par le client (SSE-C)

SSE permet au client de stocker un objet et de le chiffrer à l'aide d'une clé unique fournie par le client avec l'objet. Lorsque l'objet est demandé, la même clé doit être fournie pour décrypter et renvoyer l'objet.

**Exemple SSE-C.**

- Vous pouvez créer une clé de chiffrement à des fins de test ou de démonstration
  - Créez une clé de chiffrement

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

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

- Placer un objet avec la clé générée

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

- Dirigez l'objet

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

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

Si vous ne fournissez pas la clé de cryptage, vous recevrez une erreur « une erreur s'est produite (404) lors de l'appel de l'opération HeadObject : introuvable ».

- Obtenir l'objet

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

> Si vous ne fournissez pas la clé de cryptage, vous recevrez une erreur "une erreur s'est produite (InvalidRequest) lors de l'appel de l'opération GetObject: L'objet a été stocké à l'aide d'une forme de chiffrement côté serveur. Les paramètres corrects doivent être fournis pour récupérer l'objet. »

## Chiffrement côté serveur godet (SSE-S3)

SSE-S3 permet au client de définir un comportement de cryptage par défaut pour tous les objets stockés dans un compartiment. Les objets sont chiffrés avec une clé unique gérée par StorageGRID. À la demande de l'objet, celui-ci est décrypté par la clé stockée dans StorageGRID.

### Exemple de godet SSE-S3

- Créez un compartiment et définissez une règle de chiffrement par défaut
  - Créer un nouveau compartiment

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

  - Put bucket Encryption

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

- Placer un objet dans le godet

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

- Dirigez l'objet

```
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": {}
}
```

- OBTENIR l'objet

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

*Par Aron Klein*

# Testez et faites une démonstration du verrouillage d'objet S3 sur StorageGRID

Le verrouillage d'objet fournit un modèle WORM pour éviter que les objets ne soient supprimés ou remplacés. L'implémentation StorageGRID du verrouillage d'objet est une fonctionnalité qui est évaluée afin de respecter les exigences réglementaires, et qui prend en charge le mode de conservation légale et de conformité pour la conservation des objets et les règles de conservation des compartiments par défaut.

Ce guide présente l'API de verrouillage d'objet S3.

## Obligation légale

- La mise en attente légale de verrouillage d'objet est un état activé/désactivé simple appliqué à un objet.

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

- Vérifiez-le avec une opération GET.

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

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

- Désactiver la mise en attente légale

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

- Vérifiez-le avec une opération GET.

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

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

## Mode de conformité

- La conservation de l'objet s'effectue avec une conservation jusqu'à l'horodatage.

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

- Vérifiez l'état de la rétention

```
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"
    }
}
```

## Conservation par défaut

- Définissez la période de conservation en jours et années par rapport à une date de conservation définie avec l'api par objet.

```
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
```

- Vérifiez l'état de la rétention

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

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

- Placer un objet dans le godet

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

- La durée de conservation définie dans le compartiment est convertie en horodatage de conservation sur l'objet.

```
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"
    }
}
```

## Test de la suppression d'un objet avec une rétention définie

Le verrouillage d'objet est basé sur la gestion des versions. La conservation est définie sur une version de l'objet. Si une tentative de suppression d'un objet avec une rétention définie et qu'aucune version n'est spécifiée, un marqueur de suppression est créé comme version actuelle de l'objet.

- Supprimez l'objet dont la conservation est définie

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

- Lister les objets dans le compartiment

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

  - Notez que l'objet n'est pas répertorié.
- Répertorier les versions pour voir le marqueur de suppression et la version verrouillée d'origine

```
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": "566223993089512949 26"
            }
        }
    ],
    "DeleteMarkers": [
        {
            "Owner": {
                "DisplayName": "Tenant01",
                "ID": "566223993089512949 26"
            },
            "Key": "file01.txt",
            "VersionId":
"QjVDQzgzOTAtQ0FGNi0xMUVDLThFMzgtQ0RGMjAwQjk0MjM1",
            "IsLatest": true,
            "LastModified": "2022-05-03T15:35:50.248000+00:00"
        }
    ]
}
```

- Supprimer la version verrouillée de l'objet

```
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
```

*Par Aron Klein*

# Exemples de règles de compartiment et de groupe (IAM

Voici des exemples de politiques de compartiment et de règles de groupe (règles IAM).

## Stratégies de groupe (IAM)

### Accès au compartiment de style Home Directory

Cette stratégie de groupe autorise uniquement les utilisateurs à accéder aux objets du compartiment nommé nom d'utilisateur utilisateurs.

```
"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}/*"
    }

  ]
}
```

### Refuser la création de compartiments de verrouillage d'objet

Cette stratégie de groupe empêche les utilisateurs de créer un compartiment avec le verrouillage d'objet activé sur le compartiment.

> Cette règle n'est pas appliquée dans l'interface utilisateur de StorageGRID et elle n'est appliquée que par l'API S3.

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

**Limite de conservation du verrouillage des objets**

Cette stratégie de compartiment limite la durée de conservation du verrouillage de l'objet à 10 jours ou moins

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

### Empêcher les utilisateurs de supprimer des objets par ID de version

Cette stratégie de groupe empêche les utilisateurs de supprimer des objets multiversion par ID de version

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

Cette stratégie de compartiment empêche un utilisateur (identifié par l'ID utilisateur « 56622399308951294926 ») de supprimer des objets multiversion par l'ID de version

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

**Restriction du compartiment à un seul utilisateur avec un accès en lecture seule**

Cette stratégie permet à un seul utilisateur de disposer d'un accès en lecture seule à un compartiment et d'accéder explicitement à tous les autres utilisateurs. Le regroupement des déclarations de refus en haut de la politique est une bonne pratique pour une évaluation plus rapide.

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

**Limiter un groupe à un sous-répertoire unique (préfixe) avec accès en lecture seule**

Cette règle permet aux membres du groupe d'accéder en lecture seule à un sous-répertoire (préfixe) au sein d'un compartiment. Le nom du compartiment est « Study » et le sous-répertoire est « 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/*"
            ]
        }
    ]
}
```

# Rapports techniques

## NetApp StorageGRID et l’analytique Big Data

### Utilisations de NetApp StorageGRID

La solution de stockage objet NetApp StorageGRID offre évolutivité, disponibilité des données, sécurité et hautes performances. Les entreprises de toutes tailles et de tous secteurs utilisent StorageGRID S3 pour un large éventail d’utilisations. Étudions quelques scénarios types :

**Analytique Big Data :** StorageGRID S3 est fréquemment utilisé comme data Lake, où les entreprises stockent de grandes quantités de données structurées et non structurées à des fins d’analyse à l’aide d’outils tels que Apache Spark, Splunk Smartstore et Dremio.

**Tiering des données :** les clients NetApp utilisent la fonctionnalité FabricPool d’ONTAP pour déplacer automatiquement les données entre un niveau local haute performance et StorageGRID. Le Tiering libère un stockage Flash coûteux pour les données actives tout en maintenant les données inactives disponibles dans un stockage objet à faible coût. Cela optimise les performances et les économies.

**Sauvegarde des données et reprise après incident :** les entreprises peuvent utiliser StorageGRID S3 comme une solution fiable et économique pour sauvegarder des données critiques et les restaurer en cas d’incident.

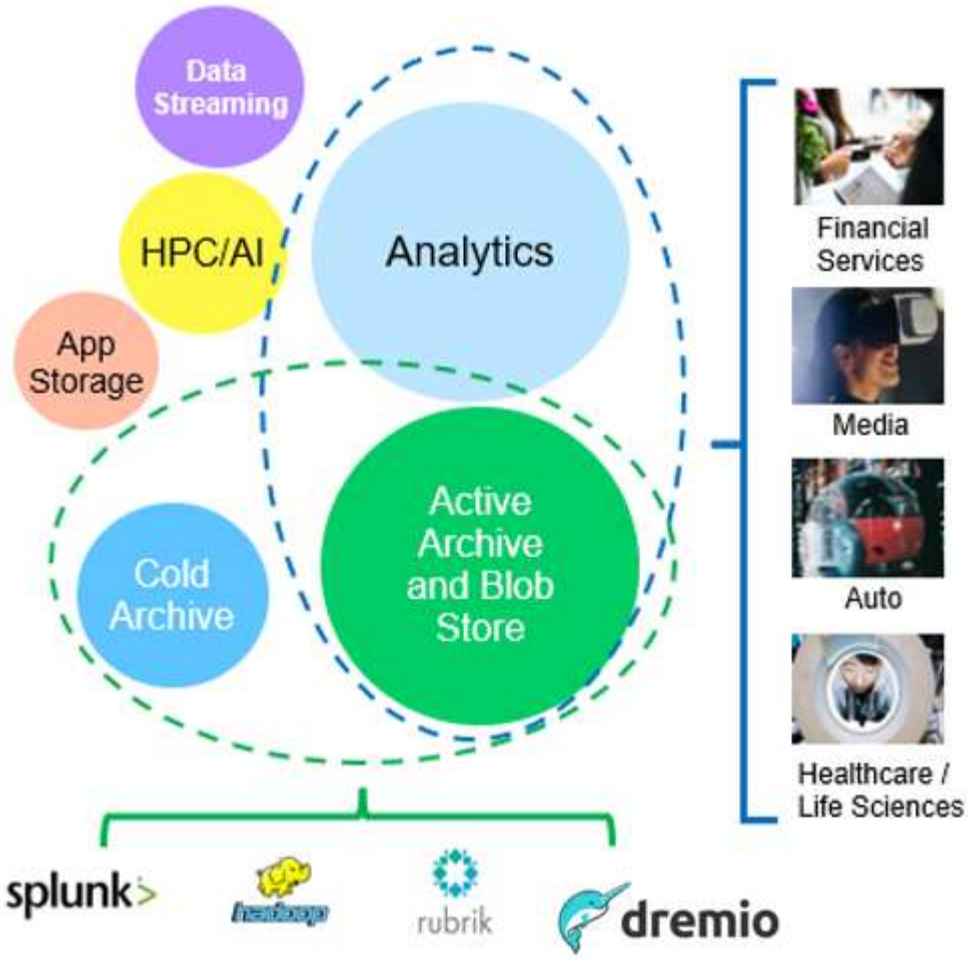
**Stockage des données pour les applications :** StorageGRID S3 peut être utilisé comme backend de stockage pour les applications, ce qui permet aux développeurs de stocker et de récupérer facilement des fichiers, des images, des vidéos et d’autres types de données.

**Diffusion de contenu :** StorageGRID S3 peut être utilisé pour stocker et fournir aux utilisateurs du monde entier du contenu statique, des fichiers multimédias et des téléchargements logiciels, en exploitant la répartition géographique et l’espace de noms global de StorageGRID pour une diffusion de contenu rapide et fiable.

**Tiering des données :** les clients NetApp utilisent la fonction ONTAP FabricPool pour déplacer automatiquement les données entre un niveau local hautes performances vers StorageGRID. Le Tiering libère du stockage Flash coûteux pour les données actives tout en maintenant les données inactives disponibles dans un stockage objet à faible coût. Cela optimise les performances et les économies.

**Archives de données :** StorageGRID offre différents types de stockage et prend en charge la hiérarchisation vers des options de stockage public à faible coût à long terme, en faisant une solution idéale pour l’archivage et la conservation à long terme des données qui doivent être conservées à des fins de conformité ou d’historique.

**Cas d’utilisation du stockage objet**

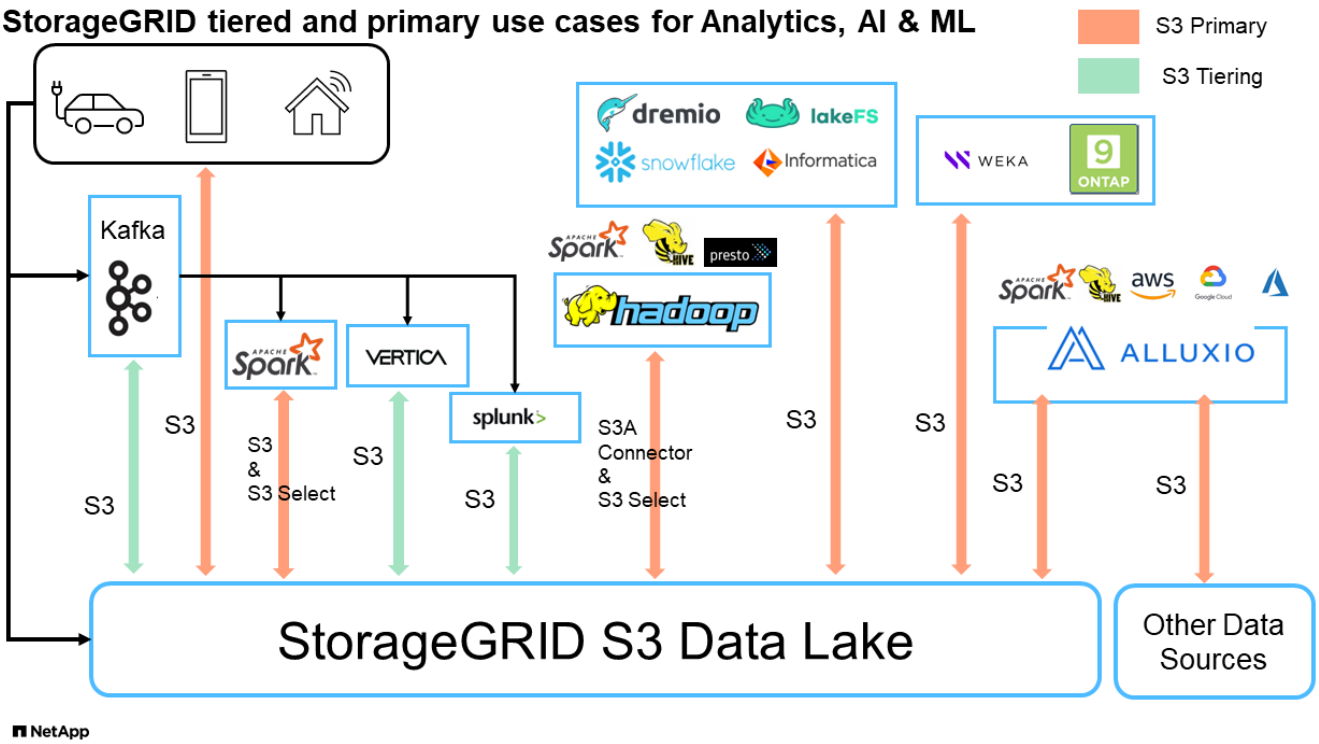



Parmi ces cas d'usage, l'analytique Big Data est l'un des plus utilisés, et son utilisation est en hausse.

## Pourquoi choisir StorageGRID pour les data Lakes ?

- Collaboration renforcée : colocation multisite partagée massive avec accès API standard
- Coûts d'exploitation réduits : simplicité opérationnelle d'une seule architecture à autorétablissement
- Évolutivité : contrairement aux solutions Hadoop et d'entrepôt de données classiques, le stockage objet StorageGRID S3 dissocie le stockage des ressources de calcul et de données pour vous permettre de faire évoluer vos besoins de stockage au fur et à mesure de leur croissance.
- Durabilité et fiabilité : StorageGRID garantit une durabilité de 99.999999999 %, ce qui signifie que les données stockées sont hautement résistantes à la perte de données. Il assure également une haute disponibilité, garantissant ainsi un accès permanent aux données.
- Sécurité : StorageGRID offre plusieurs fonctionnalités de sécurité, notamment le chiffrement, les règles de contrôle d'accès, la gestion du cycle de vie des données, le verrouillage d'objets et la gestion des versions pour protéger les données stockées dans des compartiments S3

**StorageGRID S3 Data Lakes**




## Quel data warehouse ou data Lake fonctionne le mieux avec le stockage objet S3

NetApp a évalué StorageGRID avec trois écosystèmes d'entrepôts de données et de maisons de lac - Hive, Delta Lake et Dremio. "Apache Iceberg : guide de référence" inclut une brève introduction du data warehouse et du data lake house, ainsi que les avantages/inconvénients de ces deux architectures.

- Outil de référence - TPC-DS - https://www.tpc.org/tpcds/
- Les écosystèmes Big Data
  - Cluster de 5 machines virtuelles, chacune avec 128 G de RAM et 24 vCPU, stockage SSD pour le disque système
  - Hadoop 3.3.5 avec Hive 3.1.3 (1 nœud de nom + 4 nœuds de données)
  - Delta Lake avec Spark 3.2.0 (1 maître + 4 employés) et Hadoop 3.3.5
  - Dremio v23 (1 maître + 4 exécuteurs)
- Stockage objet
  - NetApp® StorageGRID® 11.6 avec 3 x SG6060 + 1 équilibreur de charge SG1000
  - Protection objet : 2 copies
- Taille de base de données : 1 000 Go
- Cache désactivé sur les 3 écosystèmes pour obtenir un résultat cohérent pour chaque test de requête.

TPC-DS est fourni avec 99 requêtes SQL complexes pour l'analyse comparative des requêtes. Nous avons mesuré le nombre total de minutes nécessaires pour effectuer les 99 requêtes et nous avons analysé le résultat plus en détail en dépanne le type et le nombre de requêtes S3. Le premier tableau ci-dessous présente la durée totale des 99 requêtes et le second tableau résume le nombre et les types de requêtes S3 envoyées à chaque écosystème à StorageGRID.

**Résultat de la requête TPC-DS**

| Écosystème | Ruche | Delta Lake | Dremio |
|---|---|---|---|
| La couche de stockage | NetApp® StorageGRID® | NetApp® StorageGRID® | NetApp® StorageGRID® |
| Type de disque | DISQUES DURS | DISQUES DURS | DISQUES DURS |
| Format de tableau | Parquet | Parquet | Parquet [1] |
| Taille de la base de données | 1 000 G | 1 000 G | 1 000 G |
| Requêtes TPCDS 99 nombre total de minutes | 1084 [2] | 55 | 47 |

[1] testé les formats de table parquet et Iceberg, le résultat est similaire.

[2] Hive Impossible de compléter la requête numéro 72.

**Requêtes TPC-DS - ventilation des requêtes S3**

| Requêtes S3 | Ruche | Delta Lake | Dremio |
|---|---|---|---|
| OBTENEZ | 1,117,184 | 2,074,610 | 4,414,227 |
| observation: Tous les ACCÈS à la gamme | Plage de 80 % de 2 Ko à 2 Mo à partir d'objets de 32 Mo, 50 à 100 requêtes/sec | Plage de 73 % inférieure à 100 Ko pour les objets de 32 Mo, 1000 à 1400 requêtes/sec | 90 % plage d'octets de 1 Mo provenant d'objets de 256 Mo, 2000 à 2300 requêtes/sec |
| Liste des objets | 312,053 | 24,158 | 240 |
| TÊTE (objet inexistant) | 156,027 | 12,103 | 192 |
| TÊTE (objet existant) | 982,126 | 922,732 | 1,845 |
| Nombre total de demandes | 2,567,390 | 3,033,603 | 4,416,504 |

À partir de la première table, nous pouvons voir Delta Lake et Dremio sont beaucoup plus rapides que Hive. À partir du second tableau, Hive a envoyé de nombreuses demandes d'objets de liste S3, qui sont généralement lentes dans toutes les plateformes de stockage objet, en particulier si le compartiment contient de nombreux objets. Cela augmente considérablement la durée globale des requêtes. Une autre observation est Dremio a pu envoyer un grand nombre de requêtes GET en parallèle, 2,000 à 2,300 requêtes par seconde contre 50 à 100 requêtes par seconde à Hive. Le système de fichiers standard du modèle ruve et Hadoop S3A contribue à la lenteur de Hive dans le stockage objet S3.

Pour utiliser Hadoop (HDFS ou le stockage objet S3) avec Hive ou Spark, il est nécessaire de disposer de connaissances approfondies sur Hadoop et Hive/Spark et sur l'interaction entre les paramètres de chaque service. Ensemble, ils disposent de plus de 1000 paramètres. Très souvent, les paramètres sont interdépendants et ne peuvent pas être modifiés seuls. Il faut beaucoup de temps et d'efforts pour trouver la combinaison optimale de paramètres et de valeurs à utiliser.

Dremio est un moteur de data Lake qui utilise Apache Arrow de bout en bout pour améliorer considérablement les performances de requêtes. Apache Arrow propose un format de mémoire standard par colonnes pour un partage efficace des données et une analyse rapide. Arrow utilise une approche indépendante du langage, conçue pour éliminer le besoin de sérialisation et de désérialisation des données, améliorant ainsi les

performances et l'interopérabilité entre les processus et les systèmes de données complexes.

Les performances de Dremio sont principalement déterminées par la puissance de calcul sur le cluster Dremio. Bien que Dremio utilise le connecteur S3A de Hadoop pour la connexion de stockage d'objets S3, Hadoop n'est pas nécessaire et la plupart des paramètres fs.s3a de Hadoop ne sont pas utilisés par Dremio. Cela facilite le réglage des performances de Dremio sans passer de temps à apprendre et à tester différents paramètres Hadoop s3a.

À partir de ce résultat du banc d'essai, nous pouvons conclure que le système d'analytique Big Data optimisé pour la charge de travail S3 constitue un facteur de performance majeur. Dremio optimise l'exécution des requêtes, utilise efficacement les métadonnées et fournit un accès transparent aux données S3. Il offre ainsi de meilleures performances que Hive avec le stockage S3. Se reporter à ceci "page" Pour configurer la source de données Dremio S3 avec StorageGRID.

Cliquez sur les liens ci-dessous pour découvrir comment StorageGRID et Dremio travaillent en collaboration pour fournir une infrastructure de data Lake moderne et efficace, et comment NetApp a migré de Hive + HDFS vers Dremio + StorageGRID pour améliorer considérablement l'efficacité de l'analyse Big Data.

- "Optimisez les performances de vos Big Data avec NetApp StorageGRID"
- "Infrastructure de data Lake moderne, puissante et efficace avec StorageGRID et Dremio"
- "Comment NetApp redéfinit l'expérience client avec l'analytique des produits"

# Réglage Hadoop S3A

Le connecteur Hadoop S3A facilite l'interaction transparente entre les applications Hadoop et le stockage objet S3. Le réglage du connecteur Hadoop S3A est essentiel pour optimiser les performances lorsque vous travaillez avec le stockage objet S3. Avant d'entrer dans les détails d'ajustement, analysons très bien Hadoop et ses composants.

## Qu'est-ce que Hadoop ?

**Hadoop** est une structure open source puissante conçue pour gérer le traitement et le stockage de données à grande échelle. Il permet le stockage distribué et le traitement parallèle sur des clusters d'ordinateurs.

Ces trois composants sont les suivants :

- **Hadoop HDFS (Hadoop Distributed File System)** : gère le stockage, décode les données en blocs et les distribue entre les nœuds.
- **Hadoop MapReduce** : responsable du traitement des données en divisant les tâches en petits blocs et en les exécutant en parallèle.
- **FIL Hadoop (encore un autre négociateur de ressources):** "Gère les ressources et planifie les tâches de manière efficace"

## Connecteur HDFS et S3A Hadoop

HDFS est un composant essentiel de l'écosystème Hadoop, qui joue un rôle essentiel dans l'efficacité du traitement des Big Data. HDFS assure un stockage et une gestion fiables. Elle assure un traitement parallèle et un stockage des données optimisé, ce qui accélère l'accès aux données et leur analyse.

Dans le traitement du Big Data, HDFS se distingue par son excellente tolérance aux pannes pour le stockage de datasets volumineux. Pour cela, il s'agit de la réplication des données. Il peut stocker et gérer d'importants volumes de données structurées et non structurées dans un environnement de data warehouse. De plus, il

s'intègre en toute transparence aux principales structures de traitement des Big Data, comme Apache Spark, Hive, Pig et Flink, pour un traitement des données évolutif et efficace. Il est compatible avec les systèmes d'exploitation Unix (Linux), ce qui en fait un choix idéal pour les entreprises qui préfèrent utiliser des environnements Linux pour leur traitement Big Data.

Comme le volume de données s'est accru au fil du temps, l'approche consistant à ajouter de nouvelles machines au cluster Hadoop avec leurs propres ressources de calcul et de stockage s'avère inefficace. L'évolutivité linéaire engendre des défis pour utiliser les ressources efficacement et gérer l'infrastructure.

Pour relever ces défis, le connecteur Hadoop S3A offre des E/S haute performance par rapport au stockage objet S3. L'implémentation d'un workflow Hadoop avec S3A vous permet d'exploiter le stockage objet en tant que référentiel de données et de séparer les ressources de calcul et de stockage. Vous pouvez ainsi faire évoluer indépendamment les ressources de calcul et de stockage. Grâce à la dissociation du calcul et du stockage, vous pouvez également dédier la bonne quantité de ressources pour vos tâches de calcul et fournir la capacité requise en fonction de la taille de votre jeu de données. Par conséquent, vous pouvez réduire votre TCO global pour les workflows Hadoop.

## Réglage du connecteur S3A Hadoop

S3 se comporte différemment de HDFS et certaines tentatives de préservation de l'apparence d'un système de fichiers ne sont pas totalement optimales. Des ajustements/tests/tests rigoureux sont nécessaires pour optimiser l'utilisation des ressources S3.

Les options Hadoop présentées dans ce document sont basées sur Hadoop 3.3.5, voir "Hadoop 3.3.5 core-site.xml" pour toutes les options disponibles.

Remarque – la valeur par défaut de certains paramètres Hadoop fs.s3a est différente dans chaque version de Hadoop. Vérifiez la valeur par défaut spécifique à votre version Hadoop actuelle. Si ces paramètres ne sont pas spécifiés dans Hadoop core-site.xml, la valeur par défaut sera utilisée. Vous pouvez remplacer la valeur au moment de l'exécution à l'aide des options de configuration Spark ou Hive.

Vous devez accéder à cette page "Page Apache Hadoop" pour comprendre chaque option fs.s3a. Si possible, testez-les dans un cluster Hadoop non productif pour trouver les valeurs optimales.

Vous devriez lire "Optimisation des performances lors de l'utilisation du connecteur S3A" pour obtenir d'autres recommandations de réglage.

Examinons quelques points clés à prendre en compte :

**1. Compression des données**

N'activez pas la compression StorageGRID. La plupart des systèmes Big Data utilisent l'option GET de plage d'octets au lieu de récupérer l'objet entier. L'utilisation de la plage d'octets GET avec des objets compressés dégrade considérablement les performances GET.

**2. S3A committers**

En général, le Comitter Magic s3a est recommandé. Se reporter à ceci "Page des options de renvoi S3A courantes" pour mieux comprendre le comitter magique et ses paramètres s3a associés.

Magic Committer :

Le Magic Committer s'appuie spécifiquement sur S3Guard pour offrir des listes de répertoires cohérentes sur le magasin d'objets S3.

Avec S3 cohérent (ce qui est désormais le cas), le Magic Committer peut être utilisé en toute sécurité avec n'importe quel compartiment S3.

Choix et expérimentation :

Selon votre cas d'utilisation, vous pouvez choisir entre la variable de transfert (qui s'appuie sur un système de fichiers HDFS de cluster) et la variable Magic Committer.

Testez les deux pour déterminer celle qui convient le mieux à votre workload et à vos besoins.

En résumé, les committers S3A constituent une solution au défi fondamental de l'engagement de sortie cohérent, haute performance et fiable pour S3. Leur conception interne garantit un transfert de données efficace tout en préservant l'intégrité des données.

| 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. Threads, tailles de pool de connexions et taille de bloc

- Chaque client **S3A** interagissant avec un seul compartiment dispose de son propre pool dédié de connexions HTTP 1.1 ouvertes et de threads pour les opérations de téléchargement et de copie.
- "Vous pouvez régler la taille de ces pools de manière à trouver un équilibre entre les performances et l'utilisation de la mémoire/des threads".
- Lors du téléchargement de données vers S3, elles sont divisées en blocs. La taille de bloc par défaut est de 32 Mo. Vous pouvez personnaliser cette valeur en définissant la propriété fs.s3a.block.size.
- Des blocs plus volumineux peuvent améliorer les performances lors du chargement de données volumineuses en réduisant la surcharge liée à la gestion des pièces à part multiple lors du téléchargement. La valeur recommandée est de 256 Mo ou plus pour les jeux de données volumineux.

| 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. Téléchargement partitionné

s3a committers **toujours** utiliser MPU (téléchargement partitionné) pour charger des données dans le compartiment s3. Ceci est nécessaire pour permettre : l'échec de tâche, l'exécution spéculative des tâches et les abandons de travail avant la validation. Voici quelques spécifications clés relatives aux téléchargements partitionnés :

- Taille maximale des objets : 5 Tio (téraoctets).
- Nombre maximum de pièces par téléchargement: 10,000.
- Numéros de pièce : compris entre 1 et 10,000 (inclus).
- Taille de la pièce : entre 5 Mio et 5 Gio. En particulier, il n'existe pas de limite de taille minimale pour la dernière partie de votre téléchargement partitionné.

L'utilisation d'une taille de pièce plus petite pour les téléchargements partitionnés S3 présente à la fois des avantages et des inconvénients.

**Avantages** :

- Récupération rapide à partir des problèmes réseau : lorsque vous chargez des pièces plus petites, l'impact du redémarrage d'un téléchargement échoué en raison d'une erreur réseau est réduit. Si une pièce

échoue, il vous suffit de télécharger à nouveau cette pièce spécifique plutôt que l'objet entier.

- Meilleure parallélisation : plus de pièces peuvent être téléchargées en parallèle, ce qui permet de tirer parti du multithreading ou des connexions simultanées. Cette parallélisation améliore les performances, en particulier pour les fichiers volumineux.

**Désavantage** :

- Surcharge réseau : une taille de pièce plus petite signifie plus de parties à télécharger, chaque partie nécessite sa propre requête HTTP. Le nombre de requêtes HTTP augmente la charge de lancement et de traitement des requêtes individuelles. La gestion d'un grand nombre de petites pièces peut avoir un impact sur les performances.
- Complexité : la gestion de la commande, le suivi des pièces et la garantie de la réussite des téléchargements peuvent s'avérer fastidieux. Si le téléchargement doit être abandonné, tous les articles déjà téléchargés doivent être suivis et purgés.

Pour Hadoop, la taille de pièce de 256 Mo ou plus est recommandée pour fs.s3a.multipart.size. Définissez toujours la valeur fs.s3a.mutipart.threshold sur 2 x fs.s3a.multipart.size. Par exemple, si fs.s3a.multipart.size = 256M, fs.s3a.mutlipart.threshold doit être de 512M.

Utiliser une taille de pièce plus grande pour un jeu de données volumineux. Il est important de choisir une taille de pièce qui équilibre ces facteurs en fonction de votre cas d'utilisation et des conditions réseau spécifiques.

Un téléchargement partitionné est un "processus en trois étapes":

1. Le téléchargement est lancé, StorageGRID renvoie un ID de téléchargement
2. Les parties d'objet sont chargées à l'aide de l'ID de téléchargement
3. Une fois toutes les parties d'objet chargées, envoie une demande de téléchargement partitionné complète avec upload-ID StorageGRID construit l'objet à partir des pièces téléchargées, et le client peut accéder à l'objet.

Si la demande complète de téléchargement partitionné n'est pas envoyée correctement, les pièces restent dans StorageGRID et ne créeront aucun objet. Cela se produit lorsque les travaux sont interrompus, en échec ou abandonnés. Les pièces restent dans la grille jusqu'à ce que le téléchargement partitionné soit terminé ou abandonné ou que StorageGRID purge ces pièces si 15 jours se sont écoulés depuis le lancement du téléchargement. S'il y a beaucoup (quelques centaines de milliers à plusieurs millions) de téléchargements partitionnés en cours dans un compartiment, lorsque Hadoop envoie des « téléchargements partiaux-listes » (cette requête ne filtre pas par identifiant de téléchargement), la demande peut prendre un certain temps ou finir par se terminer. Vous pouvez envisager de définir fs.s3a.mutipart.purge sur TRUE avec une valeur fs.s3a.multipart.purge.age appropriée (par exemple, 5 à 7 jours, n'utilisez pas la valeur par défaut de 86400, c'est-à-dire 1 jour). Ou faites appel au support NetApp pour étudier la situation.

| 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. Mémoire tampon pour écrire les données en mémoire**

Pour améliorer les performances, vous pouvez mettre en mémoire tampon l'écriture des données en mémoire avant de les télécharger dans S3. Cela permet de réduire le nombre d'écritures de petite taille et d'améliorer l'efficacité.

| 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 |

N’oubliez pas que S3 et HDFS fonctionnent différemment. Des ajustements/tests/expériences minutieux sont nécessaires pour utiliser de manière optimale les ressources S3.

# Blogs NetApp StorageGRID

Vous trouverez d'excellents blogs NetApp StorageGRID ici :

- Mai 10 : "Lab on Demand est votre meilleur outil de vente pour StorageGRID"
- Mai 24 : "Modernisez vos workloads d'analytique avec NetApp et Alluxio"
- Mai 26 : "StorageGRID : stockage et gestion des données de réplication et de sauvegarde sur site"
- 9 juin : "Utilisez le connecteur Cloudera Hadoop S3A avec StorageGRID"
- 26 juillet : "Consultez la liste croissante des solutions partenaires validées pour StorageGRID"
- 5 août : "NetApp StorageGRID obtient la certification de sécurité Common Criteria"
- 16 août : "Intégration de StorageGRID à la pile ELK open source pour améliorer l'expérience client"
- 17 août : "Tout commence par le verrouillage d'objet… Création d'un écosystème de stockage S3 pour les applications de sauvegarde stratégiques"
- 23 août : "Bâtissez votre data Lake sur StorageGRID"
- 1er septembre : "Prenez ces mesures et Graph IT"
- 19 septembre : "Prise en charge de DataLock et de la protection contre les ransomware pour StorageGRID"
- 26 septembre : "NetApp StorageGRID pour les fournisseurs de services"
- 5 octobre : "Décongelez vos données sur StorageGRID pour Snowflake"
- 5 octobre : "NetApp Cloud Insights ajoute les tableaux de bord de la galerie StorageGRID"
- Novembre 7 : "Prise en charge d'StorageGRID et d'ONTAP S3 : différences, similarités et intégration"
- Novembre 23 : "Découvrez l'IA explicable avec MLOps optimisée par NetApp et Modzy"
- Décembre 6 : "StorageGRID obtient la certification de conformité KPMG"
- Janvier 16 : "StorageGRID renouvelle la certification de conformité NF203 et ISO/IEC 25051"
- Janvier 18 : "Verrouillage objet StorageGRID S3 validé pour Veritas NetBackup"
- 14 février : "Qu'ont en commun le chocolat, le ski, les montres et les gros systèmes ?"
- 14 mars : "Comment sauvegarder des bases de données DME des systèmes Epic à l'aide d'une seule commande dans une architecture 3:2:1"
- 30 mars : "Utilisez BlueXP pour protéger les dossiers médicaux électroniques Epic avec une règle de sauvegarde conforme à 3:2:1"
- 30 mars : "Point de montage pour Amazon S3 version alpha avec StorageGRID"
- Mai 16 : "Nouveautés de la gamme de stockage objet StorageGRID"
- Mai 16 : "Présentation d'StorageGRID 11.7 et du nouveau système SGF6112 de stockage objet 100 % Flash"
- 30 août : "Le point de montage pour Amazon S3 File System est désormais GA"
- 1er septembre : "Utilisation de Cloud Insights pour surveiller et collecter les journaux à l'aide du bit Fluent"
- 17 octobre : "Hadoop, modernisation de l'analytique avec Dremio et StorageGRID"
- Novembre 7 : "Spectra Logic On-sur-site Glacier avec StorageGRID"
- Décembre 12 : "Analytique Big Data sur StorageGRID : Dremio est 23 fois plus rapide qu'Apache Hive"

- 2 févr. : "Annonce de la description de la solution StorageGRID + lakeFS"
- 16 février : "Présentation de StorageGRID 11.8 : sécurité améliorée, simplicité et expérience utilisateur"
- 16 février : "Présentation de StorageGRID 11.8"

# Documentation NetApp StorageGRID

Tous les documents relatifs à chaque version de NetApp StorageGRID sont disponibles ici :

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

# Mentions légales

Les mentions légales donnent accès aux déclarations de copyright, aux marques, aux brevets, etc.

## Droits d’auteur

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

## Marques déposées

NetApp, le logo NETAPP et les marques mentionnées sur la page des marques commerciales NetApp sont des marques commerciales de NetApp, Inc. Les autres noms de sociétés et de produits peuvent être des marques commerciales de leurs propriétaires respectifs.

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

## Brevets

Vous trouverez une liste actuelle des brevets appartenant à NetApp à l’adresse suivante :

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

## Politique de confidentialité

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

## Source ouverte

Les fichiers de notification fournissent des informations sur les droits d’auteur et les licences de tiers utilisés dans le logiciel NetApp.

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

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

## Informations sur le copyright



## Informations sur les marques commerciales