# Guias de ferramentas e aplicações

StorageGRID solutions and resources

NetApp
December 10, 2025

# Índice

# Guias de ferramentas e aplicações

## Use o conetor Cloudera Hadoop S3A com StorageGRID

*Por Angela Cheng*

Hadoop tem sido um favorito dos cientistas de dados há algum tempo. O Hadoop permite o processamento distribuído de grandes conjuntos de dados entre clusters de computadores usando estruturas de programação simples. O Hadoop foi projetado para escalar de servidores únicos para milhares de máquinas, com cada máquina possuindo computação e armazenamento locais.

### Por que usar o S3A para fluxos de trabalho Hadoop?

À medida que o volume de dados cresceu com o tempo, a abordagem de adicionar novas máquinas com sua própria computação e storage tornou-se ineficiente. O dimensionamento linear cria desafios para o uso eficiente de recursos e o gerenciamento da infraestrutura.

Para lidar com esses desafios, o cliente Hadoop S3A oferece e/S de alto desempenho em relação ao storage de objetos S3. A implementação de um fluxo de trabalho do Hadoop com o S3A ajuda você a utilizar o storage de objetos como repositório de dados e permite separar a computação e o storage, o que, por sua vez, permite escalar a computação e o storage de forma independente. A dissociação da computação e do storage também permite que você dedique a quantidade certa de recursos para suas tarefas de computação e forneça capacidade com base no tamanho do conjunto de dados. Portanto, você pode reduzir o TCO geral para workflows do Hadoop.

### Configure o conetor S3A para usar o StorageGRID

#### Pré-requisitos

- Um URL de endpoint do StorageGRID S3, uma chave de acesso do locatário S3 e uma chave secreta para o teste de conexão do Hadoop S3A.
- Um cluster Cloudera e uma permissão root ou sudo para cada host no cluster para instalar o pacote Java.

Em abril de 2022, o Java 11.0.14 com Cloudera 7.1.7 foi testado contra o StorageGRID 11,5 e 11,6. No entanto, o número da versão Java pode ser diferente no momento de uma nova instalação.

#### Instale o pacote Java

1. Verifique "Matriz de suporte Cloudera" se há a versão do JDK suportada.
2. Faça o download do "Pacote Java 11.x" que corresponde ao sistema operacional do cluster Cloudera. Copie este pacote para cada host no cluster. Neste exemplo, o pacote rpm é usado para o CentOS.
3. Faça login em cada host como root ou usando uma conta com permissão sudo. Execute as seguintes etapas em cada host:
   a. Instale o pacote:

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

   b. Verifique onde o Java está instalado. Se várias versões estiverem instaladas, defina a versão recém-instalada como padrão:

```
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. Adicione esta linha ao final `/etc/profile` do . O caminho deve corresponder ao caminho da seleção acima:

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

d. Execute o seguinte comando para que o perfil entre em vigor:

```
source /etc/profile
```

**Configuração Cloudera HDFS S3A**

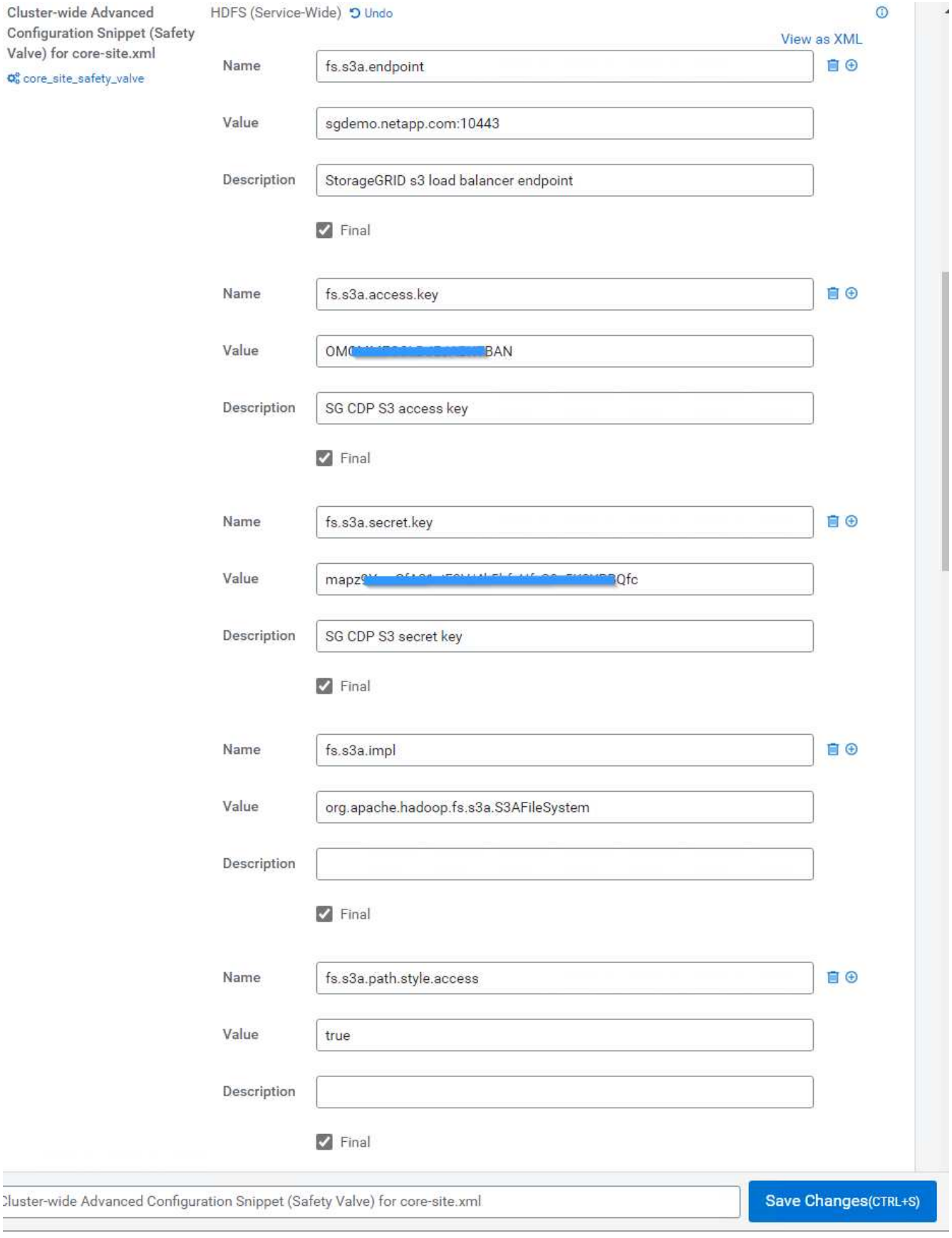
**Passos**

1. Na GUI do Cloudera Manager, selecione clusters > HDFS e selecione Configuração.
2. NA CATEGORIA, selecione Avançado e role para baixo para localizar `Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml`.
3. Clique no sinal e adicione os seguintes pares de valores.

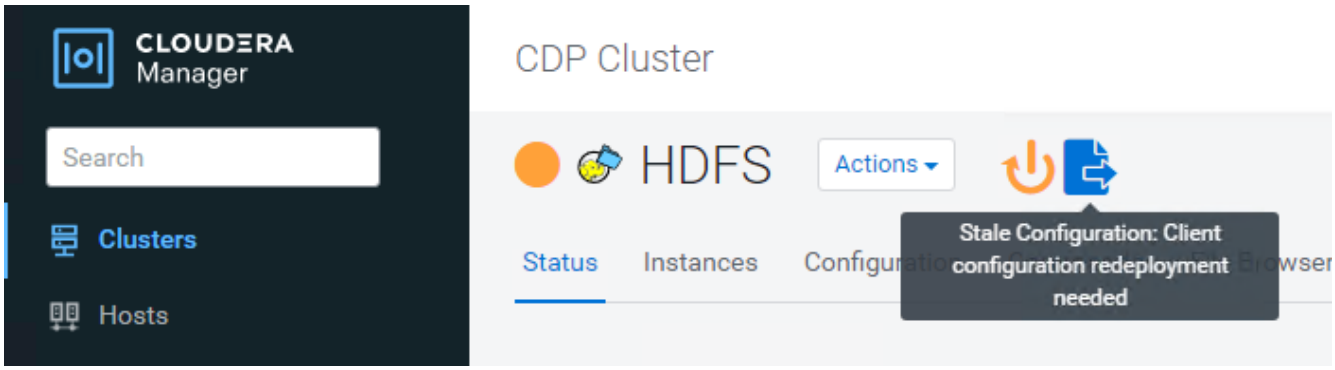
| Nome | Valor |
|---|---|
| fs.s3a.access.key | _ Chave de acesso S3 do cliente a partir de StorageGRID>_ |
| fs.s3a.secret.key | _ Chave secreta do cliente S3 da StorageGRID>_ |
| fs.s3a.connection.ssl.enabled | [true ou false] (o padrão é https se esta entrada estiver ausente) |
| fs.s3a.endpoint | _ Endpoint do cliente StorageGRID S3:port>_ |
| fs.s3a.impl | org.apache.hadoop.fs.s3a.S3AFileSystem |

| Nome | Valor |
| --- | --- |
| fs.s3a.path.style.access | [verdadeiro ou falso] (o padrão é o estilo de host virtual se essa entrada estiver ausente) |

**Captura de tela de amostra**




4. Clique no botão Salvar alterações. Selecione o ícone Configuração obsoleta na barra de menus do HDFS, selecione Reiniciar Serviços obsoletos na próxima página e selecione Reiniciar agora.




## Teste a conexão S3A com o StorageGRID

### Execute o teste básico de conexão

Faça login em um dos hosts no cluster Cloudera e `hadoop fs -ls s3a://*<bucket-name>*/`digite .

O exemplo a seguir usa syle de caminho com um bucket de teste hdfs pré-existente e um objeto de teste.

```
[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.
```

### Solução de problemas

#### Cenário 1

Use uma conexão HTTPS com o StorageGRID e obtenha um `handshake_failure` erro após um tempo limite de 15 minutos.

**Motivo:** versão antiga do JRE/JDK usando pacote de codificação TLS desatualizado ou não suportado para conexão com o StorageGRID.

- Exemplo de mensagem de erro*

```
[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
```

**Resolução:** Certifique-se de que o JDK 11.x ou posterior esteja instalado e definido como padrão a biblioteca Java. Consulte Instale o pacote Javaa secção para obter mais informações.

**Cenário 2:**

Falha ao se conetar ao StorageGRID com mensagem de erro `Unable to find valid certification path to requested target`.

**Razão:** o certificado do servidor de endpoint StorageGRID S3 não é confiável pelo programa Java.

Exemplo de mensagem de erro:

```
[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
```

**Resolução:** a NetApp recomenda o uso de um certificado de servidor emitido por uma autoridade pública de assinatura de certificado conhecida para garantir que a autenticação seja segura. Como alternativa, adicione uma CA personalizada ou certificado de servidor ao armazenamento de confiança Java.

Siga as etapas a seguir para adicionar uma CA personalizada do StorageGRID ou um certificado de servidor ao armazenamento de confiança do Java.

1. Faça backup do arquivo Java cacerts padrão existente.

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

2. Importe o cert de endpoint do StorageGRID S3 para o armazenamento de confiança 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>
```

#### Dicas de solução de problemas

1. Aumente o nível de log do hadoop para DEPURAR.

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

2. Execute o comando e direcione as mensagens de log para error.log.

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

*Por Angela Cheng*

# Use o S3cmd para testar e demonstrar o acesso S3 no StorageGRID

*Por Aron Klein*

S3cmd é uma ferramenta de linha de comando gratuita e cliente para operações S3. Você pode usar o s3cmd para testar e demonstrar o acesso S3 no StorageGRID.

## Instale e configure o S3cmd

Para instalar o S3cmd em uma estação de trabalho ou servidor, faça o download do "Linha de comando S3 cliente". o s3cmd é pré-instalado em cada nó do StorageGRID como uma ferramenta para auxiliar na solução de problemas.

## Etapas iniciais de configuração

1. s3cmd --configure
2. Forneça apenas access_key e secret_key, para que o resto mantenha os padrões.
3. Testar o acesso com as credenciais fornecidas? [Y/n]: N (ignorar o teste, pois ele falhará)
4. Guardar definições? [y/N] y
   a. Configuração guardada em '/root/.s3cfg'
5. Em .s3cfg, deixe os campos host_base e host_bucket vazios após o sinal "
   a. base_de_host
   b. host_bucket

Se você especificar host_base e host_bucket na etapa 4, não será necessário especificar um endpoint com --host na CLI. Exemplo:

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

## Exemplos básicos de comandos

- **Crie um bucket:**

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

- **Liste todos os baldes:**

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

- **Liste todos os baldes e seus conteúdos:**

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

- **Liste objetos em um bucket específico:**

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

- **Excluir um balde:**

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

- **Coloque um objeto:**

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

- **Obter um objeto:**

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

- **Excluir um objeto:**

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

# Banco de dados do modo Eon usando NetApp StorageGRID como armazenamento comunitário

*Por Angela Cheng*

Este guia descreve o procedimento para criar um banco de dados do modo Vertica Eon com armazenamento comunitário no NetApp StorageGRID.

## Introdução

Vertica é um software de gerenciamento de banco de dados analítico. É uma plataforma de armazenamento colunar projetada para lidar com grandes volumes de dados, o que permite um desempenho de consulta muito rápido em um cenário tradicionalmente intensivo. Um banco de dados Vertica é executado em um dos dois modos: EON ou Enterprise. Você pode implantar os dois modos no local ou na nuvem.

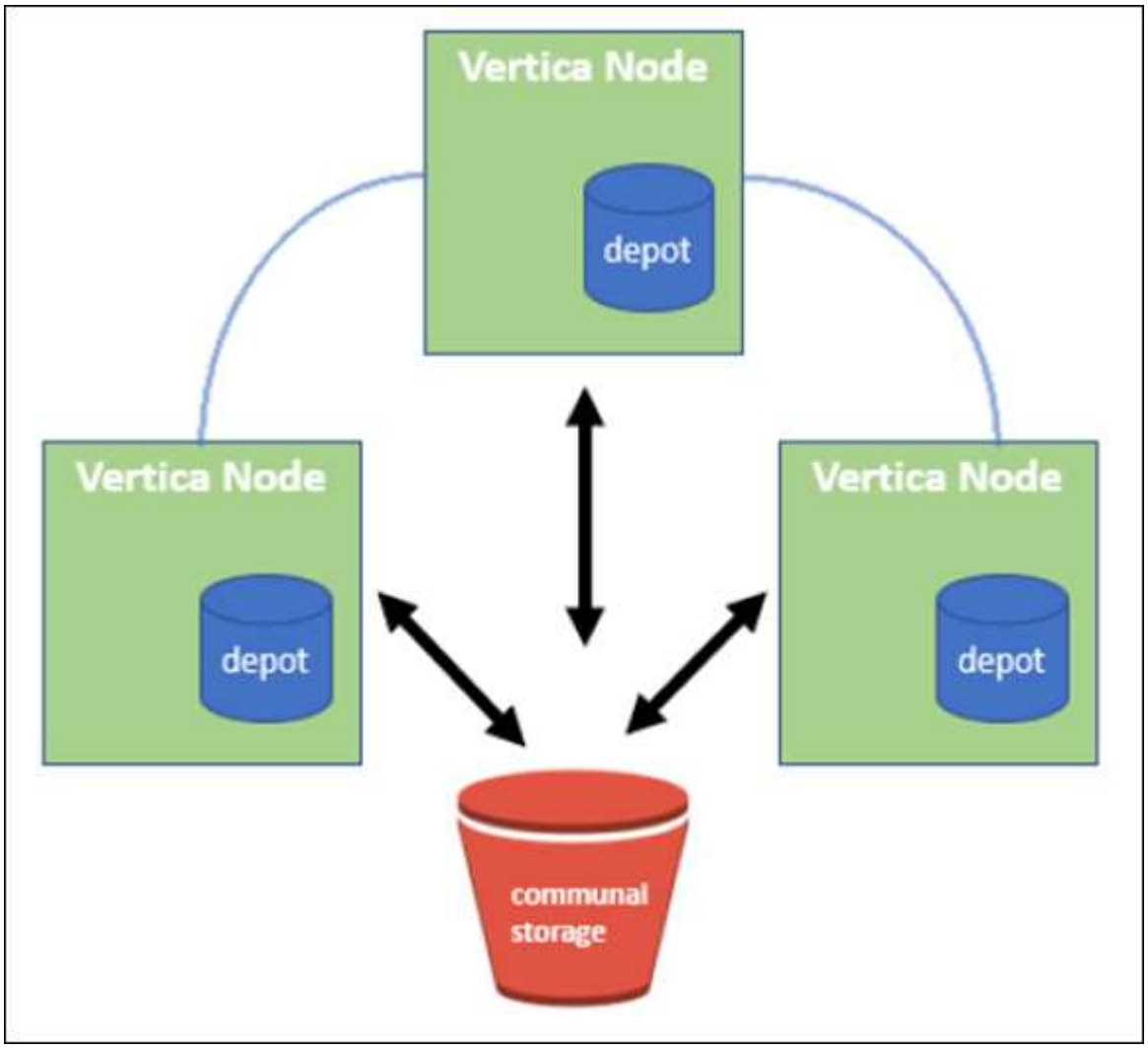
Os modos EON e Enterprise diferem principalmente no local onde armazenam dados:

- As bases de dados do modo EON utilizam armazenamento comunitário para os seus dados. Isso é recomendado pela Vertica.
- Os bancos de dados do modo empresarial armazenam dados localmente no sistema de arquivos de nós que compõem o banco de dados.

**Arquitetura do modo EON**

O modo EON separa os recursos computacionais da camada de armazenamento comum do banco de dados, o que permite que a computação e o armazenamento sejam dimensionados separadamente. O Vertica no modo Eon é otimizado para lidar com cargas de trabalho variáveis e isolá-las umas das outras usando recursos de computação e armazenamento separados.

O modo EON armazena dados em um armazenamento de objetos compartilhado chamado armazenamento comunitário - um bucket do S3, hospedado no local ou no Amazon S3.




**Armazenamento comunitário**

Em vez de armazenar dados localmente, o modo Eon usa um único local de armazenamento comunitário para todos os dados e o catálogo (metadados). O armazenamento comum é o local de armazenamento centralizado do banco de dados, compartilhado entre os nós do banco de dados.

O armazenamento comunitário tem as seguintes propriedades:

- O armazenamento comum na nuvem ou no local de objetos é mais resiliente e menos suscetível à perda de dados devido a falhas de armazenamento do que armazenamento em disco em máquinas individuais.
- Todos os dados podem ser lidos por qualquer nó usando o mesmo caminho.
- A capacidade não é limitada pelo espaço de disco nos nós.
- Como os dados são armazenados em comunidade, você pode dimensionar elasticamente seu cluster para atender às demandas em constante mudança. Se os dados fossem armazenados localmente nos nós, adicionar ou remover nós exigiria a movimentação de quantidades significativas de dados entre nós para movê-los dos nós que estão sendo removidos ou para nós recém-criados.

**O depósito**

Uma desvantagem do armazenamento comunitário é a sua velocidade. Acessar dados de um local compartilhado na nuvem é mais lento do que lê-los a partir do disco local. Além disso, a conexão com o armazenamento comunitário pode se tornar um gargalo se muitos nós estiverem lendo dados de uma só vez. Para melhorar a velocidade de acesso aos dados, os nós em um banco de dados do modo Eon mantêm um cache de dados de disco local chamado de depósito. Ao executar uma consulta, os nós primeiro verificam se os dados de que precisam estão no depósito. Se for, então ele termina a consulta usando a cópia local dos dados. Se os dados não estiverem no depósito, o nó buscará os dados do armazenamento comunitário e salvará uma cópia no depósito.

## Recomendações do NetApp StorageGRID

Vertica armazena dados de banco de dados para armazenamento de objetos como milhares (ou milhões) de objetos compatados (o tamanho observado é de 200 a 500MB por objeto. Quando um usuário executa consultas de banco de dados, o Vertica recupera o intervalo de dados selecionado desses objetos compatados em paralelo usando a chamada DE RECEBIMENTO DE intervalo de bytes. Cada intervalo de bytes GET é de aproximadamente 8KB.

Durante o teste de consultas de usuários do 10TBo depósito do banco de dados, 4.000 a 10.000 solicitações GET (byte-range GET) por segundo foram enviadas para a grade. Ao executar esse teste usando dispositivos SG6060, embora a % de utilização de CPU por nó de appliance seja baixa (cerca de 20% a 30%), 2/3x do tempo de CPU está aguardando a e/S. Uma porcentagem muito pequena (0% a 0,5%) de espera de e/S é observada no SGF6024.

Devido à alta demanda de IOPS pequenos com requisitos de latência muito baixos (a média deve ser inferior a 0,01 segundos), a NetApp recomenda o uso do SFG6024 para serviços de storage de objetos. Se o SG6060 for necessário para tamanhos de banco de dados muito grandes, o cliente deve trabalhar com a equipe de contas Vertica no dimensionamento do depósito para oferecer suporte ao conjunto de dados ativamente consultado.

Para o nó Admin e o nó API Gateway, o cliente pode usar o SG100 ou o SG1000. A escolha depende do número de solicitações de consulta dos usuários em paralelo e tamanho do banco de dados. Se o cliente preferir usar um balanceador de carga de terceiros, a NetApp recomenda um balanceador de carga dedicado para workloads de demanda de alta performance. Para dimensionamento do StorageGRID, consulte a equipe de conta do NetApp.

Outras recomendações de configuração do StorageGRID incluem:

- **Topologia de grade**. Não misture o SGF6024 com outros modelos de dispositivos de armazenamento no mesmo local da grade. Se você preferir usar o SG6060 para proteção de arquivo de longo prazo, mantenha o SGF6024 com um balanceador de carga de grade dedicado em seu próprio local de grade (local físico ou lógico) para um banco de dados ativo para melhorar o desempenho. Misturar diferentes modelos de aparelho no mesmo local reduz o desempenho geral no local.

- **Proteção de dados**. Use cópias replicadas para proteção. Não use codificação de apagamento para um banco de dados ativo. O cliente pode usar a codificação de apagamento para proteção a longo prazo de bancos de dados inativos.
- **Não ative a compressão da grade**. Vertica compacta objetos antes de armazenar em armazenamento de objetos. Ativar a compressão de grade não economiza ainda mais o uso de armazenamento e reduz significativamente o desempenho DA faixa de bytes.
- * Conexão de endpoint HTTP versus HTTPS S3*. Durante o teste de benchmark, observamos uma melhoria de desempenho de cerca de 5% ao usar uma conexão HTTP S3 do cluster Vertica para o ponto de extremidade do balanceador de carga StorageGRID. Esta escolha deve basear-se nos requisitos de segurança do cliente.

As recomendações para uma configuração Vertica incluem:

- **As configurações padrão do depósito do banco de dados Vertica estão ativadas (valor de 1) para operações de leitura e gravação**. A NetApp recomenda fortemente que essas configurações do depósito estejam ativadas para aprimorar o desempenho.
- **Desativar limitações de streaming**. Para obter detalhes de configuração, consulte a secção Desativação das limitações de streaming.

## Instalação do modo Eon no local com armazenamento comunitário no StorageGRID

As seções a seguir descrevem o procedimento para instalar o modo Eon no local com armazenamento comunitário no StorageGRID. O procedimento para configurar o armazenamento de objetos compatível com o Simple Storage Service (S3) no local é semelhante ao procedimento no guia Vertica, "Instale um banco de dados do modo Eon no local".

A seguinte configuração foi usada para o teste funcional:

- StorageGRID 11.4.0.4
- Vertica 10.1.0
- Três máquinas virtuais (VMs) com CentOS 7.x os para nós Vertica formarem um cluster. Esta configuração é apenas para o teste funcional, não para o cluster de banco de dados de produção Vertica.

Esses três nós são configurados com uma chave Secure Shell (SSH) para permitir SSH sem uma senha entre os nós dentro do cluster.

### Informações necessárias da NetApp StorageGRID

Para instalar o modo Eon no local com armazenamento comunitário no StorageGRID, você deve ter as seguintes informações pré-requisitos.

- Endereço IP ou nome de domínio totalmente qualificado (FQDN) e número da porta do endpoint StorageGRID S3. Se você estiver usando HTTPS, use uma autoridade de certificação personalizada (CA) ou um certificado SSL autoassinado implementado no endpoint do StorageGRID S3.
- Nome do intervalo. Ele deve pré-existir e estar vazio.
- Acesse o ID da chave e a chave de acesso secreta com acesso de leitura e gravação ao bucket.

### Criando um arquivo de autorização para acessar o endpoint S3

Os pré-requisitos a seguir se aplicam ao criar um arquivo de autorização para acessar o endpoint S3:

- Vertica está instalado.
- Um cluster está configurado, configurado e pronto para criação de banco de dados.

Para criar um arquivo de autorização para acessar o endpoint S3, siga estas etapas:

1. Faça login no nó Vertica onde você será executado `admintools` para criar o banco de dados do modo Eon.

   O usuário padrão é `dbadmin`, criado durante a instalação do cluster Vertica.

2. Use um editor de texto para criar um arquivo sob o `/home/dbadmin` diretório. O nome do arquivo pode ser o que você quiser, por exemplo `sg_auth.conf`, .
3. Se o endpoint S3 estiver usando uma porta HTTP 80 padrão ou uma porta HTTPS 443, ignore o número da porta. Para usar HTTPS, defina os seguintes valores:

   ◦ `awsenablehttps = 1`, caso contrário, defina o valor como `0`.

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

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

   Para usar uma CA personalizada ou um certificado SSL autoassinado para a conexão HTTPS de endpoint do StorageGRID S3, especifique o caminho completo do arquivo e o nome do arquivo do certificado. Esse arquivo deve estar no mesmo local em cada nó Vertica e ter permissão de leitura para todos os usuários. Ignore esta etapa se o certificado SSL do StorageGRID S3 for assinado pela CA publicamente conhecida.

   - `awscafile = <filepath/filename>`

   Por exemplo, veja o seguinte arquivo de exemplo:

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

+

> Em um ambiente de produção, o cliente deve implementar um certificado de servidor assinado por uma CA publicamente conhecida em um endpoint do balanceador de carga do StorageGRID S3.

### Escolhendo um caminho de depósito em todos os nós Vertica

Escolha ou crie um diretório em cada nó para o caminho do storage de depósito. O diretório que você fornece para o parâmetro caminho do storage de depósito deve ter o seguinte:

- O mesmo caminho em todos os nós do cluster (por exemplo, `/home/dbadmin/depot`)
- Seja legível e gravável pelo usuário dbadmin

- Armazenamento suficiente

  Por padrão, o Vertica usa 60% do espaço do sistema de arquivos que contém o diretório para armazenamento de depósito. Você pode limitar o tamanho do depósito usando o `--depot-size` argumento no `create_db` comando. "Dimensionamento do seu cluster Vertica para um banco de dados do modo Eon"consulte o artigo para obter diretrizes gerais de dimensionamento Vertica ou consulte o seu gerente de conta Vertica.

  A `admintools create_db` ferramenta tenta criar o caminho do depósito para você se não existir um.

### Criando o banco de dados Eon on-premises

Para criar o banco de dados Eon on-premises, siga estas etapas:

1. Para criar o banco de dados, use a `admintools create_db` ferramenta.

   A lista a seguir fornece uma breve explicação dos argumentos usados neste exemplo. Consulte o documento Vertica para obter uma explicação detalhada de todos os argumentos necessários e opcionais.

   - -x caminho/nome do ficheiro de autorização criado em "Criando um arquivo de autorização para acessar o endpoint S3" >.

     Os detalhes da autorização são armazenados no banco de dados após a criação bem-sucedida. Você pode remover esse arquivo para evitar expor a chave secreta S3.

   - --communal-storage-localização inferior a s3://StorageGRID bucketname>
   - Lista separada por vírgulas de nós Vertica a serem usados para este banco de dados>
   - -d nome do banco de dados a ser criado>
   - a palavra-passe a ser definida para esta nova base de dados>. Por exemplo, veja o seguinte comando de exemplo:

     ```
     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>'
     ```

     A criação de um novo banco de dados leva vários minutos de duração, dependendo do número de nós para o banco de dados. Ao criar banco de dados pela primeira vez, você será solicitado a aceitar o Contrato de Licença.

Por exemplo, veja o seguinte arquivo de autorização de exemplo e `create db` comando:

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

[dbadmin@vertica-vm1 ~]$ admintools -t create_db -x sg_auth.conf
```

```
--communal-storage-location=s3://vertica --depot-path=/home/dbadmin/depot
--shard-count=6 -s vertica-vm1,vertica-vm2,vertica-vm3 -d vmart -p
'xxxxxxxx'
Default depot size in use
Distributing changes to cluster.
    Creating database vmart
    Starting bootstrap node v_vmart_node0007 (10.45.74.19)
    Starting nodes:
        v_vmart_node0007 (10.45.74.19)
    Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (DOWN)
    Node Status: v_vmart_node0007: (UP)
    Creating database nodes
    Creating node v_vmart_node0008 (host 10.45.74.29)
    Creating node v_vmart_node0009 (host 10.45.74.39)
    Generating new configuration information
    Stopping single node db before adding additional nodes.
    Database shutdown complete
    Starting all nodes
Start hosts = ['10.45.74.19', '10.45.74.29', '10.45.74.39']
    Starting nodes:
        v_vmart_node0007 (10.45.74.19)
        v_vmart_node0008 (10.45.74.29)
        v_vmart_node0009 (10.45.74.39)
    Starting Vertica on all nodes. Please wait, databases with a large
catalog may take a while to initialize.
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (DOWN) v_vmart_node0008: (DOWN)
v_vmart_node0009: (DOWN)
    Node Status: v_vmart_node0007: (UP) v_vmart_node0008: (UP)
v_vmart_node0009: (UP)
Creating depot locations for 3 nodes
Communal storage detected: rebalancing shards

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

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

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
|---|---|
| 61 | s3://vertica/051/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a07_0_0.dfs |
| 145 | s3://vertica/2c4/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a3d_0_0.dfs |
| 146 | s3://vertica/33c/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a1d_0_0.dfs |
| 40 | s3://vertica/382/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a31_0_0.dfs |

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
|---|---|
| 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 |
| 33 | s3://vertica/acd/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a29_0_0.dfs |

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
|---|---|
| 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 |
| 8937984 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000002187a/026d63ae9d4a33237bf0e2c2cf2a794a00a000000002187a.tar |

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
| --- | --- |
| 56260608 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000218b2.tar |
| 53947904 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219ba.tar |
| 44932608 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de/026d63ae9d4a33237bf0e2c2cf2a794a00a00000000219de.tar |
| 256306688 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021a6e.tar |
| 8062464 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e34.tar |
| 20024832 | s3://vertica/metadata/VMart/Libraries/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70/026d63ae9d4a33237bf0e2c2cf2a794a00a0000000021e70.tar |
| 10444 | s3://vertica/metadata/VMart/cluster_config.json |
| 823266 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/chkpt_1.cat.gz |
| 254 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c13_13/completed |

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
| --- | --- |
| 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 |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0016/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |

| Tamanho do objeto (byte) | Caminho completo da chave do balde/objeto |
| --- | --- |
| 822922 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz |
| 232 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed |
| 822930 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz |
| 755033 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0017/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |
| 822922 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/chkpt_1.cat.gz |
| 232 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Checkpoints/c14_7/completed |
| 822930 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_14_g7.cat.gz |
| 755033 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/Txnlogs/txn_15_g8.cat |
| 0 | s3://vertica/metadata/VMart/nodes/v_vmart_node0018/Catalog/859703b06a3456d95d0be28575a673/tiered_catalog.cat |

### Desativação das limitações de streaming

Este procedimento é baseado no guia Vertica para outro armazenamento de objetos no local e deve ser

aplicável ao StorageGRID.

1. Depois de criar o banco de dados, desative o `AWSStreamingConnectionPercentage` parâmetro de configuração definindo-o como `0`. Esta configuração é desnecessária para uma instalação no local do modo Eon com armazenamento comunitário. Este parâmetro de configuração controla o número de conexões ao armazenamento de objetos que o Vertica usa para leituras de streaming. Em um ambiente de nuvem, essa configuração ajuda a evitar que os dados de streaming do armazenamento de objetos usem todas as alças de arquivo disponíveis. Ele deixa algumas alças de arquivo disponíveis para outras operações de armazenamento de objetos. Devido à baixa latência de armazenamentos de objetos no local, essa opção é desnecessária.
2. Use uma `vsql` instrução para atualizar o valor do parâmetro. A senha é a senha do banco de dados que você definiu em "criando o banco de dados on-premises Eon". Por exemplo, veja a seguinte saída de amostra:

```
[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
```

#### Verificando as configurações do depósito

As configurações padrão de depósito do banco de dados Vertica são ativadas (valor de 1) para operações de leitura e gravação. A NetApp recomenda fortemente que essas configurações do depósito estejam ativadas para aprimorar o desempenho.

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

#### Carregamento de dados de amostra (opcional)

Se este banco de dados for para teste e será removido, você pode carregar dados de amostra para este banco de dados para teste. O Vertica vem com um conjunto de dados de amostra, VMart, encontrado em `/opt/vertica/examples/VMart_Schema/` cada nó Vertica. Você pode encontrar mais informações sobre este conjunto de "aqui"dados de amostra .

Siga estes passos para carregar os dados de amostra:

1. Faça login como dbadmin em um dos nós Vertica: cd /opt/vertica/examples/VMart_Schema/
2. Carregue dados de amostra para o banco de dados e insira a senha do banco de dados quando solicitado nas subetapas c e d:
   a. `cd /opt/vertica/examples/VMart_Schema`

b. `./vmart_gen`

c. `vsql < vmart_define_schema.sql`

d. `vsql < vmart_load_data.sql`

3. Existem várias consultas SQL predefinidas, você pode executar algumas delas para confirmar que os dados de teste são carregados com sucesso no banco de dados. Por exemplo: `vsql < vmart_queries1.sql`

## Onde encontrar informações adicionais

Para saber mais sobre as informações descritas neste documento, consulte os seguintes documentos e/ou sites:

- "Documentação do produto NetApp StorageGRID 11,7"
- "Folha de dados do StorageGRID"
- "Documentação do produto Vertica 10,1"

## Histórico de versões

| Versão | Data | Histórico de versões do documento |
| --- | --- | --- |
| Versão 1,0 | Setembro de 2021 | Lançamento inicial. |

*Por Angela Cheng*

# Análises de log do StorageGRID usando o ELK stack

*Por Angela Cheng*

Com o recurso de encaminhamento de syslog do StorageGRID, você pode configurar um servidor syslog externo para coletar e analisar mensagens de log do StorageGRID. ELK (Elasticsearch, Logstash, Kibana) tornou-se uma das soluções de análise de logs mais populares. Assista ao "Análise de log do StorageGRID usando o vídeo ELK" para exibir uma configuração DO ELK de exemplo e como ele pode ser usado para identificar e solucionar problemas de solicitações S3 com falha. O StorageGRID 11,9 suporta a exportação de log de acesso de endpoint do balanceador de carga para o servidor syslog externo. Assista a isso "Vídeo do YouTube" para saber mais sobre esse novo recurso. este artigo fornece arquivos de exemplo de configuração do Logstash, consultas do Kibana, gráficos e painel para dar a você um início rápido para o gerenciamento e análise de logs do StorageGRID.

## Requisitos

- StorageGRID 11.6.0.2 ou superior
- ELK (Elasticsearch, Logstash e Kibana) 7,1x ou superior instalado e em operação

## Arquivos de exemplo

- "Faça o download do pacote de arquivos de amostra Logstash 7.x" **md5 checksum** 148c23d0021d9a4bb4a6c0287464deab e **sha256 checksum** f51ec9e2e3f842d5a7861566b167a561b4373038b4e7bb3c8b8be3d522adf2d6

- "Faça o download do pacote de arquivos de amostra Logstash 8.x" **md5 checksum** e11bae3a662f87c310ef363d0fe06835 e **sha256 checksum** 5c670755742cfdfd5aa723a596ba087e0153a65bcaef3934afdb682f6cd278d
- "Faça o download do pacote de arquivos de amostra Logstash 8.x para o StorageGRID 11,9" **md5 checksum** 41272857c4a54600f95995f6ed74800d e **sha256 checksum** 67048e8661052719990851e1ad960d4902fe537a6e135e8600177188da677c9

## Suposição

Os leitores estão familiarizados com a terminologia e operações do StorageGRID e ELK.

## Instrução

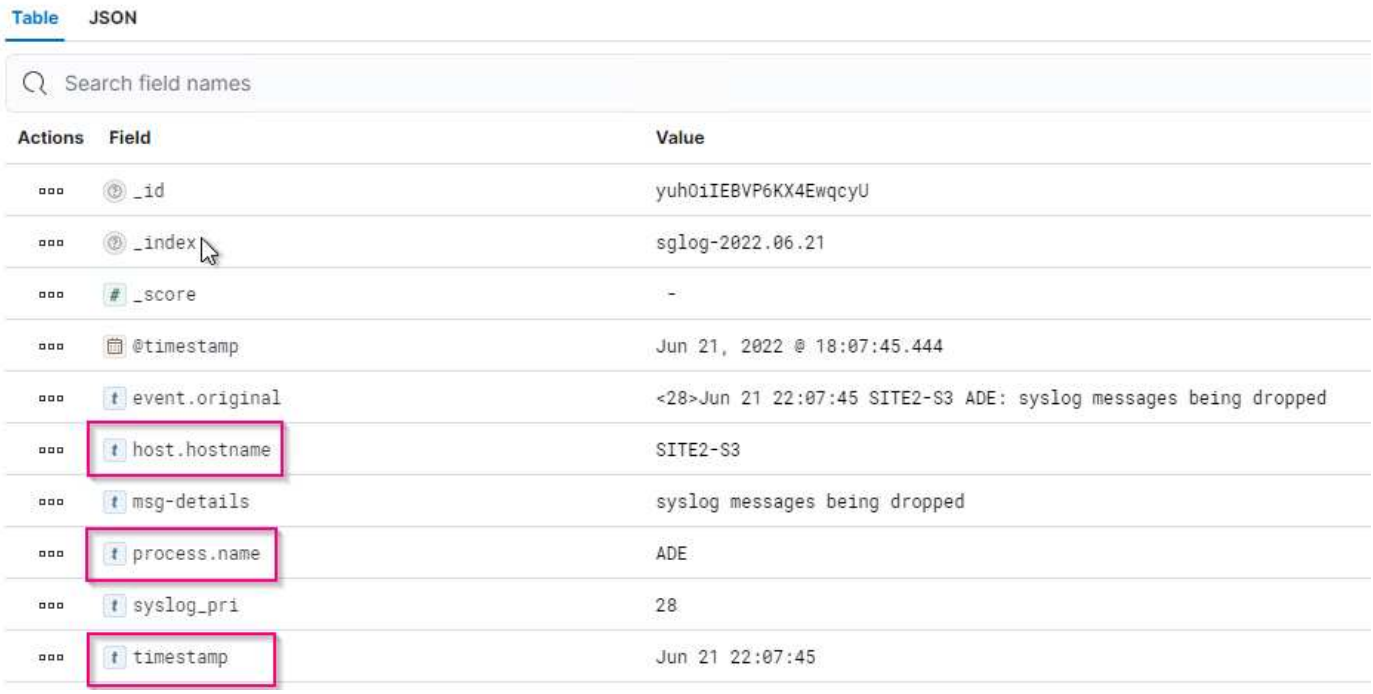
Duas versões de amostra são fornecidas devido a diferenças nos nomes definidos pelos padrões de grok. Por exemplo, o padrão de grok SYSLOGBASE no arquivo de configuração Logstash define nomes de campo de forma diferente, dependendo da versão instalada do Logstash.

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

- Logstash 7,17 amostra*

| 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 amostra*




### Passos

1. Descompacte a amostra fornecida com base na versão ELK instalada. * Sglog-2-file.conf:* este arquivo de configuração envia mensagens de log do StorageGRID para um arquivo no Logstash sem transformação de dados. Você pode usar isso para confirmar que o Logstash está recebendo mensagens do StorageGRID ou para ajudar a entender os padrões de log do StorageGRID. **Sglog-2-es.conf:** este arquivo de configuração transforma mensagens de log do StorageGRID usando vários padrões e filtros. Ele inclui exemplos de instruções drop, que deixam cair mensagens com base em padrões ou filtros. A saída é enviada ao Elasticsearch para indexação. Personalize o arquivo de configuração selecionado de acordo com a instrução dentro do arquivo.
2. Teste o arquivo de configuração personalizado:

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

   Se a última linha retornada for semelhante à linha abaixo, o arquivo de configuração não tem erros de sintaxe:

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

3. Copie o arquivo conf personalizado para a configuração do servidor Logstash: /Etc/logstash/conf.d se você não tiver habilitado o config.reload.automatic em /etc/logstash/logstash.yml, reinicie o serviço Logstash. Caso contrário, aguarde até que o intervalo de recarga da configuração passe.

```
 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
```

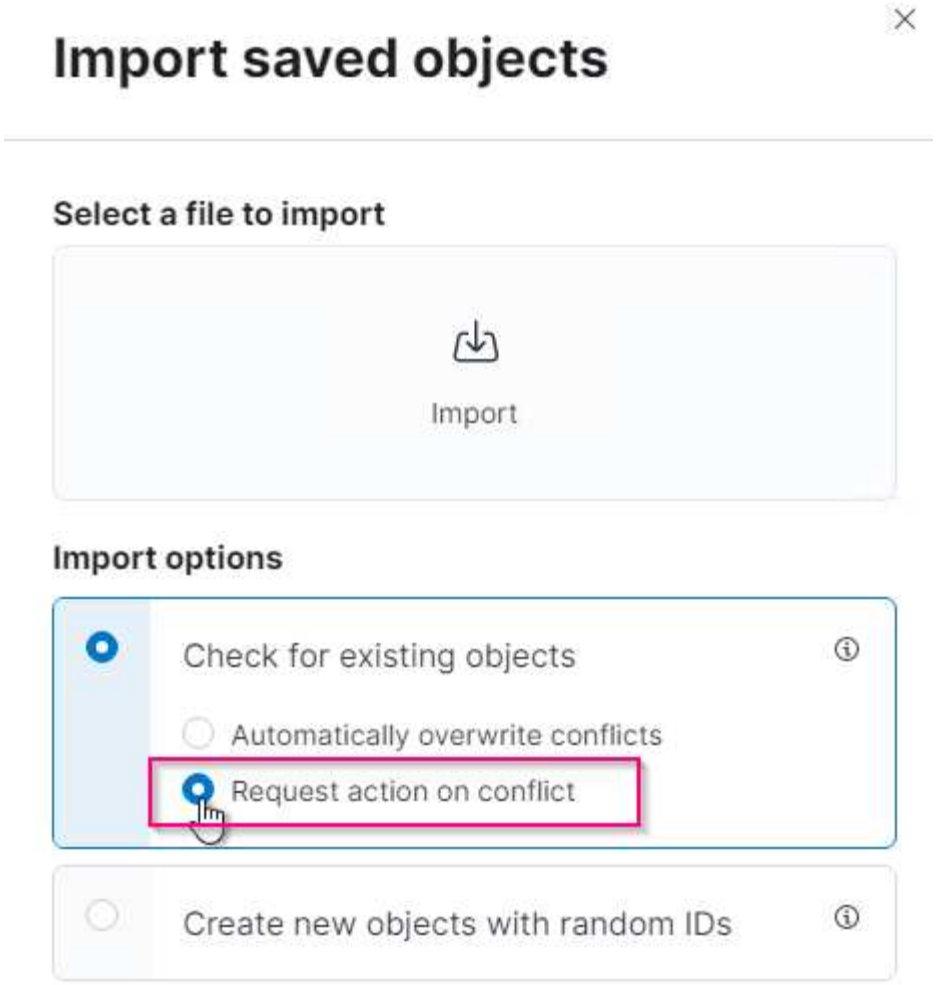
4. Verifique /var/log/logstash/logstash-plain.log e confirme que não há erros iniciando o Logstash com o novo arquivo de configuração.
5. Confirme se a porta TCP foi iniciada e escutada. Neste exemplo, a porta TCP 5000 é usada.

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

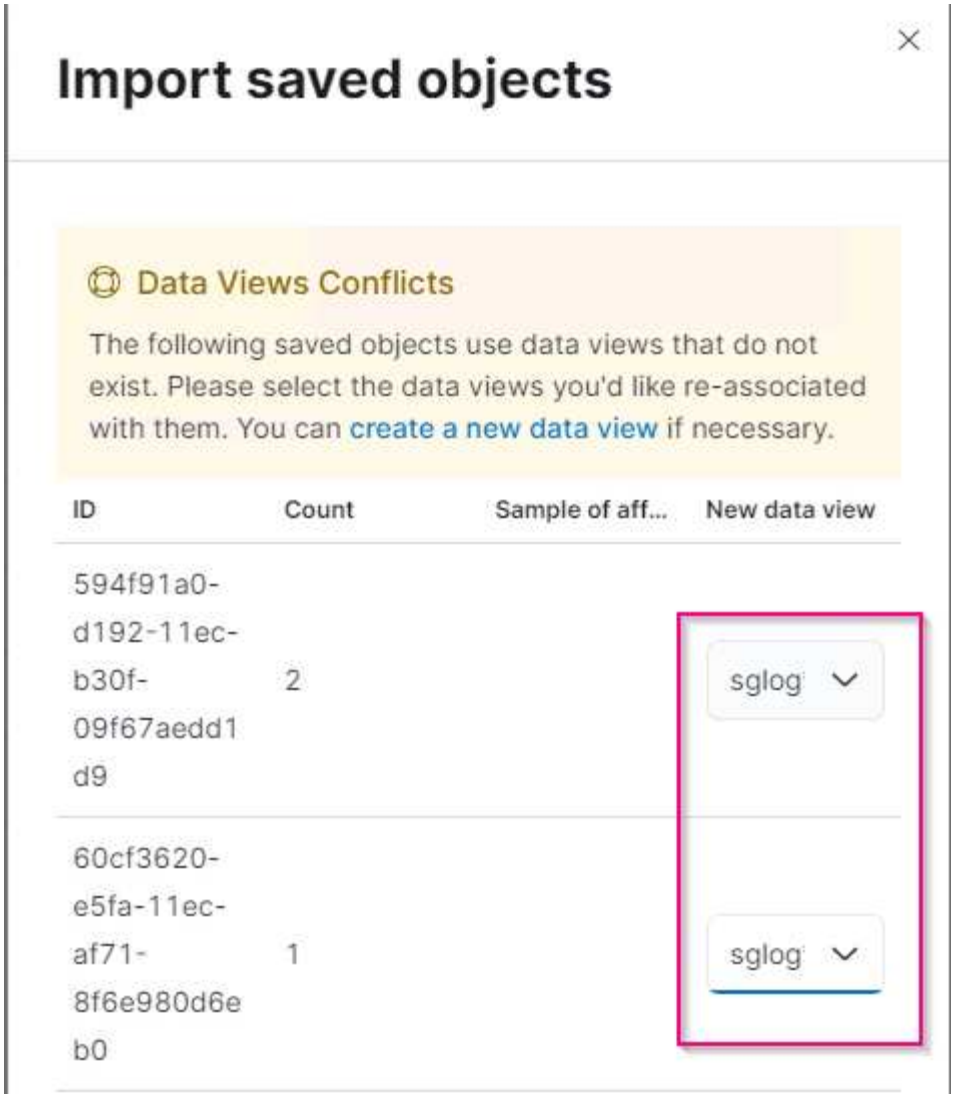
6. A partir da GUI do gerenciador do StorageGRID, configure o servidor syslog externo para enviar mensagens de log para o Logstash. Consulte "vídeo de demonstração" para obter mais informações.
7. Você precisa configurar ou desativar o firewall no servidor Logstash para permitir a conexão de nós StorageGRID à porta TCP definida.
8. Na GUI do Kibana, selecione Gerenciamento → Ferramentas de desenvolvimento. Na página Console, execute este comando GET para confirmar que novos índices são criados no Elasticsearch.

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

9. A partir do Kibana GUI, crie um padrão de índice (ELK 7.x) ou visualização de dados (ELK 8.x).
10. Na GUI do Kibana, digite 'objetos salvos' na caixa de pesquisa que está localizada no centro superior. Na página objetos salvos, selecione Importar. Em Opções de importação, selecione "solicitar ação em conflito"




Importe elk<version>-query-chart-sample.ndjson. Quando solicitado a resolver o conflito, selecione o padrão de índice ou a exibição de dados que você criou na etapa 8.




Os seguintes objetos Kibana são importados: * Consulta * * * auditoria-msg-s3rq S3-orlm * Registro de bycast S3 S3 mensagens relacionadas com * aviso de nível de log ou acima * evento de segurança com falha * nginx-gw Registro de acesso de endpoint (disponível apenas em elk8-sample-for-sg119.zip) * Gráfico * S3 pedidos contam com base em bycast.log * Código de status HTTP *

Agora você está pronto para executar a análise de log do StorageGRID usando o Kibana.

## Recursos adicionais

- "syslog101"
- "O que é a pilha ELK"
- "Lista de padrões Grok"
- "Um guia para iniciantes para Logstash: Grok"
- "Um guia prático para o Logstash: Syslog Deep Dive"
- "Guia Kibana – explore o documento"
- "Referência de mensagens de log de auditoria do StorageGRID"

# Use Prometheus e Grafana para estender a retenção de métricas

*Por Aron Klein*

Este relatório técnico fornece instruções detalhadas para configurar o NetApp StorageGRID com serviços externos Prometheus e Grafana.

## Introdução

O StorageGRID armazena métricas usando Prometheus e fornece visualizações dessas métricas por meio de dashboards Grafana integrados. As métricas Prometheus podem ser acessadas com segurança a partir do StorageGRID configurando certificados de acesso de cliente e habilitando o acesso prometheus para o cliente especificado. Hoje, a retenção desses dados métricos é limitada pela capacidade de storage do nó de administração. Para obter uma duração mais longa e uma capacidade de criar visualizações personalizadas dessas métricas, implantaremos um novo servidor Prometheus e Grafana, configuraremos nosso novo servidor para raspar as métricas da instância StorageGRIDs e construir um painel com as métricas que são importantes para nós. Você pode obter mais informações sobre as métricas do Prometheus coletadas no "Documentação do StorageGRID".

## Federado Prometheus

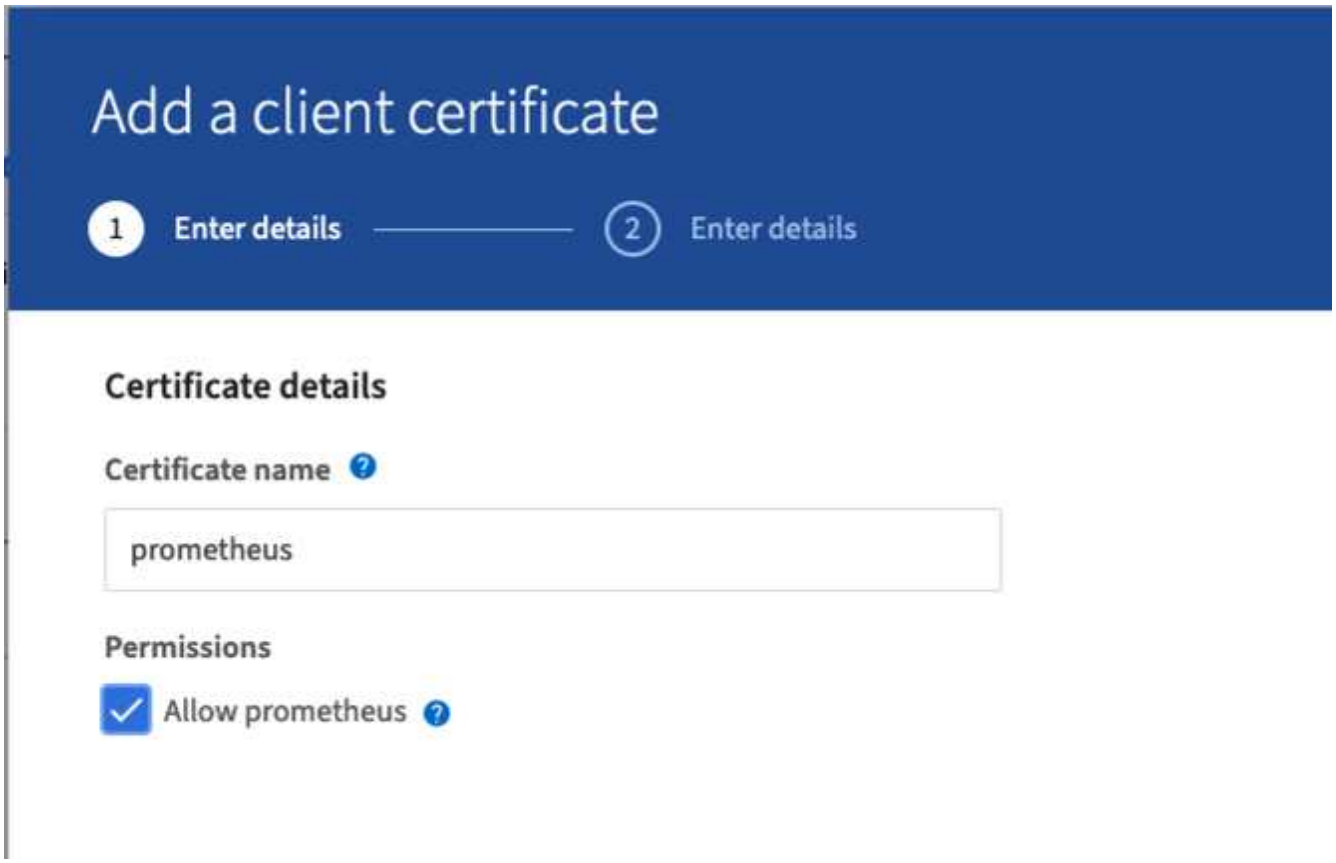
### Detalhes do laboratório

Para os propósitos deste exemplo, eu vou usar todas as máquinas virtuais para nós StorageGRID 11,6 e um servidor Debian 11. A interface de gerenciamento do StorageGRID é configurada com um certificado de CA publicamente confiável. Este exemplo não passará pela instalação e configuração do sistema StorageGRID ou instalação do Debian linux. Você pode usar qualquer versão do Linux que desejar que seja suportada por Prometheus e Grafana. Tanto o Prometheus quanto o Grafana podem instalar como contentores docker, compilar a partir de fontes ou binários pré-compilados. Neste exemplo eu estarei instalando ambos binários Prometheus e Grafana diretamente no mesmo servidor Debian. Faça o download e siga as instruções básicas de instalação https://prometheus.io de e https://grafana.com/grafana/, respetivamente.

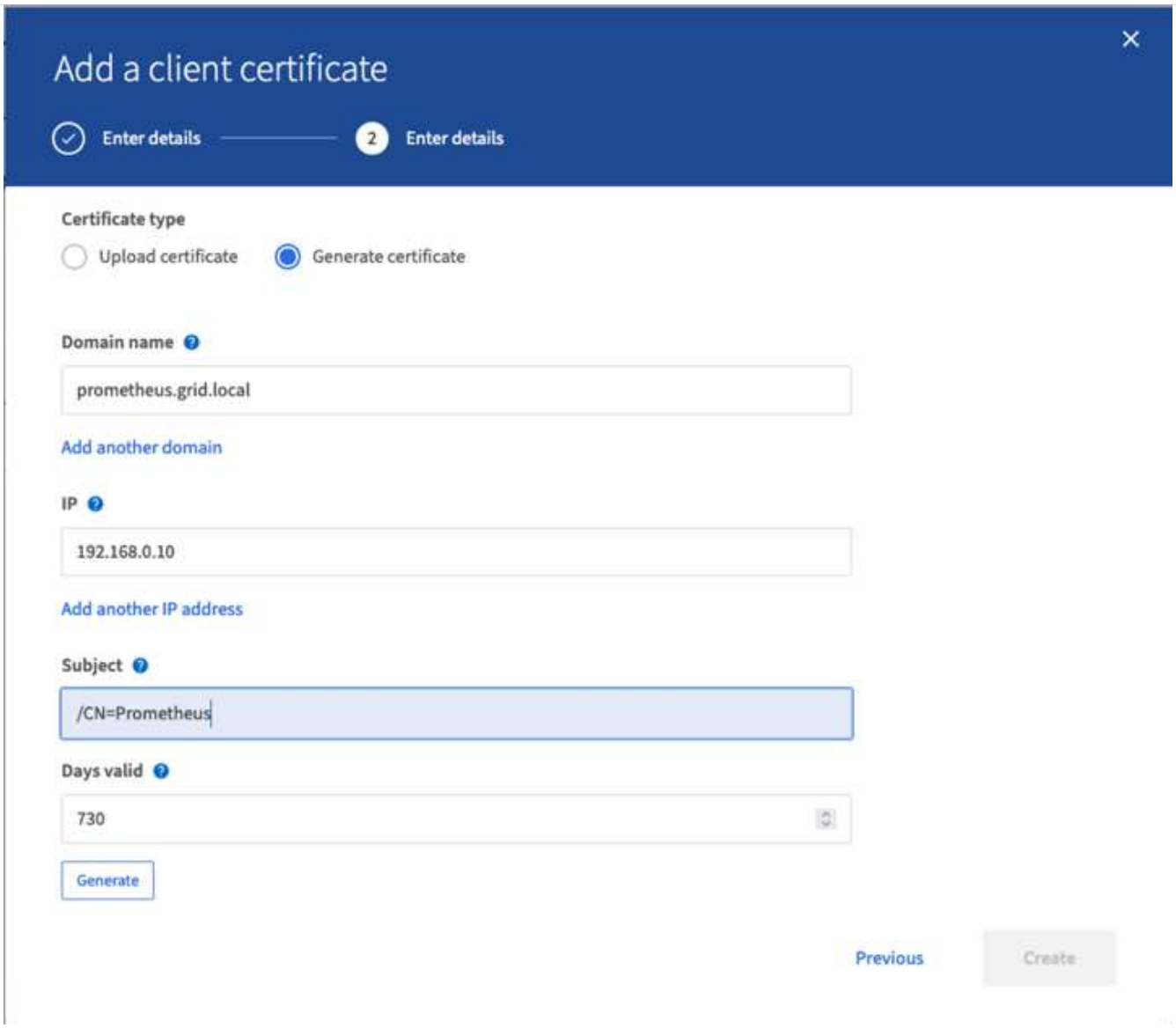
### Configurar o StorageGRID para acesso ao cliente Prometheus

Para obter acesso às métricas do prometheus armazenadas pelo StorageGRIDs, você deve gerar ou carregar um certificado de cliente com chave privada e habilitar a permissão para o cliente. A interface de gerenciamento do StorageGRID deve ter um certificado SSL. Esse certificado deve ser confiável pelo servidor prometheus por uma CA confiável ou manualmente confiável se ele for autoassinado. Para ler mais, visite o "Documentação do StorageGRID".

1. Na interface de gerenciamento do StorageGRID, selecione "CONFIGURAÇÃO" no lado inferior esquerdo e, na segunda coluna em "Segurança", clique em certificados.
2. Na página certificados, selecione a guia "Cliente" e clique no botão "Adicionar".
3. Forneça um nome para o cliente que será concedido acesso e use este certificado. Clique na caixa em "permissões", na frente de "permitir Prometheus" e clique no botão continuar.

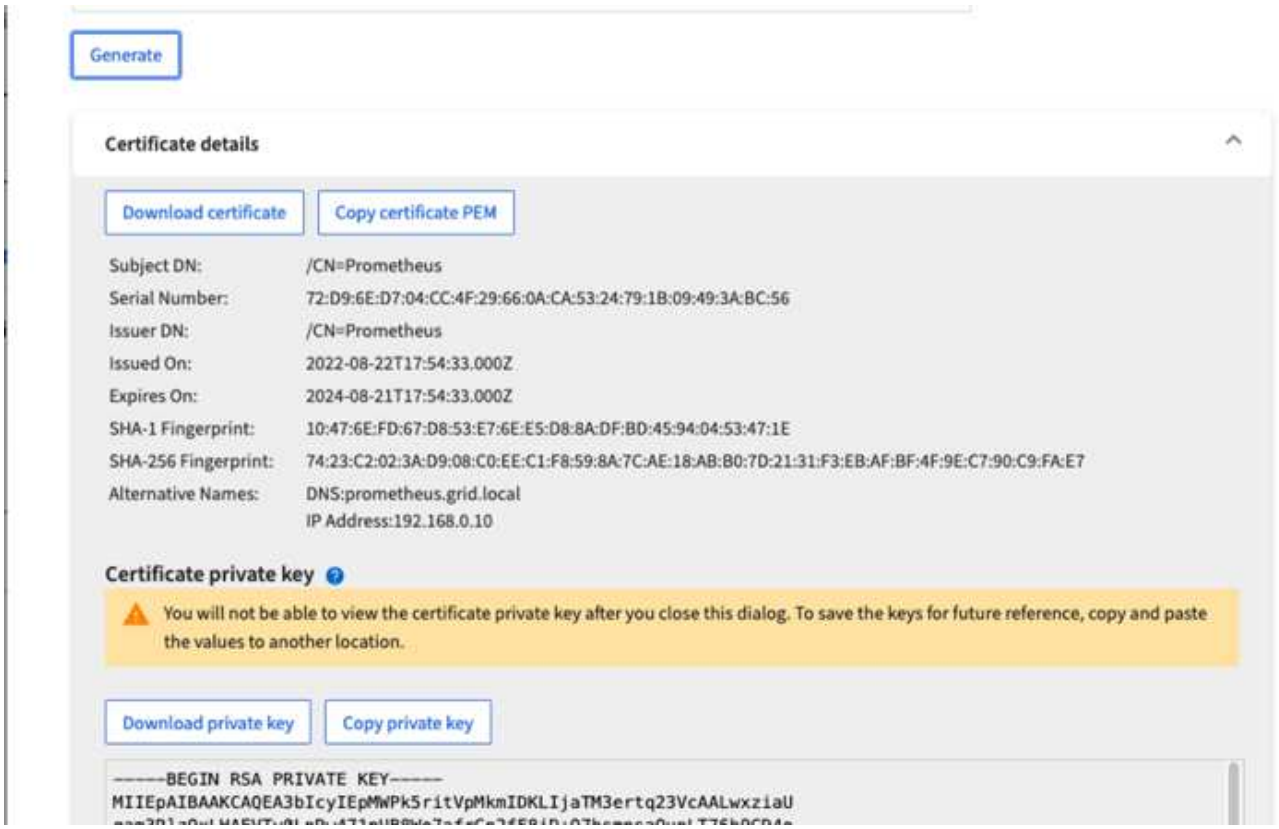



4. Se você tiver um certificado assinado pela CA, você pode selecionar o botão de opção "carregar certificado", mas, no nosso caso, vamos permitir que o StorageGRID gere o certificado do cliente selecionando o botão de opção "gerar certificado". Os campos obrigatórios serão exibidos para serem preenchidos. Insira o FQDN para o servidor cliente, o IP do servidor, o assunto e dias válidos. Em seguida, clique no botão "gerar".




> 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. Baixe o arquivo pem de certificado e o arquivo pem de chave privada.




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

**Prepare o servidor Linux para a instalação do Prometheus**

Antes de instalar o Prometheus, eu quero preparar meu ambiente com um usuário Prometheus, a estrutura de diretórios e configurar a capacidade para o local de armazenamento de métricas.

1. Crie o usuário Prometheus.

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

2. Crie os diretórios para Prometheus, certificado de cliente e dados de métricas.

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

3. Formatei o disco que estou usando para retenção de métricas com um sistema de arquivos ext4.

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

4. Eu então montei o sistema de arquivos para o diretório de métricas do Prometheus.

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

5. Obtenha o uuid do disco que você está usando para seus dados de métricas.

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

6. Adicionando uma entrada em /etc/fstab/ fazendo com que a montagem persista em reinicializações usando o uuid de /dev/sdb.

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

**Instale e configure Prometheus**

Agora que o servidor está pronto, posso iniciar a instalação do Prometheus e configurar o serviço.

1. Extraia o pacote de instalação do Prometheus

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

2. Copie os binários para /usr/local/bin e altere a propriedade para o usuário prometheus criado anteriormente

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

3. Copie os consoles e bibliotecas para /etc/prometheus

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

4. Copie o certificado do cliente e os arquivos pem de chave privada baixados anteriormente do StorageGRID para /etc/prometheus/certs
5. Crie o arquivo yaml de configuração prometheus

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

6. Insira a seguinte configuração. O nome do trabalho pode ser qualquer coisa que você desejar. Altere o "-targets: ["]" para o FQDN do nó admin e, se você alterou os nomes dos nomes dos arquivos de certificado e chave privada, atualize a seção tls_config para corresponder. Em seguida, salve o arquivo. Se sua interface de gerenciamento de grade estiver usando um certificado autoassinado, baixe o certificado e coloque-o com o certificado de cliente com um nome exclusivo, e na seção tls_config adicione CA_file: /Etc/prometheus/cert/UIcert.pem

   a. Neste exemplo, estou coletando todas as métricas que começam com alertmanager, cassandra, node e StorageGRID. Você pode ver mais informações sobre as métricas do Prometheus no "Documentação do 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']
```

Se a interface de gerenciamento de grade estiver usando um certificado autoassinado, baixe o certificado e coloque-o com o certificado do cliente com um nome exclusivo. Na seção tls_config adicione o certificado acima do certificado do cliente e das linhas de chave privada

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

1. Altere a propriedade de todos os arquivos e diretórios em /etc/prometheus e /var/lib/prometheus para o usuário prometheus

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

2. Crie um arquivo de serviço prometheus em /etc/systemd/system

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

3. Insira as linhas a seguir, observe o número—storage.tsdb.retension.time.1y que define a retenção dos dados métricos para 1 ano. Como alternativa, você pode usar 300GiB para basear a retenção nos limites de armazenamento. Este é o único local para definir a retenção de métricas.

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

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

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

4. Recarregue o serviço systemd para Registrar o novo serviço prometheus. Em seguida, inicie e ative o serviço prometheus.

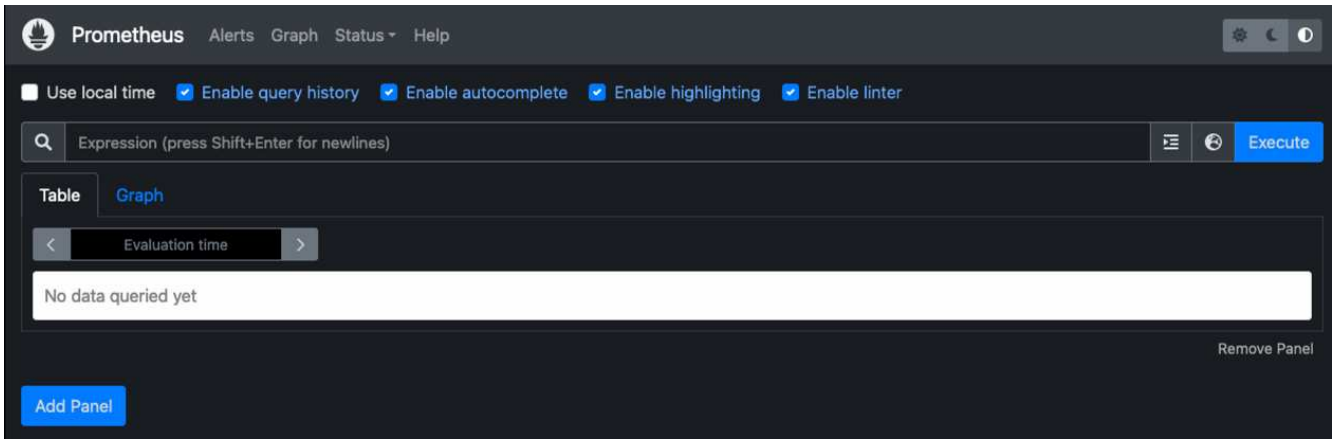
```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Verifique se o serviço está funcionando corretamente

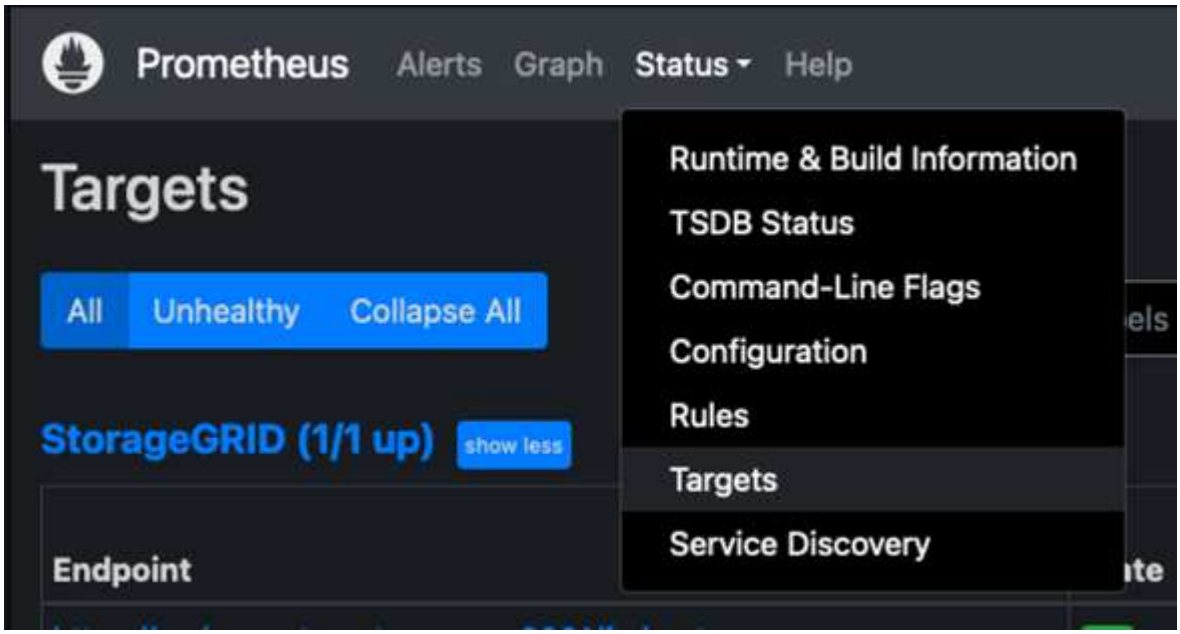
```
sudo systemctl status prometheus
```

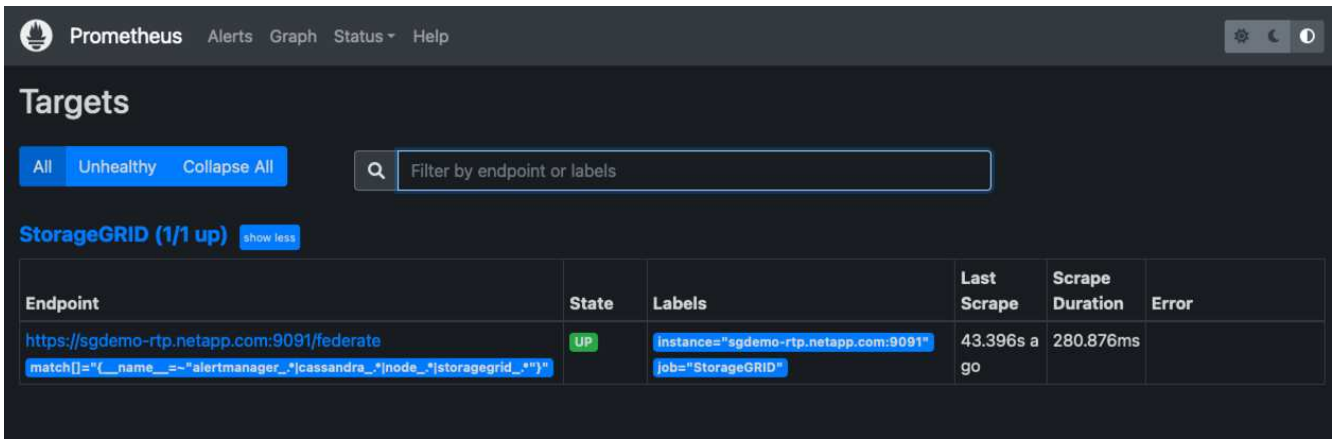
```
● prometheus.service - Prometheus Time Series Collection and Processing
Server
     Loaded: loaded (/etc/systemd/system/prometheus.service; enabled;
vendor preset: enabled)
     Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago
   Main PID: 6498 (prometheus)
      Tasks: 13 (limit: 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. Agora você deve ser capaz de navegar até a IU do seu servidor prometheus http://Prometheus-server:9090 e ver a IU

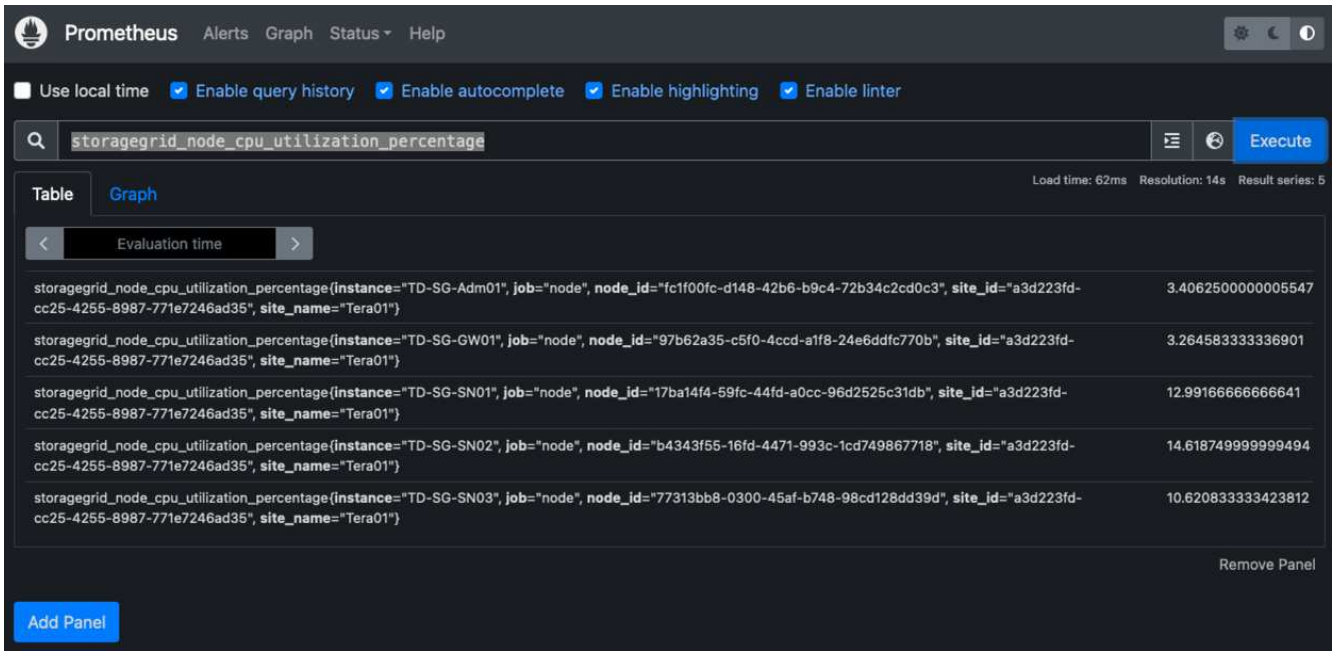



7. Em destinos "Status", você pode ver o status do endpoint do StorageGRID que configuramos em prometheus.yml




8. Na página Gráfico, você pode executar uma consulta de teste e verificar se os dados estão sendo raspados com sucesso. Por exemplo, digite "StorageGRID_node_cpu_utilization_percentage" na barra de consulta e clique no botão Executar.




## Instale e configure o Grafana

Agora que o prometheus está instalado e funcionando, podemos passar para a instalação do Grafana e configurar um dashboard

### Instalação do Grafana

1. Instale a mais recente edição corporativa do 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. Adicione este repositório para versões estáveis:

```
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. Depois de adicionar o repositório.

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

4. Recarregue o serviço systemd para Registrar o novo serviço grafana. Em seguida, inicie e ative o serviço Grafana.

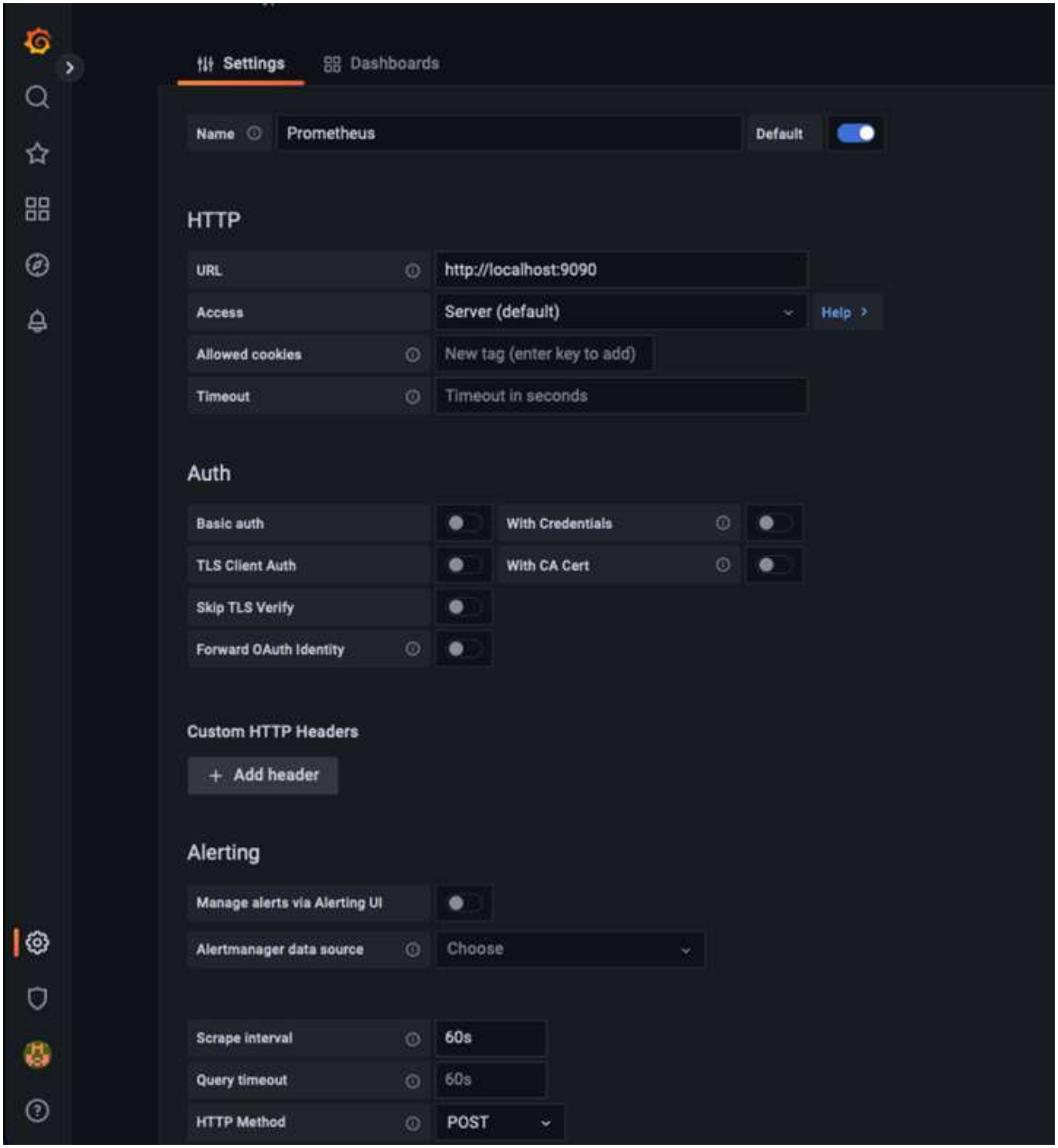
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. O Grafana agora está instalado e em execução. Quando você abre um navegador para HTTP://Prometheus-server:3000 você será recebido com a página de login do Grafana.
6. As credenciais de login padrão são admin/admin, e você deve definir uma nova senha como ela solicita.

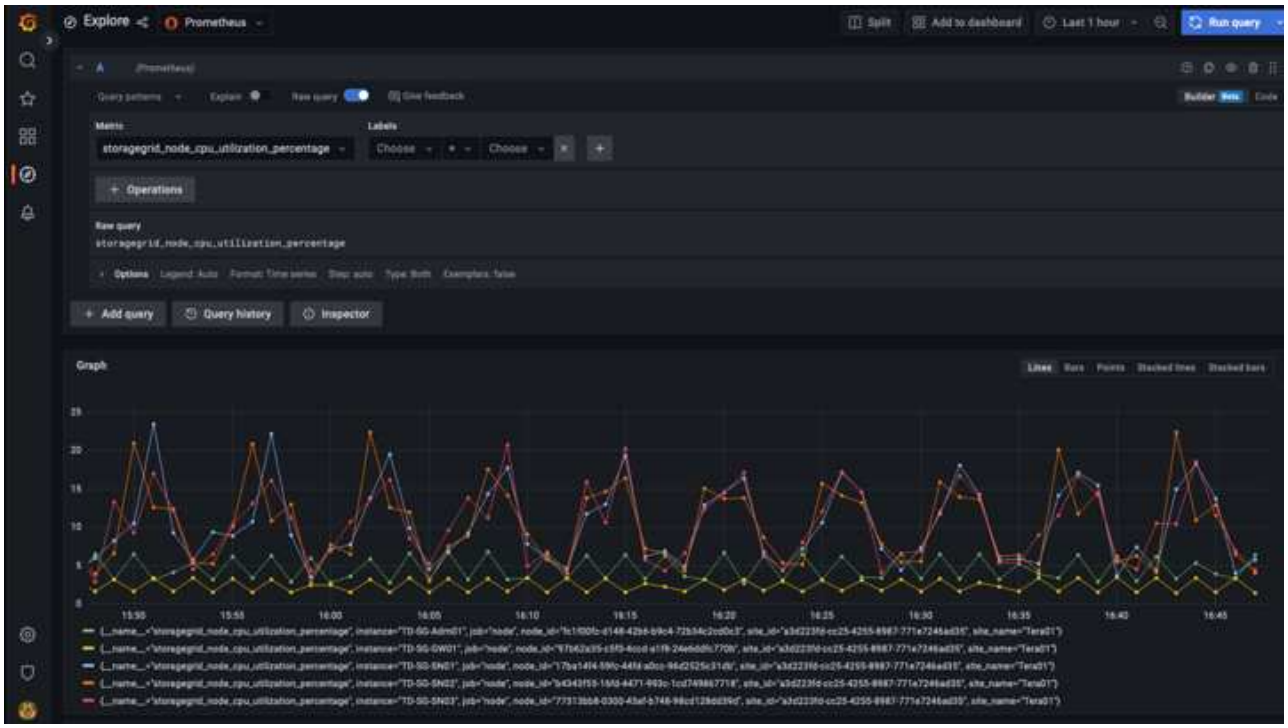
**Crie um painel Grafana para StorageGRID**

Com Grafana e Prometheus instalados e em execução, agora é hora de conetar os dois criando uma fonte de dados e construindo um painel

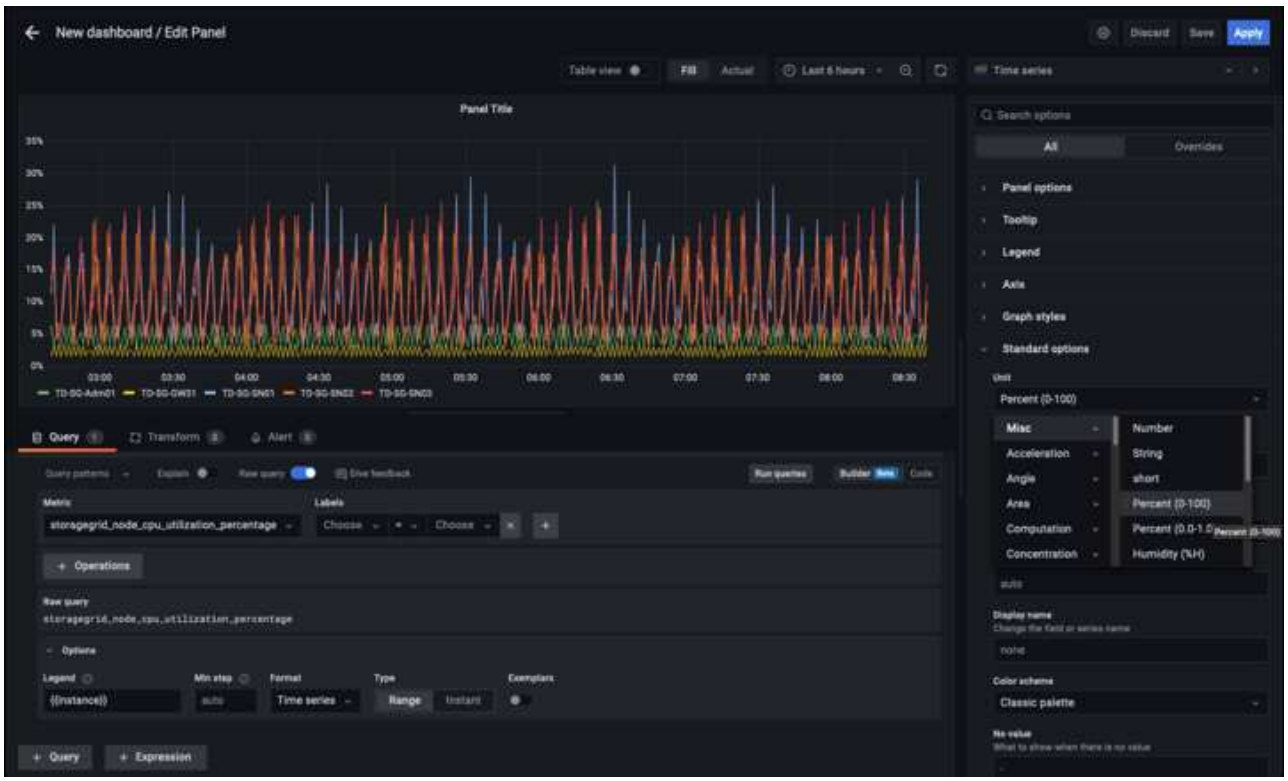
1. No painel esquerdo, expanda "Configuration" (Configuração) e selecione "Data Sources" (fontes de dados) e, em seguida, clique no botão "Add Data source" (Adicionar fonte de dados)
2. Prometheus será uma das principais fontes de dados para escolher. Se não estiver, use a barra de pesquisa para localizar "Prometheus"
3. Configure a fonte Prometheus inserindo o URL da instância prometheus e o intervalo de raspagem para corresponder ao intervalo Prometheus. Eu também desabilitei a seção de alerta, pois não configurei o gerenciador de alertas no prometheus.




4. Com as configurações desejadas inseridas, role para baixo até a parte inferior e clique em "Salvar e testar"
5. Depois que o teste de configuração for bem-sucedido, clique no botão explorar.
    a. Na janela explorar você pode usar a mesma métrica que testamos Prometheus com "StorageGRID_node_cpu_utilization_percentage" e clicar no botão "Executar consulta"



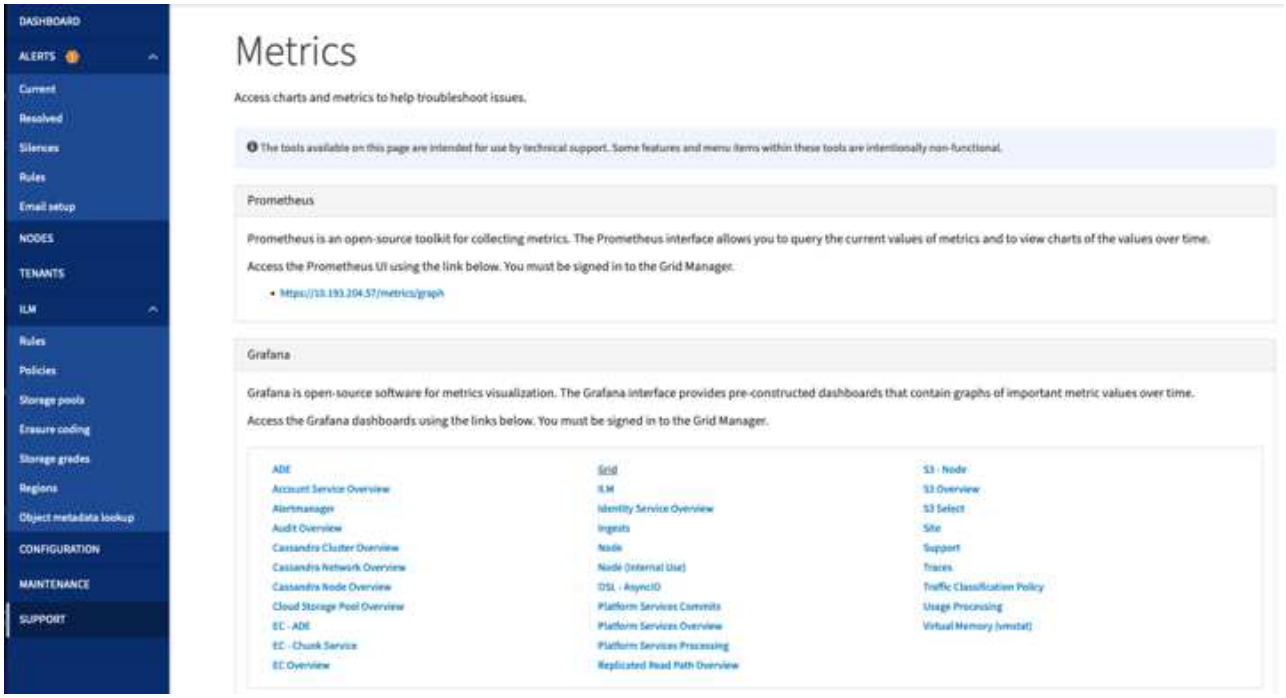

6. Agora que temos a fonte de dados configurada, podemos criar um dashboard.
   a. No painel esquerdo, expanda "Dashboards" e selecione "novo painel"
   b. Selecione "Adicionar um novo painel"
   c. Configure o novo painel selecionando uma métrica, novamente vou usar "StorageGRID_node_cpu_utilization_percentage", digite um título para o painel, expanda "Opções" na parte inferior e para a legenda mudar para personalizado e digite "_instância" para definir os nomes dos nós" e no painel direito em "Opções padrão" defina "Unidade" para "Misc/percent(0-100)". Em seguida, clique em "aplicar" para salvar o painel no painel.



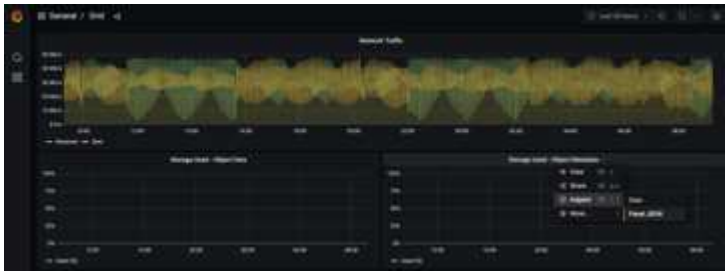
7. Poderíamos continuar a construir nosso painel como esse para cada métrica que quisermos, mas felizmente o StorageGRID já tem painéis com painéis que podemos copiar em nossos painéis personalizados.

   a. No painel esquerdo da interface de gerenciamento do StorageGRID, selecione "suporte" e, na parte inferior da coluna "Ferramentas", clique em "métricas".

   b. Dentro das métricas, vou selecionar o link "Grid" na parte superior da coluna do meio.

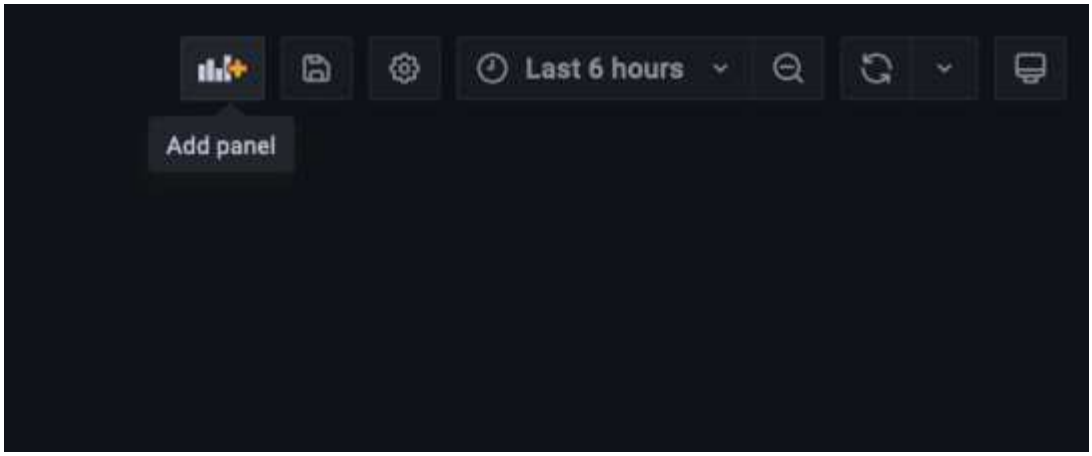



   c. No painel Grid, permite selecionar o painel "Storage Used - Object Metadata" (armazenamento usado - metadados de objetos). Clique na pequena seta para baixo e no final do título do painel para soltar um menu. Neste menu, selecione "Inspecionar" e "Painel JSON".



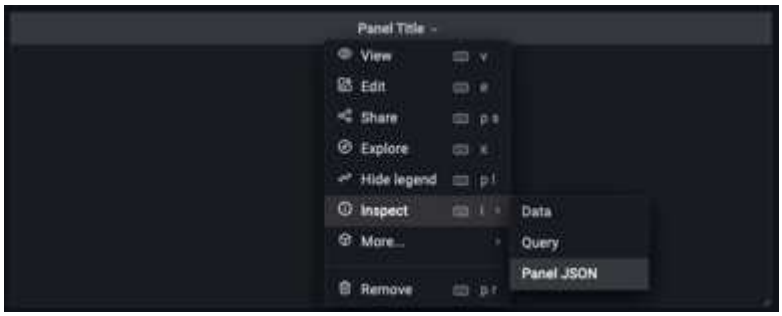
   d. Copie o código JSON e feche a janela.


e. No nosso novo painel, clique no ícone para adicionar um novo painel.

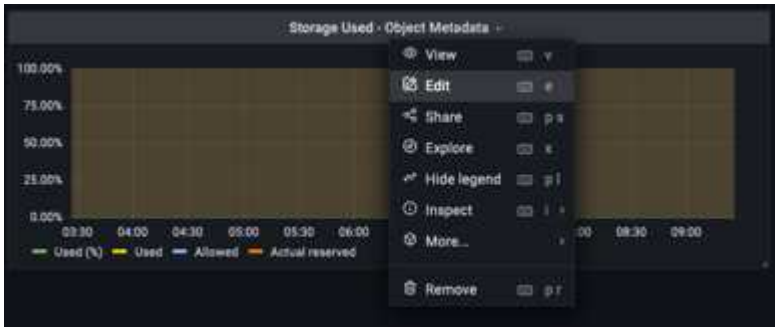


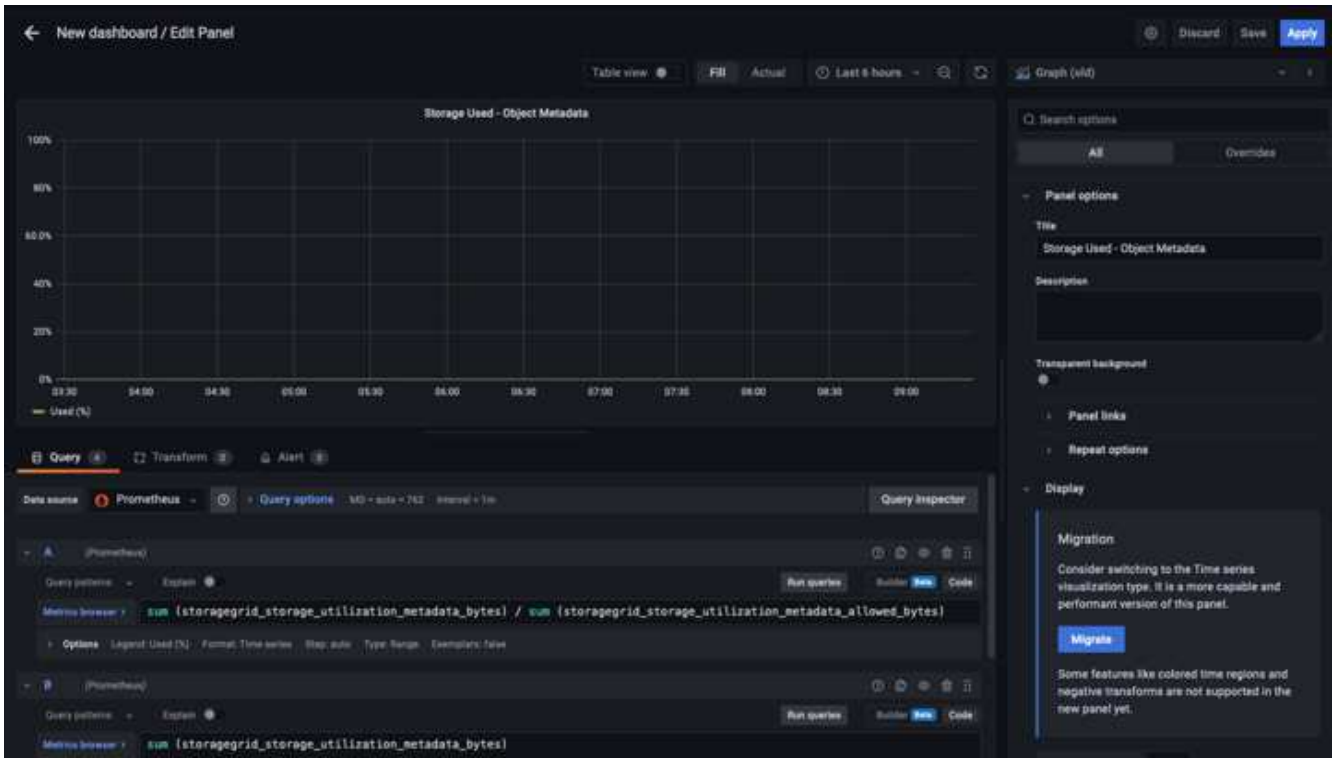

f. Aplique o novo painel sem fazer alterações

g. Assim como no painel StorageGRID, inspecione o JSON. Remova todo o código JSON e substitua-o pelo código copiado do painel StorageGRID.




h. Edite o novo painel e, no lado direito, você verá uma mensagem de migração com um botão "migrar". Clique no botão e, em seguida, clique no botão "aplicar".






8. Depois de ter todos os painéis no lugar e configurado como quiser. Salve o painel clicando no ícone do disco no canto superior direito e dê um nome ao painel.

### Conclusão

Agora temos um servidor Prometheus com capacidade de armazenamento e retenção de dados personalizáveis. Com isso, podemos continuar construindo nossos próprios painéis com as métricas mais relevantes para nossas operações. Você pode obter mais informações sobre as métricas do Prometheus coletadas no "Documentação do StorageGRID".

# Use o DNS da F5 para balancear a carga globalmente no StorageGRID.

*Por Steve Gorman (F5)*

Este relatório técnico fornece instruções detalhadas para configurar o NetApp StorageGRID com os serviços DNS da F5 para balanceamento de carga global, visando oferecer melhor disponibilidade de dados, maior consistência de dados e otimizar o roteamento de transações do S3 quando sua grade estiver distribuída em vários sites e/ou grupos de alta disponibilidade.

## Introdução

A solução F5 BIG-IP DNS, anteriormente chamada de BIG-IP GTM (Global Traffic Manager) e informalmente conhecida como GSLB (Global Server Load Balancing), permite o acesso contínuo em vários grupos HA ativo-ativo e a implementação eficaz de soluções StorageGRID multi-site ativo-ativo.

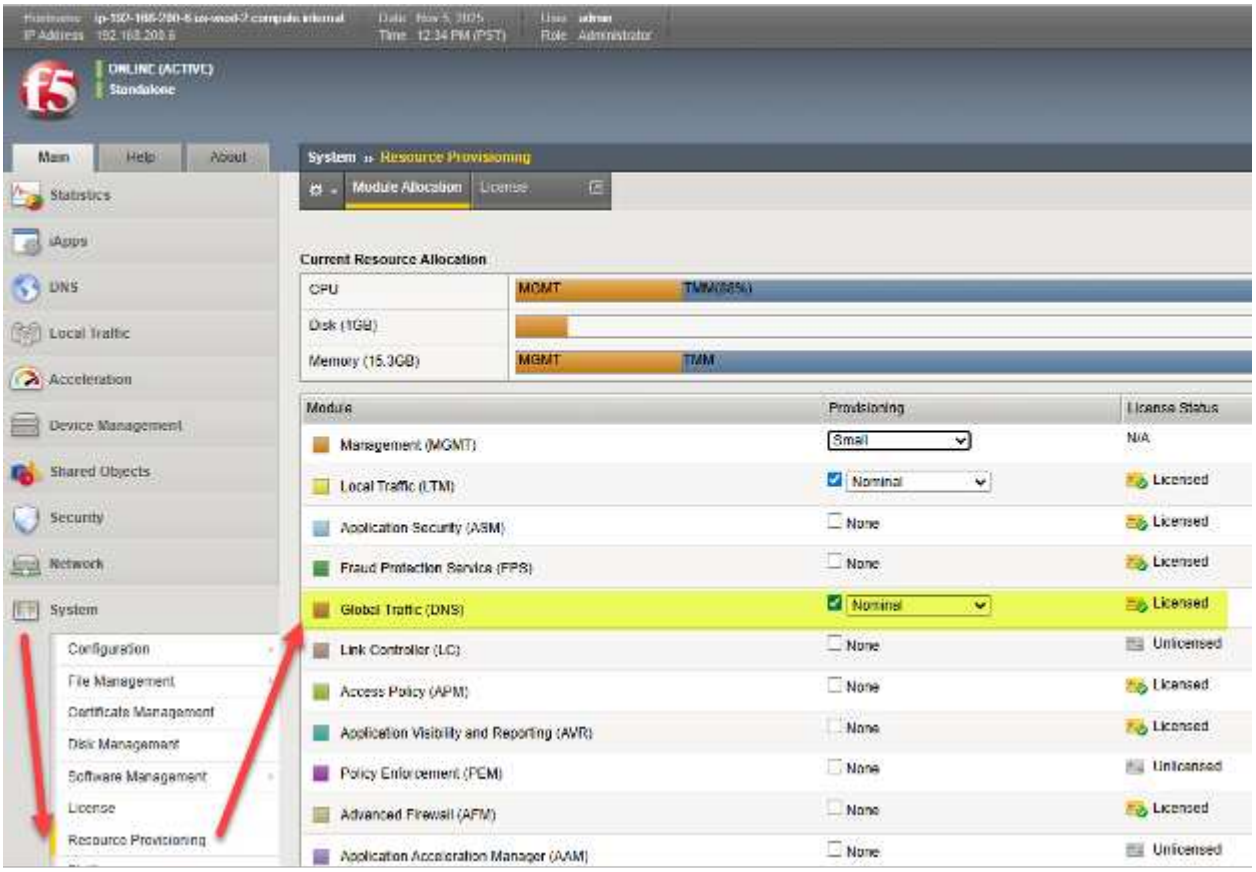
## Configuração F5 BIG-IP StorageGRID em vários locais

Independentemente do número de sites StorageGRID a serem suportados, no mínimo dois dispositivos BIG-IP, físicos ou virtuais, devem ter o módulo BIG-IP DNS habilitado e configurado. Quanto mais dispositivos DNS

forem instalados, maior será o grau de redundância do qual uma empresa se beneficiará.

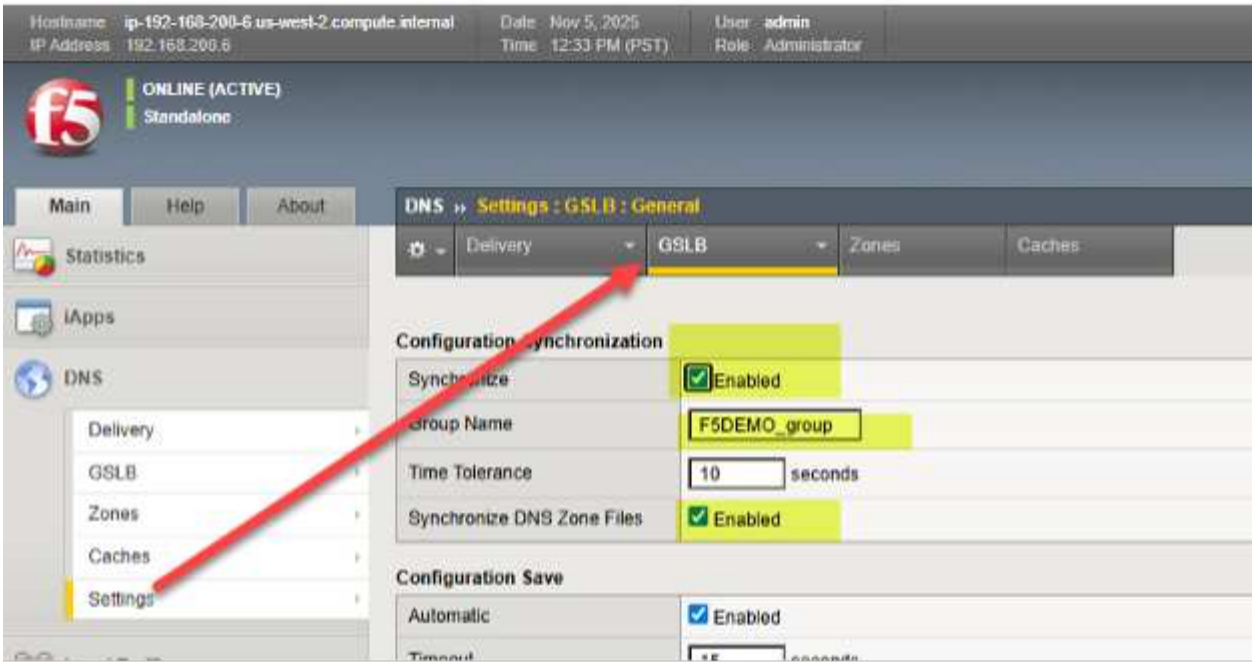
### BIG-IP DNS - Primeiros passos na configuração inicial

Após o dispositivo BIG-IP ter passado pelo menos pelo provisionamento inicial, use um navegador da web para acessar a interface TMUI (GUI do BIG-IP) e selecione Sistema → Provisionamento de Recursos. Conforme destacado, certifique-se de que o módulo “Tráfego Global (DNS)” esteja marcado e licenciado. Observe que, como mostra a imagem, é comum que o "Tráfego Local (LTM)" possa ser provisionado no mesmo dispositivo.

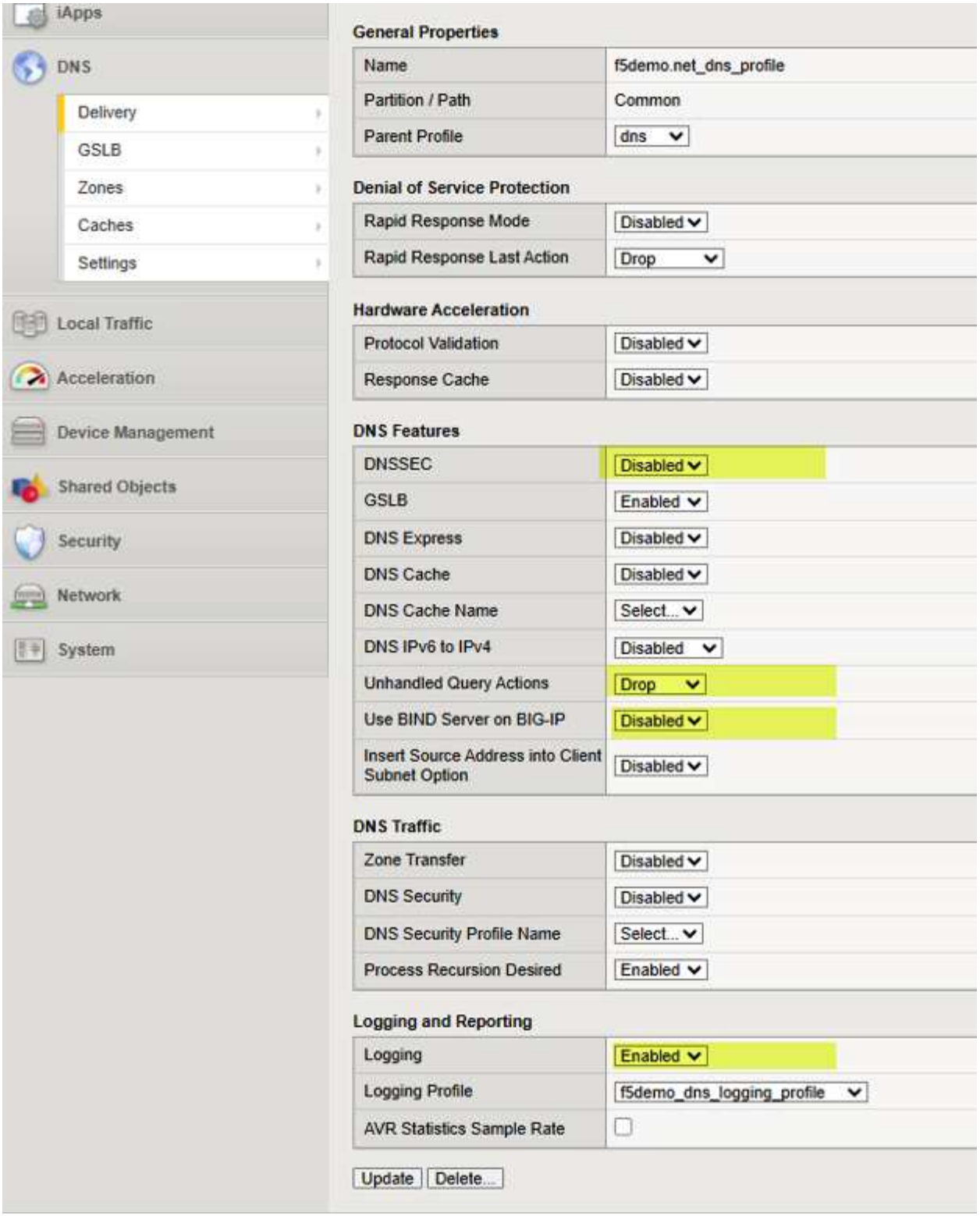



### Configurar os elementos fundamentais do protocolo DNS

O primeiro passo para o gerenciamento de tráfego global para sites StorageGRID é selecionar a guia DNS, onde praticamente todo o direcionamento de tráfego global será configurado, e escolher Configurações → GLSB. Ative as duas opções de sincronização e escolha um nome de grupo DNS que será compartilhado entre os dispositivos BIG-IP participantes.



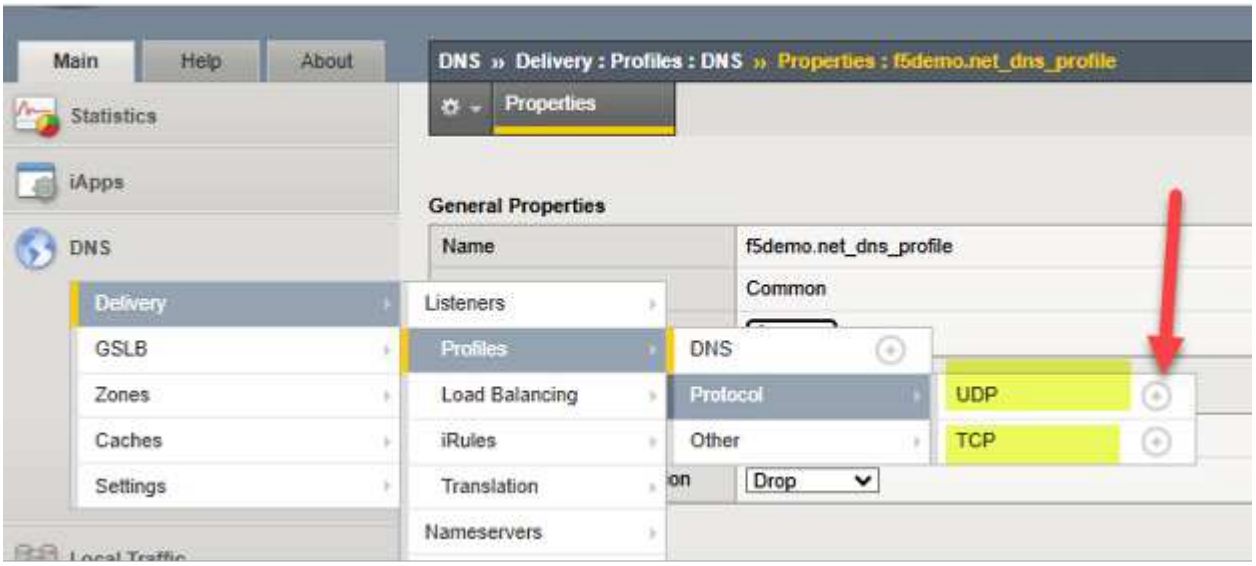

Em seguida, navegue até DNS > Entrega > Perfis > DNS: Criar e crie um perfil que irá controlar os recursos de DNS que você deseja ativar ou desativar. Consulte o link anterior para o guia de sala de aula sobre DNS, caso tenha interesse em gerar registros DNS específicos. Aqui está um exemplo de um perfil DNS funcional; observe os quatro destaques que representam configurações com valores importantes. Para sua informação, cada configuração possível é explicada no seguinte artigo da Base de Conhecimento (KB) da F5. "aqui".



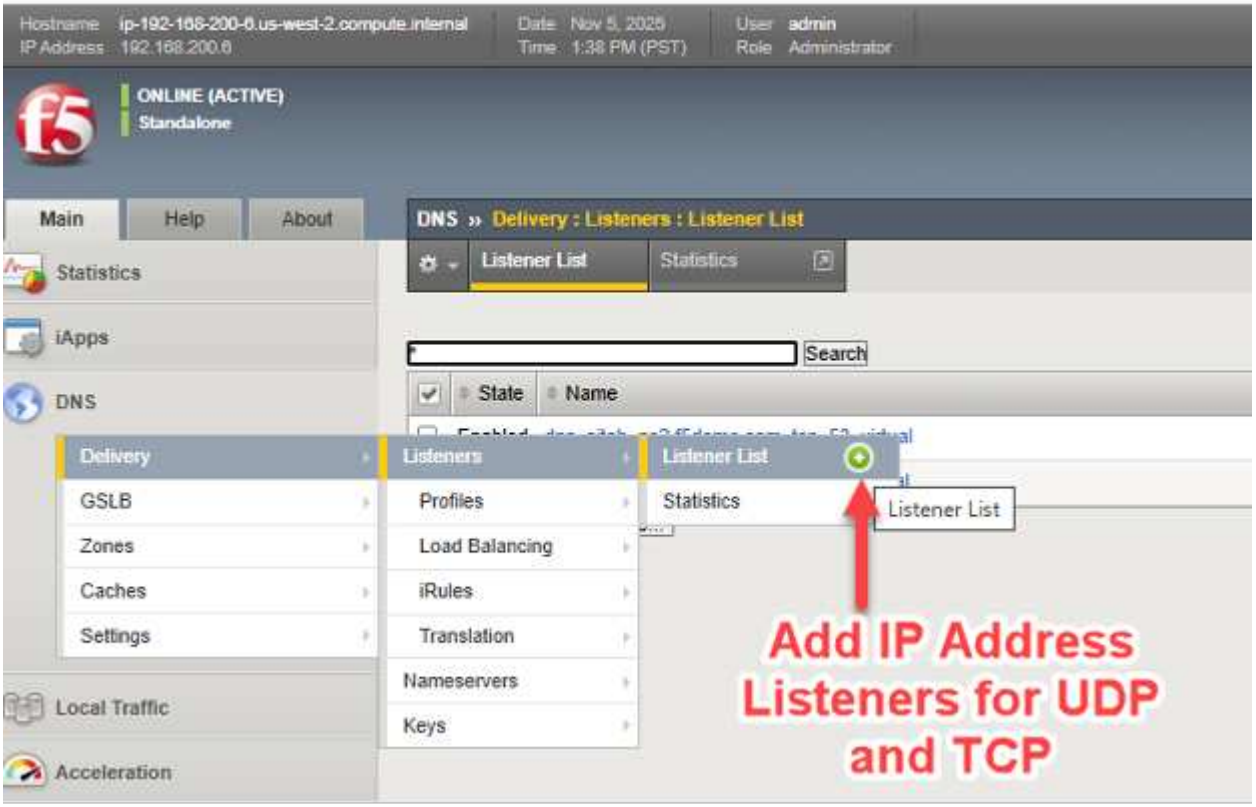

Neste ponto, podemos ajustar as características dos protocolos UDP e TCP, através da criação de "perfis", que podem transportar tráfego DNS envolvendo o BIG-IP. Basta criar um novo perfil para UDP e outro para TCP. Partindo do pressuposto de que o tráfego DNS irá atravessar links WAN, uma boa prática é simplesmente herdar as características do UDP e do TCP que comprovadamente apresentam bom desempenho em ambientes WAN. Para adicionar cada um, basta clicar no ícone “+” ao lado de cada protocolo e definir o perfil principal da seguinte forma:

UDP → usar o perfil “pai” “udp_gtm_dns”

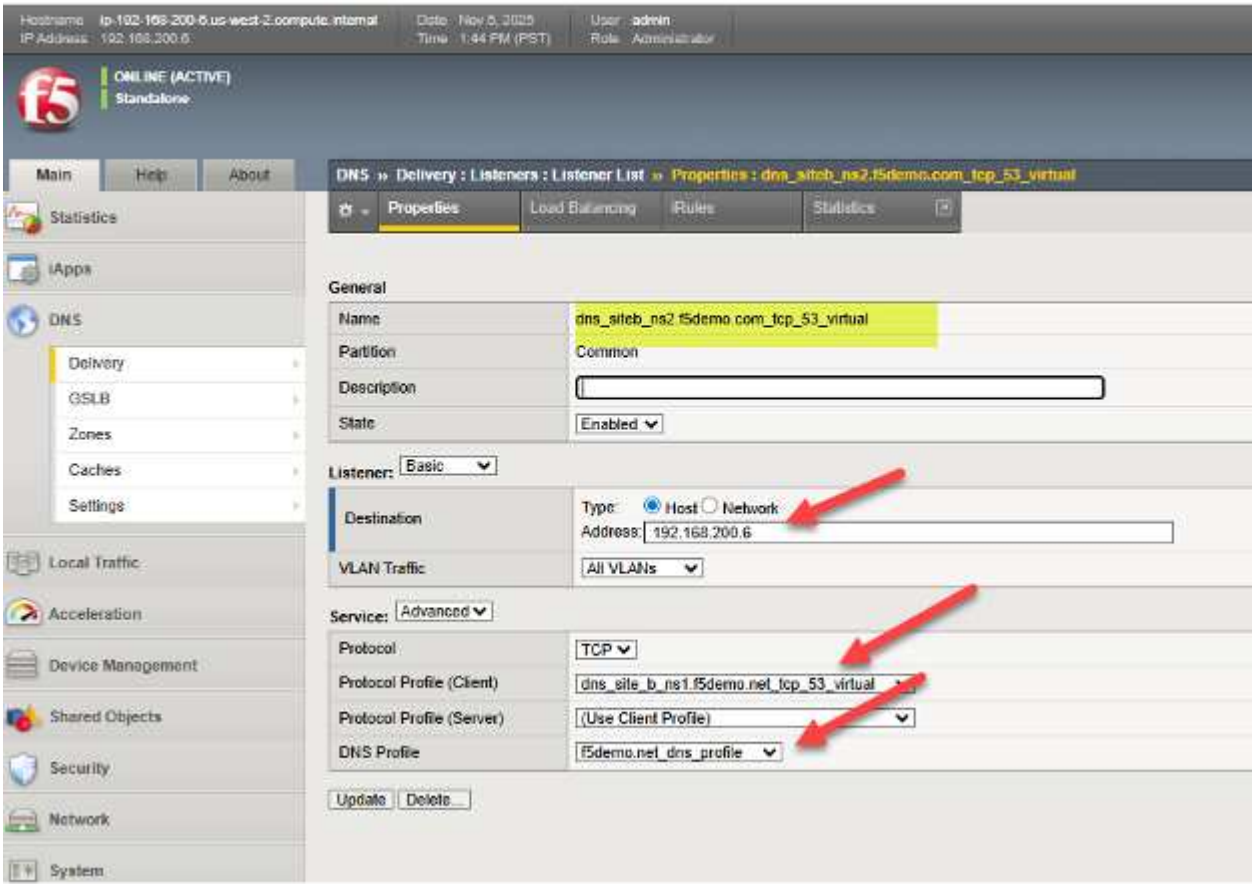
TCP → usar o perfil “pai” “f5-tcp-wan”




Agora, basta atribuirmos um endereço IP para o tráfego UDP e TCP que envolve o DNS do BIG-IP. Para quem está familiarizado com o BIG-IP LTM, isso é essencialmente a criação de servidores DNS virtuais, e servidores virtuais precisam de endereços IP de "escuta". Como mostra a captura de tela, siga as setas para criar servidores virtuais/de escuta para DNS/UDP e DNS/TCP.




Segue abaixo um exemplo de um servidor DNS BIG-IP em funcionamento, onde podemos ver as configurações do ouvinte do servidor virtual TCP e como ele integra muitas das etapas anteriores. Isso inclui referenciar o perfil DNS e o perfil de protocolo (TCP), bem como configurar um endereço IP válido para o DNS usar. Assim como ocorre com todos os objetos criados com o BIG-IP, é útil usar um nome significativo que sirva para identificar o objeto, como dns/siteb/TCP53 no exemplo de nome atribuído.



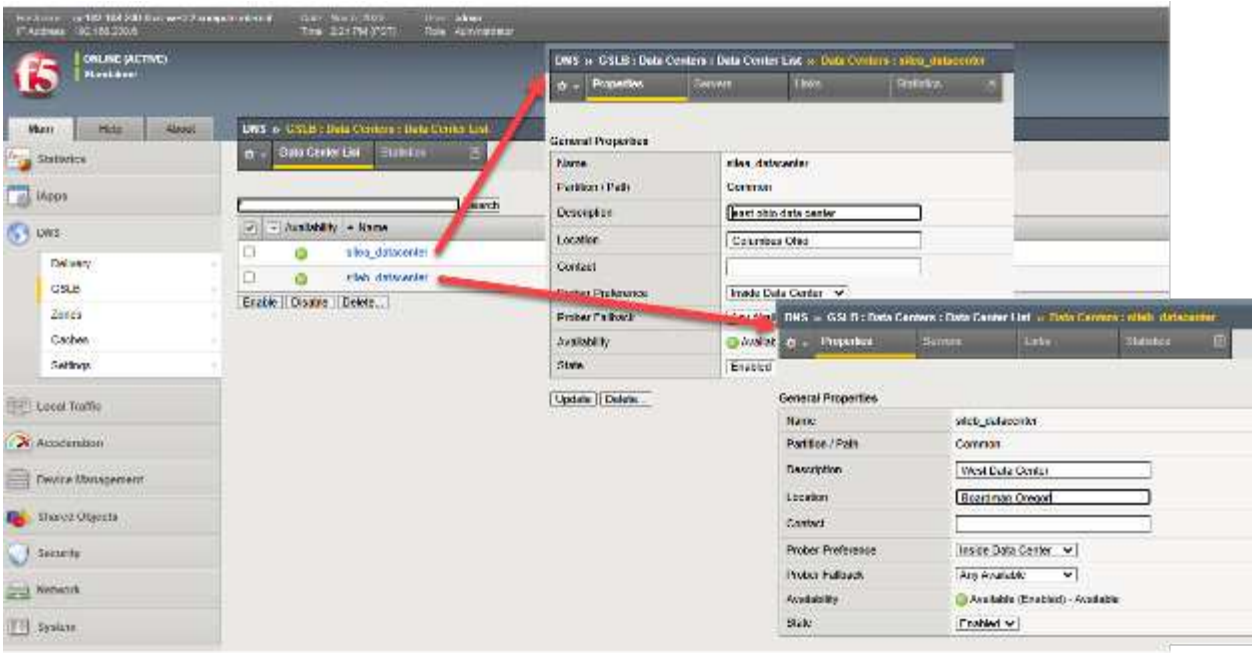

Isso conclui as etapas preliminares de configuração, geralmente realizadas "uma única vez", de um dispositivo BIG-IP com o módulo DNS ativado. Neste momento, estamos prontos para abordar os detalhes da configuração de uma solução global de gerenciamento de tráfego com nossos dispositivos, que, naturalmente, estará vinculada às características dos sites da StorageGRID .

## Configuração de Data Centers e estabelecimento de comunicação entre BIG-IPs em quatro etapas

#### Primeiro passo: Criar centros de dados

Cada site que irá hospedar clusters de nós para serem balanceados localmente pelo BIG-IP LTM deve ser inserido no BIG-IP DNS. Isso precisa ser feito em apenas um servidor DNS BIG-IP, pois estamos criando um grupo DNS sincronizado para dar suporte ao gerenciamento de tráfego; portanto, essa configuração será compartilhada entre os servidores DNS membros do grupo.

Através da interface gráfica do usuário (GUI) do TMUI, selecione DNS > GSLB > Data Centers > Lista de Data Centers e crie uma entrada para cada um dos sites do StorageGRID . Se estiver usando uma configuração de rede alinhada com a Figura 1, com o dispositivo DNS localizado em outros sites que não sejam StorageGRID , adicione Data Centers para esses sites, além dos sites de armazenamento. Neste exemplo, os sites a e b são criados em Ohio e Oregon, e os dispositivos BIG-IP são de uso duplo, com DNS e LTM.



**Etapa dois: Criar servidores (Lista de todos os dispositivos BIG-IP na solução)**

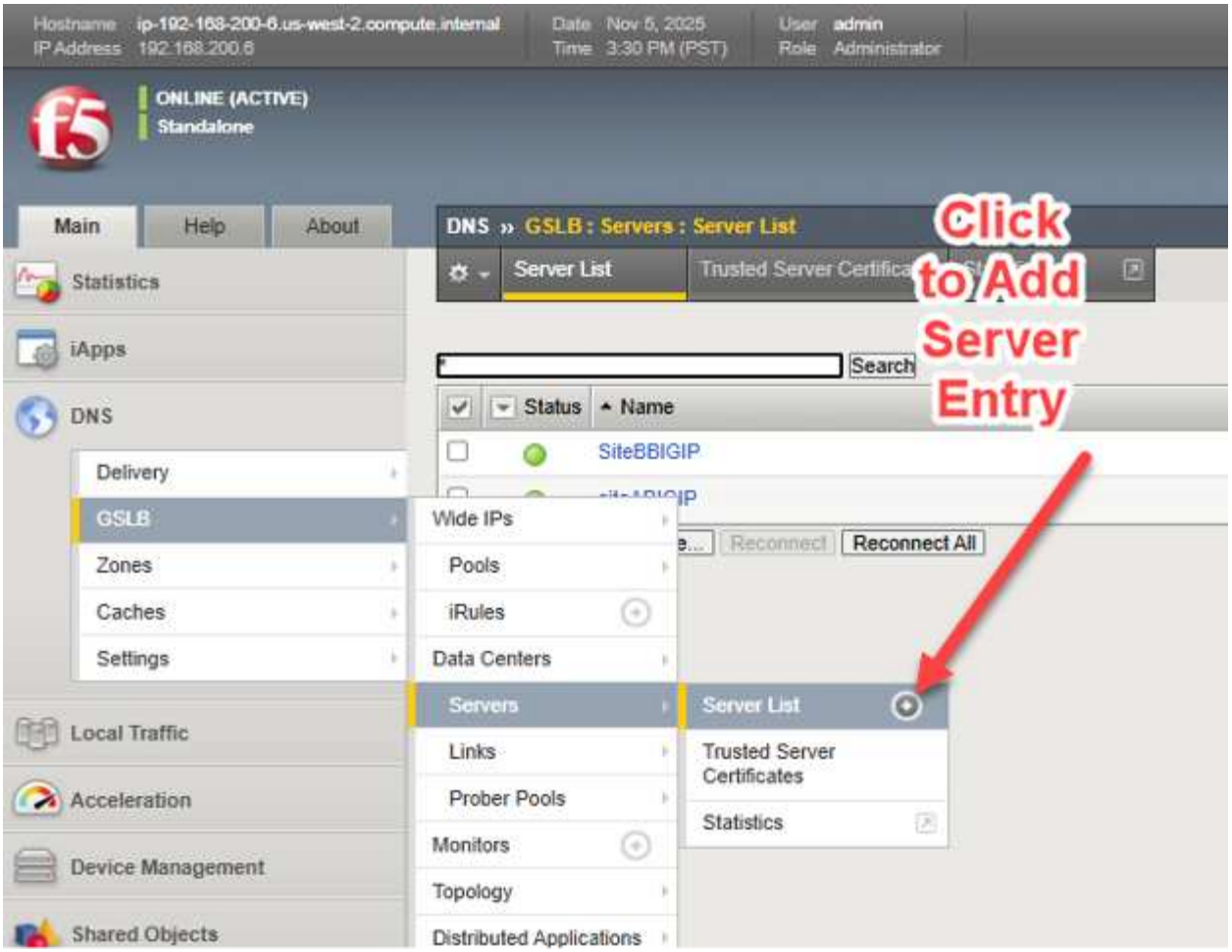
Agora estamos prontos para conectar os clusters individuais do StorageGRID à configuração de DNS do BIG-IP. Lembre-se, o dispositivo BIG-IP em cada local fará o balanceamento de carga do tráfego S3, por meio da configuração de servidores virtuais que vinculam um endereço IP/porta acessível de "front-end" a um conjunto de dispositivos Storage Node de "back-end", usando endereços IP/portas de "back-end".

Caso, por exemplo, todos os nós de armazenamento em um pool sejam desativados administrativamente, talvez devido à desativação de um site, ou inesperadamente por meio de falhas em verificações de integridade em tempo real, o tráfego será direcionado para outros sites alterando as respostas às consultas de DNS.

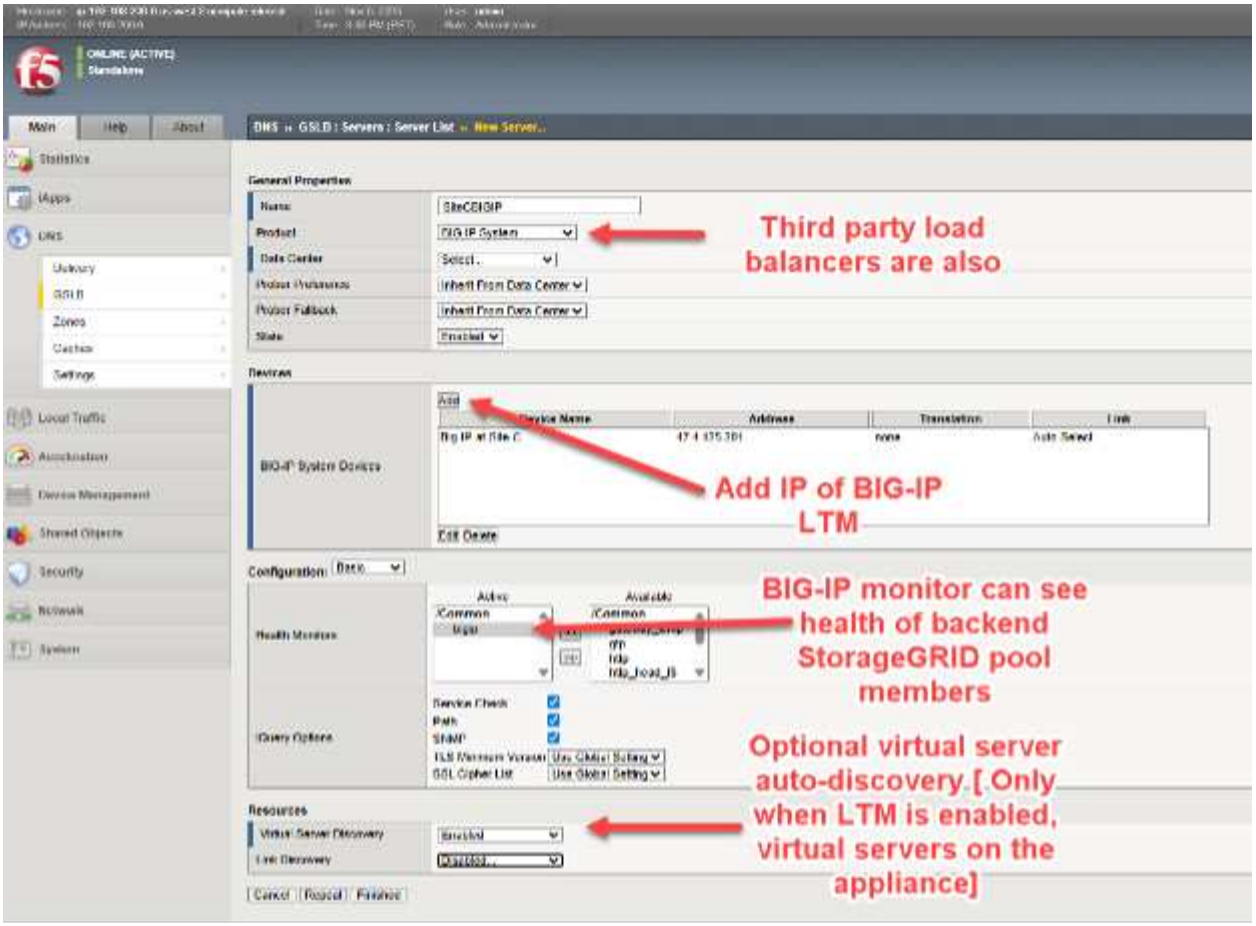
Para integrar os sites do StorageGrid, especificamente os servidores virtuais locais, à configuração de DNS do BIG-IP em cada dispositivo, a configuração precisa ser feita apenas uma vez. Em uma etapa futura, todas as configurações do grupo de dispositivos BIG-IP DNS serão sincronizadas.

Em termos simples, criaremos uma lista, denominada lista de servidores, de todos os nossos dispositivos BIG-IP, independentemente de estarem licenciados para DNS, LTM ou ambos. Esta lista principal será sincronizada com todos os dispositivos BIG-IP DNS após a conclusão do processo.

Em um dispositivo BIG-IP DNS licenciado, escolha DNS > GSLB > Servidores > Lista de servidores e clique no botão adicionar (+).




Os quatro elementos-chave ao adicionar cada BIG-IP incluem: * Selecionar o BIG-IP na lista suspensa de produtos; outros balanceadores de carga são possíveis, mas geralmente não possuem a capacidade de resposta e visibilidade em tempo real quando a integridade do nó de backend se deteriora em cada site. * Adicione o endereço IP do dispositivo BIG-IP DNS. Provavelmente, na primeira vez que um dispositivo BIG-IP DNS for adicionado, o endereço será o do dispositivo acessado pela GUI; em aplicações futuras, serão os endereços dos outros dispositivos na solução. * Escolha um monitor de integridade; use sempre “BIG-IP” quando o balanceador de carga adicionado for um appliance BIG-IP, para avaliação da integridade do nó StorageGRID de back-end. * Opcionalmente, solicite a descoberta automática do servidor virtual se o dispositivo for um dispositivo DNS/LTM duplo.

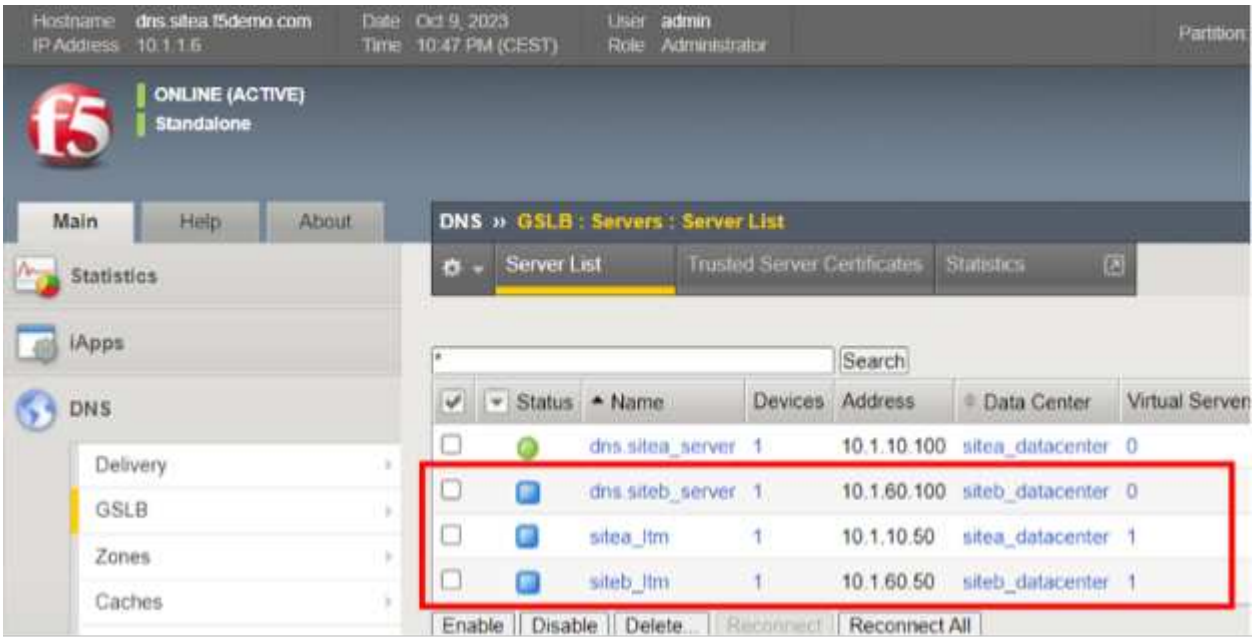



Em algumas situações, como problemas transitórios de rede ou regras de ACL de firewall entre locais de rede, ao adicionar um dispositivo remoto nesta etapa, a descoberta de servidores virtuais pode não exibir entradas para dispositivos remotos com LTM configurado. Nesses casos, após adicionar o novo dispositivo ("servidor"), é possível adicionar manualmente os servidores virtuais, conforme indicado abaixo. Ao adicionar um dispositivo BIG-IP somente para DNS, não haverá servidores virtuais a serem descobertos ou adicionados a esse dispositivo.


Precisamos adicionar essas entradas de servidor para cada dispositivo em nossa solução em todos os sites, incluindo dispositivos BIG-IP DNS, dispositivos BIG-IP LTM e quaisquer dispositivos que desempenhem as funções duplas de unidades DNS e LTM.

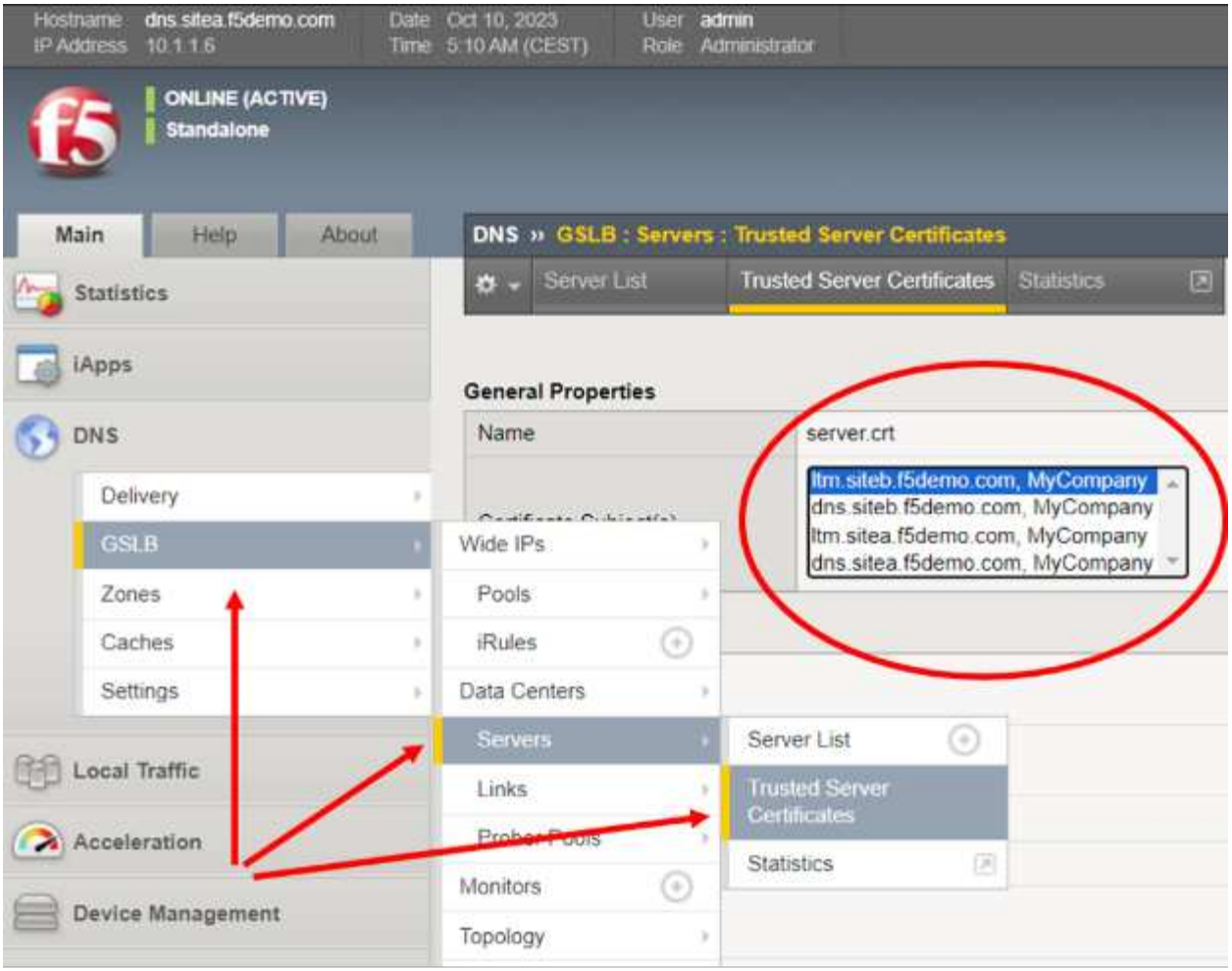
#### Etapa três: Estabelecer confiança entre todos os dispositivos BIG-IP

No exemplo a seguir, quatro dispositivos foram adicionados como servidores e estão distribuídos em dois locais. Observe que cada site possui um BIG-IP DNS e um BIG-IP LTM dedicados. No entanto, todos os dispositivos, exceto aquele em que o usuário está conectado no momento, exibem ícones azuis na coluna "Status". Isso significa que ainda não foi estabelecida uma relação de confiança com os outros dispositivos BIG-IP.




Para adicionar confiança, acesse o BIG-IP via SSH, onde os detalhes de configuração foram inseridos pela interface gráfica, e utilize a conta "root" para acessar a interface de linha de comando do BIG-IP. Execute o seguinte comando único no prompt: *bigip_add*

O comando "bigip_add" extrai o certificado de gerenciamento dos dispositivos BIG-IP de destino para uso durante a configuração do canal criptografado "iQuery" entre os servidores GSLB no cluster. O iQuery, por padrão, é executado usando a porta TCP 4353 e é o mecanismo de pulsação que permite que os membros DNS do BOG-IP permaneçam sincronizados. Ele utiliza XML e gzip no canal criptografado. Ao executar "bigip_add" sem nenhuma opção, o comando será executado em todos os dispositivos BIGIP na lista do servidor GSLB, usando o nome de usuário atual para se conectar aos endpoints. Para verificar rapidamente se tudo correu bem, basta retornar à interface gráfica do BIG-IP e confirmar se todos os servidores agora possuem certificados listados no menu suspenso exibido.




**Passo quatro: Sincronize todos os dispositivos BIG-IP DNS com o grupo DNS.**

A etapa final permitirá que todos os dispositivos BIG-IP DNS sejam totalmente configurados usando apenas a interface gráfica do usuário (GUI) do TMUI em uma única unidade. Em um exemplo hipotético, onde existem dois sites StorageGRID , isso significa usar SSH para acessar a linha de comando do DNS BIG-IP do **outro** site. Após conectar-se como root e garantir que as políticas/ACLs do firewall permitam que os dois dispositivos BIG-IP DNS se comuniquem nas portas TCP 22 (SSH), 443 (HTTPS) e 4354 (protocolo F5 iQuery), execute o seguinte comando no prompt: `_gtm_add <endereço IP do primeiro dispositivo BIG-IP DNS, onde todas as etapas da GUI foram realizadas anteriormente>`.

A partir deste ponto, todas as configurações adicionais de DNS podem ser realizadas em qualquer dispositivo BIG-IP DNS que tenha sido adicionado ao grupo. O comando acima, gtm_add, não precisa ser aplicado a membros de dispositivo que sejam somente LTM. Somente dispositivos com suporte a DNS precisam deste comando para fazer parte do grupo DNS sincronizado.

### Configuração de Data Centers e estabelecimento de comunicação entre dispositivos BIG-IP

Neste ponto, todas as etapas para criar o grupo de dispositivos BIG-IP DNS subjacente e íntegro estão concluídas. Agora podemos prosseguir com a criação de nomes, FQDNs, que apontem para nossos serviços web/S3 distribuídos, expostos em cada datacenter do StorageGRID .

Esses nomes são chamados de "Wide IPs", ou WIPs, e são FQDNs DNS normais com registros de recursos DNS A. No entanto, em vez de apontar para um servidor como um registro de recurso A tradicional, eles apontam internamente para conjuntos de servidores virtuais BIG-IP. Cada pool, individualmente, pode ser composto por um ou mais servidores virtuais. Um cliente S3 que solicita um endereço IP para resolução de

nomes receberá o endereço do servidor virtual S3 no site StorageGRID ideal, selecionado pela política.

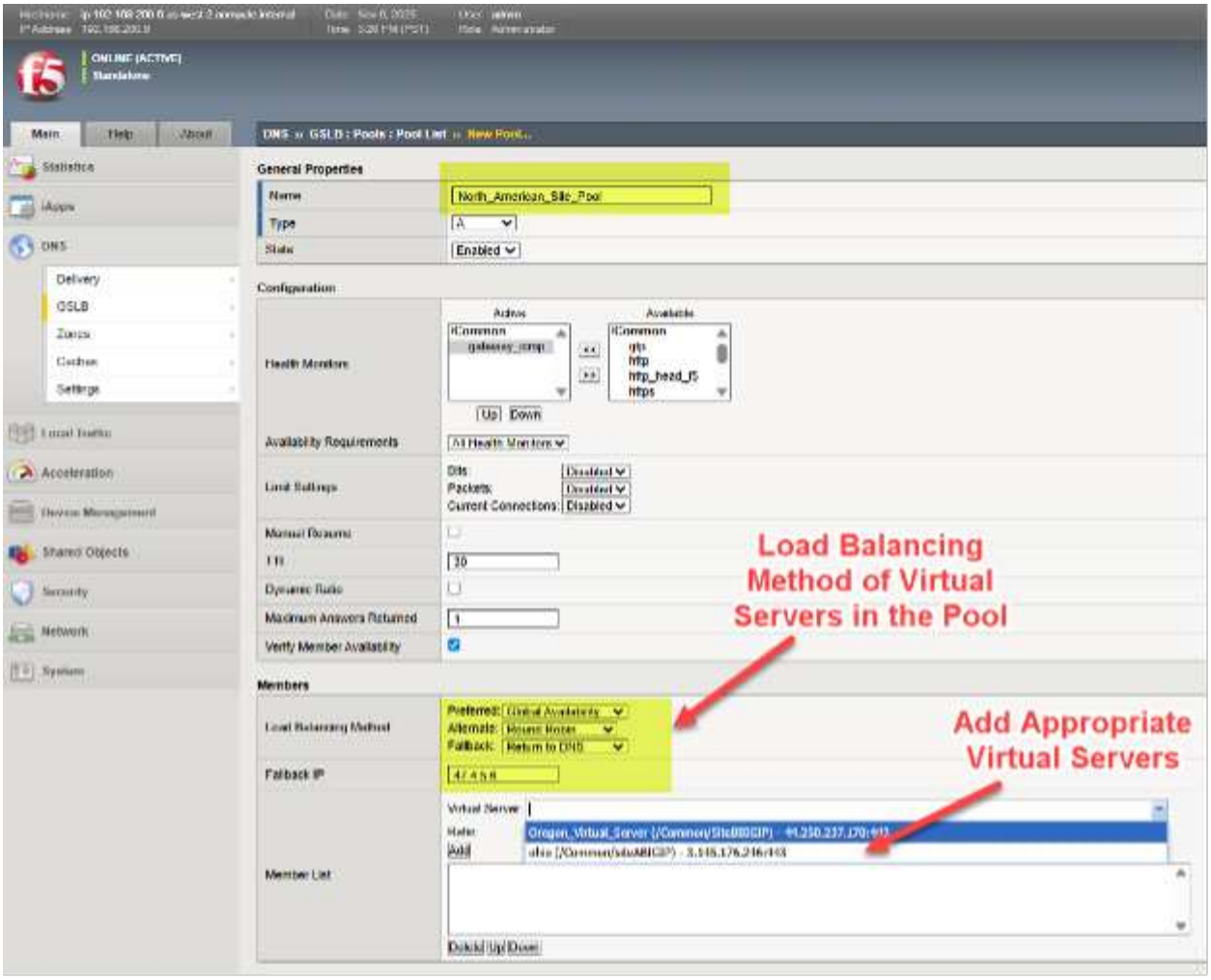
#### IPs amplos, pools e servidores virtuais em poucas palavras

Para dar um exemplo simples e fictício, um projeto em andamento para o nome **storage.quantumvault.com** poderia ver a solução DNS BIG-IP vinculada a dois conjuntos de servidores virtuais potenciais. O primeiro grupo pode ser composto por 4 locais na América do Norte; o segundo grupo pode ser composto por 3 locais na Europa.

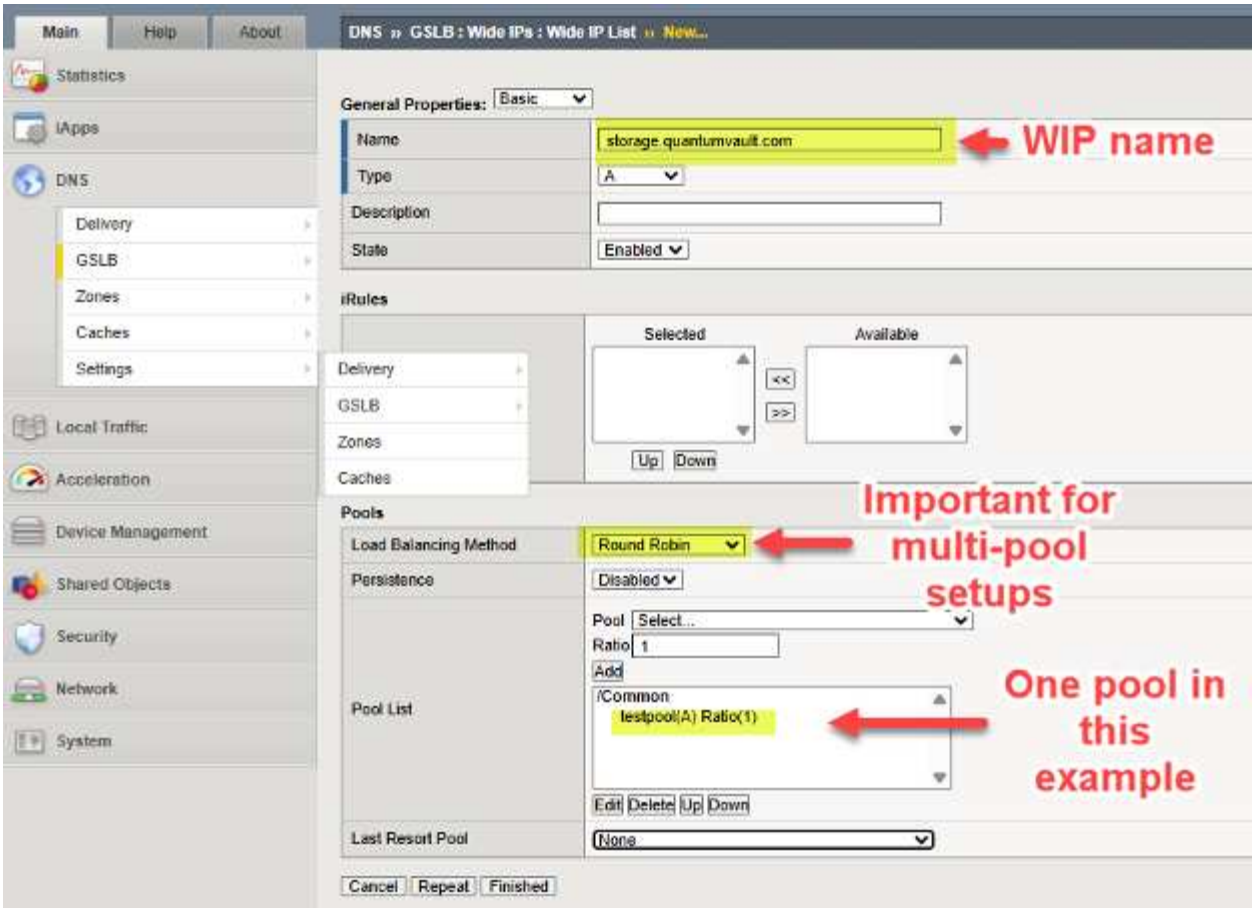
O conjunto de dispositivos selecionados pode ser definido a partir de uma série de decisões políticas; talvez uma proporção simples de 5:1 possa ser usada para direcionar a maior parte do tráfego para os sites StorageGRID da América do Norte. Talvez seja mais provável uma escolha baseada na topologia, onde o pool é escolhido de forma que, por exemplo, todo o tráfego S3 originado na Europa seja direcionado para sites europeus, e o restante do tráfego S3 mundial seja direcionado para data centers norte-americanos.

Uma vez que o BIG-IP DNS determine um pool, supondo que o pool da América do Norte tenha sido selecionado, o registro DNS A retornado para resolver storage.quantumvault.com pode ser qualquer um dos 4 servidores virtuais suportados pelo BIG-IP LTM em qualquer um dos 4 sites da América do Norte. Novamente, a escolha é orientada por políticas; existem abordagens "estáticas" simples, como o Round-Robin, enquanto seleções "dinâmicas" mais avançadas, como sondagens de desempenho para medir a latência de cada site a partir de resolvedores DNS locais, são mantidas e usadas como critérios para a seleção do site.

Para configurar um pool de servidores virtuais em um BIG-IP DNS, siga o caminho do menu **DNS > GSLB > Pools > Lista de Pools > Adicionar (+)**. Neste exemplo, podemos ver que vários servidores virtuais norte-americanos são adicionados a um pool e a abordagem preferencial para balanceamento de carga, quando este pool é selecionado, é escolhida de forma hierárquica.




Adicionamos o WIP (Wide IP), o nome do nosso serviço que será resolvido pelo DNS, a uma implantação seguindo o caminho DNS > GSLB > Wide IPs > Lista de Wide IPs > Criar (+). No exemplo a seguir, fornecemos um exemplo de trabalho em andamento (WIP) para um serviço de armazenamento habilitado para S3.




**Ajuste o DNS para suportar o gerenciamento de tráfego global.**

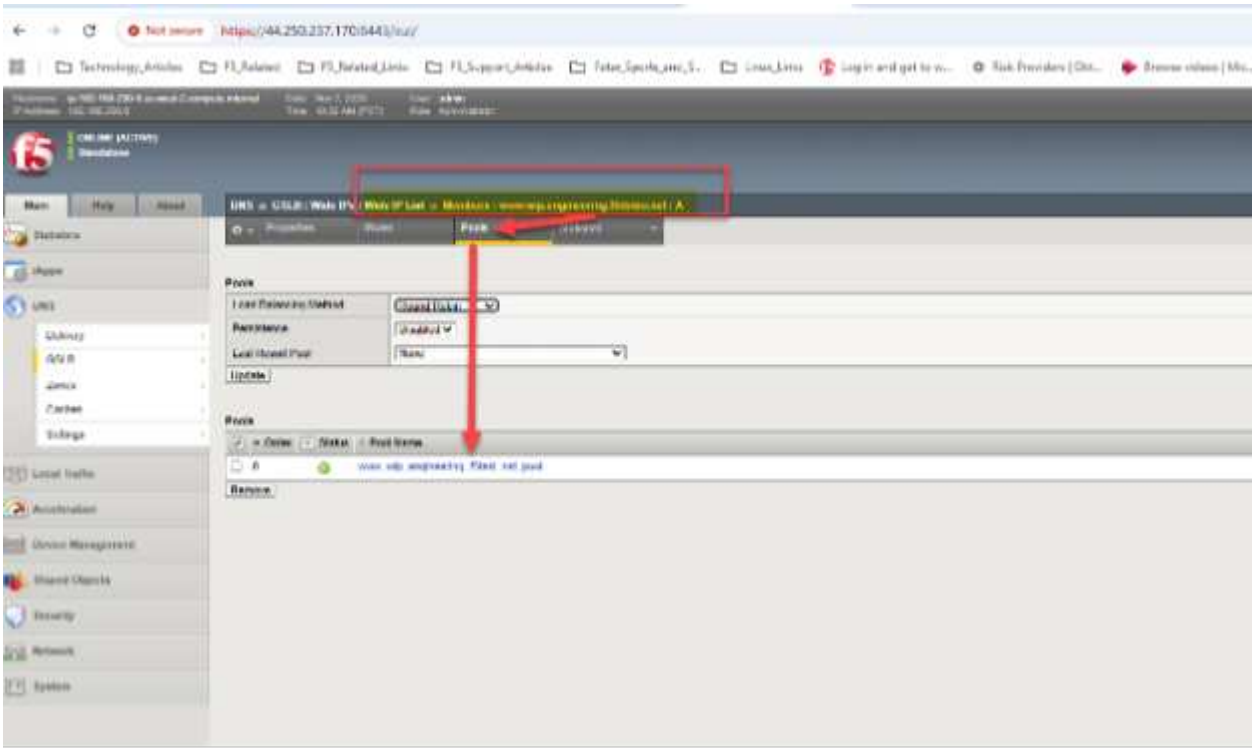
Neste ponto, todos os nossos dispositivos BIG-IP subjacentes estão prontos para executar o GSLB (balanceamento de carga global do servidor). Basta ajustarmos e atribuirmos os nomes usados para os fluxos de tráfego do S3 para aproveitarmos a solução. A abordagem geral consiste em delegar parte de um domínio DNS existente de uma empresa ao controle do BIG-IP DNS. Isso significa "reservar" uma seção do espaço de nomes, um subdomínio, e delegar o controle desse subdomínio aos dispositivos BIG-IP DNS. Tecnicamente, isso é feito garantindo que os dispositivos BIG-IP DNS tenham registros de recursos DNS (RRs) do tipo A no DNS corporativo e, em seguida, transformando esses nomes/endereços em registros de recursos DNS de servidor de nomes (NS) para o domínio delegado.

Existem várias maneiras pelas quais as empresas mantêm o DNS atualmente, sendo um método uma solução totalmente hospedada. Um exemplo disso seria operar e gerenciar o DNS por meio do Windows Server 2025. Uma abordagem alternativa para uma empresa seria utilizar provedores de DNS em nuvem, como o AWS Route53 ou o Squarespace.

Segue um exemplo fictício para fins ilustrativos. Temos o StorageGRID com suporte para leitura e gravação de objetos via protocolo S3, com um domínio existente gerenciado pelo AWS Route53. O domínio de exemplo existente é f5demo.net.

Gostaríamos de atribuir o subdomínio engineering.f5demo.net aos dispositivos BIG-IP DNS para gerenciamento de tráfego global. Para isso, criamos um novo registro de recurso NS (servidor de nomes) para engineering.f5demo.net e o direcionamos para a lista de nomes de dispositivos DNS do BIG-IP. Em nosso exemplo, temos dois dispositivos BIG-IP DNS e, portanto, criamos dois registros de recursos A para eles.

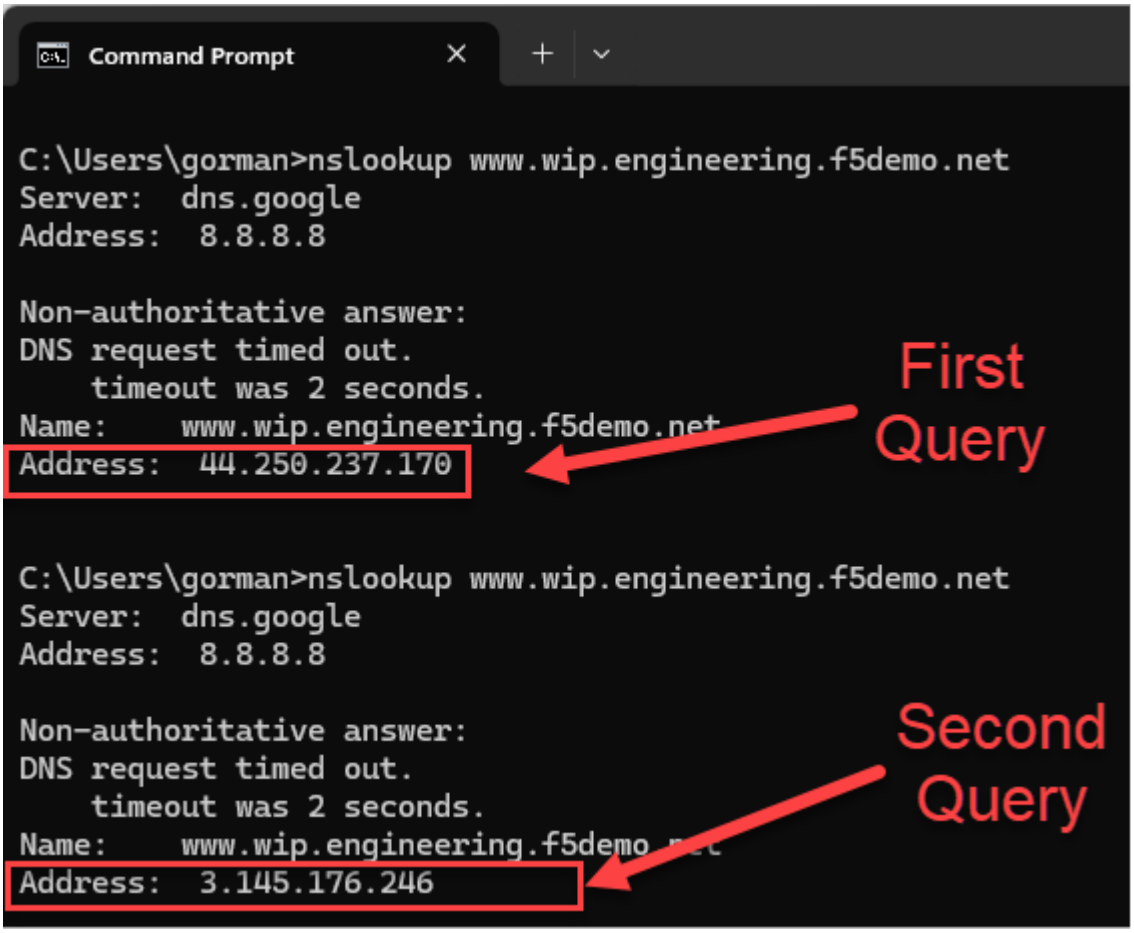

Agora, como exemplo, vamos configurar um Wide IP (WIP) em nosso DNS BIG-IP. Como o DNS usa sincronização de grupo, precisamos ajustar apenas a interface gráfica de um dos dispositivos. Na GUI do BIG-IP DNS, vá para **DNS > GSLB > Wide IPs > Wide IP List (+)**. Lembre-se que, em uma configuração tradicional de DNS com FQDN, seria necessário inserir um ou mais endereços IPv4; em nosso caso, simplesmente apontamos para um ou mais conjuntos de servidores virtuais do StorageGRID .



Em nosso exemplo, temos servidores web HTTPS genéricos localizados em sites de Ohio e Oregon. Com uma abordagem simples de "round robin", devemos conseguir ver o DNS global respondendo às consultas para os mapeamentos de registro de recurso A para *www.wip.engineering.f5demo.net* com ambos os IPs do servidor virtual.


Um teste simples pode ser feito com navegadores da web ou, no caso do S3 usando o StorageGRID, talvez com ferramentas gráficas como o S3Browser. Cada consulta DNS terá como destino o próximo site do data center no pool, devido à nossa escolha de Round Robin dentro do pool.

Em nossa configuração de exemplo, podemos usar o dig ou o nslookup para gerar rapidamente uma série de duas consultas DNS e garantir que o BIG-IP DNS esteja realmente realizando um balanceamento de carga round robin, resultando no recebimento de tráfego por ambos os sites ao longo do tempo.

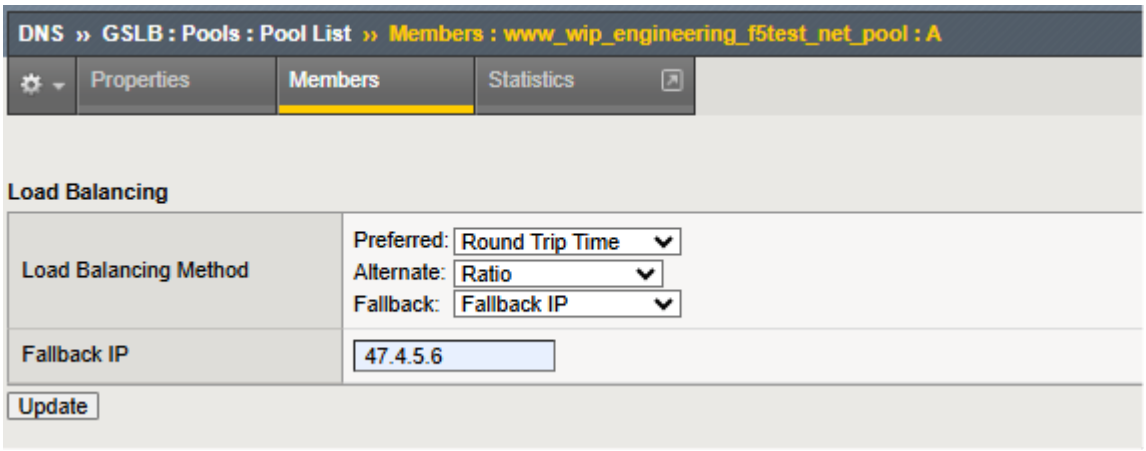



**Sugestão de exploração para técnicas mais avançadas.**

Uma das muitas abordagens possíveis seria usar o modo de "Disponibilidade Global" em vez do exemplo simples de "Round Robin" dado acima. Com a Disponibilidade Global, o tráfego pode ser direcionado para a

sequência de pools ou servidores virtuais dentro de um único pool. Dessa forma, todo o tráfego do S3 poderia, por padrão, ser direcionado, por exemplo, para um site na cidade de Nova York.

Se as verificações de integridade indicarem um problema com a disponibilidade do nó StorageGRID neste local, o tráfego poderá ser direcionado para St. Louis. Caso St. Louis enfrente problemas de saúde, um local em Frankfurt poderia, por sua vez, começar a receber transações de leitura ou gravação S3. Assim, a disponibilidade global é uma das abordagens para a resiliência geral da solução S3 StorageGRID . Outra abordagem consiste em combinar diferentes métodos de balanceamento de carga, utilizando uma abordagem em camadas.




Neste exemplo, a opção “dinâmica” é a primeira escolha de balanceamento de carga para os sites no pool configurado. No exemplo apresentado, uma abordagem de medição contínua, utilizando a sondagem ativa do desempenho do resolvedor DNS local, é mantida e serve como catalisador para a seleção do site. Caso essa abordagem não esteja disponível, os locais individuais podem ser selecionados pela proporção atribuída a cada um. Com essa proporção, sites StorageGRID maiores e com maior largura de banda podem receber mais transações S3 do que sites menores. Por fim, como possível cenário de recuperação de desastres, caso todos os sites no pool fiquem indisponíveis, o IP de fallback especificado é usado como último recurso. Um dos métodos de balanceamento de carga mais interessantes do BIG-IP DNS é o "Topology", no qual a fonte de entrada das consultas DNS, o resolvedor DNS local do usuário S3, é observada e, usando informações de topologia da Internet, o site aparentemente mais "próximo" é selecionado do conjunto.

Por fim, se os sites abrangerem o mundo inteiro, pode valer a pena considerar o uso da tecnologia de "sondagem" dinâmica, discutida em detalhes no manual do F5 BIG-IP DNS. Com sondagens, é possível monitorar fontes frequentes de consultas DNS, como por exemplo, um parceiro de negócios B2B cujo tráfego geralmente utiliza o mesmo resolvedor DNS local. As sondagens DNS do BIG-IP podem ser iniciadas a partir do BIG-IP LTM em cada local ao redor do mundo, para determinar, de forma geral, qual local em potencial provavelmente ofereceria a menor latência para transações S3. Assim sendo, o tráfego proveniente da Ásia pode ser melhor atendido por sites StorageGRID asiáticos do que por sites localizados na América do Norte ou na Europa.

## Conclusão

A integração do F5 BIG-IP com o NetApp StorageGRID resolve desafios técnicos relacionados à disponibilidade e consistência de dados em vários locais e à otimização do roteamento de transações S3. A implementação dessa solução aprimora a resiliência, o desempenho e a confiabilidade do armazenamento, tornando-a ideal para empresas que buscam uma infraestrutura de armazenamento robusta, escalável e flexível.

Para saber mais, a documentação oficial da F5 para BIG-IP DNS pode ser encontrada aqui. "link". Também é possível encontrar um guia prático para sala de aula que fornece instruções passo a passo para uma configuração de exemplo. "aqui".

# Configuração SNMP do Datadog

*Por Aron Klein*

Configure o Datadog para coletar métricas e traps do StorageGRID snmp.

## Configurar Datadog

O Datadog é uma solução de monitoramento que fornece métricas, visualizações e alertas. A seguinte configuração foi implementada com o agente linux versão 7.43.1 em um host Ubuntu 22.04.1 implantado local no sistema StorageGRID.

### Arquivos de Perfil e Trap gerados a partir do arquivo MIB do StorageGRID

O Datadog fornece um método para converter arquivos MIB do produto em arquivos de referência de datadog necessários para mapear as mensagens SNMP.

Este arquivo StorageGRID yaml para mapeamento de resolução de armadilha Datadog gerado após a instrução encontrada "aqui". Coloque este arquivo em /etc/datadog-Agent/conf.d/snmp.d/traps_dB/

- "Baixe o arquivo trap yaml" E
  - **soma de verificação md5** 42e27e4210719945a46172b98c379517
  - **soma de verificação sha256** d0fe5c8e6ca3c902d054f854b70a85f928cba8b7c76391d356f05d2cf7b6887 e

Este arquivo yaml de perfil do StorageGRID para mapeamento de métricas do Datadog gerado após a instrução encontrada "aqui". Coloque este arquivo em /etc/datadog-Agent/conf.d/snmp.d/profiles/

- "Baixe o arquivo yaml de perfil" E
  - **md5 checksum** 72bb7784f4801adda4e0c3ea77df19aa
  - **sha256 checksum** b6b7fadd33063422a8bb8e39b3ead8ab38349ee0229926eadc85f0087b8cee

### Configuração de dados SNMP para métricas

A configuração do SNMP para métricas pode ser gerenciada de duas maneiras. Você pode configurar para deteção automática fornecendo um intervalo de endereços de rede contendo o(s) sistema(s) StorageGRID ou definir os IP dos dispositivos individuais. A localização da configuração é diferente com base na decisão tomada. A descoberta automática é definida no arquivo yaml do agente de dados. Definições explícitas de dispositivo são configuradas no arquivo yaml de configuração snmp. Abaixo estão exemplos de cada um para o mesmo sistema StorageGRID.

#### Descoberta automática

configuração localizada em /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
```

#### Dispositivos individuais

/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'
```

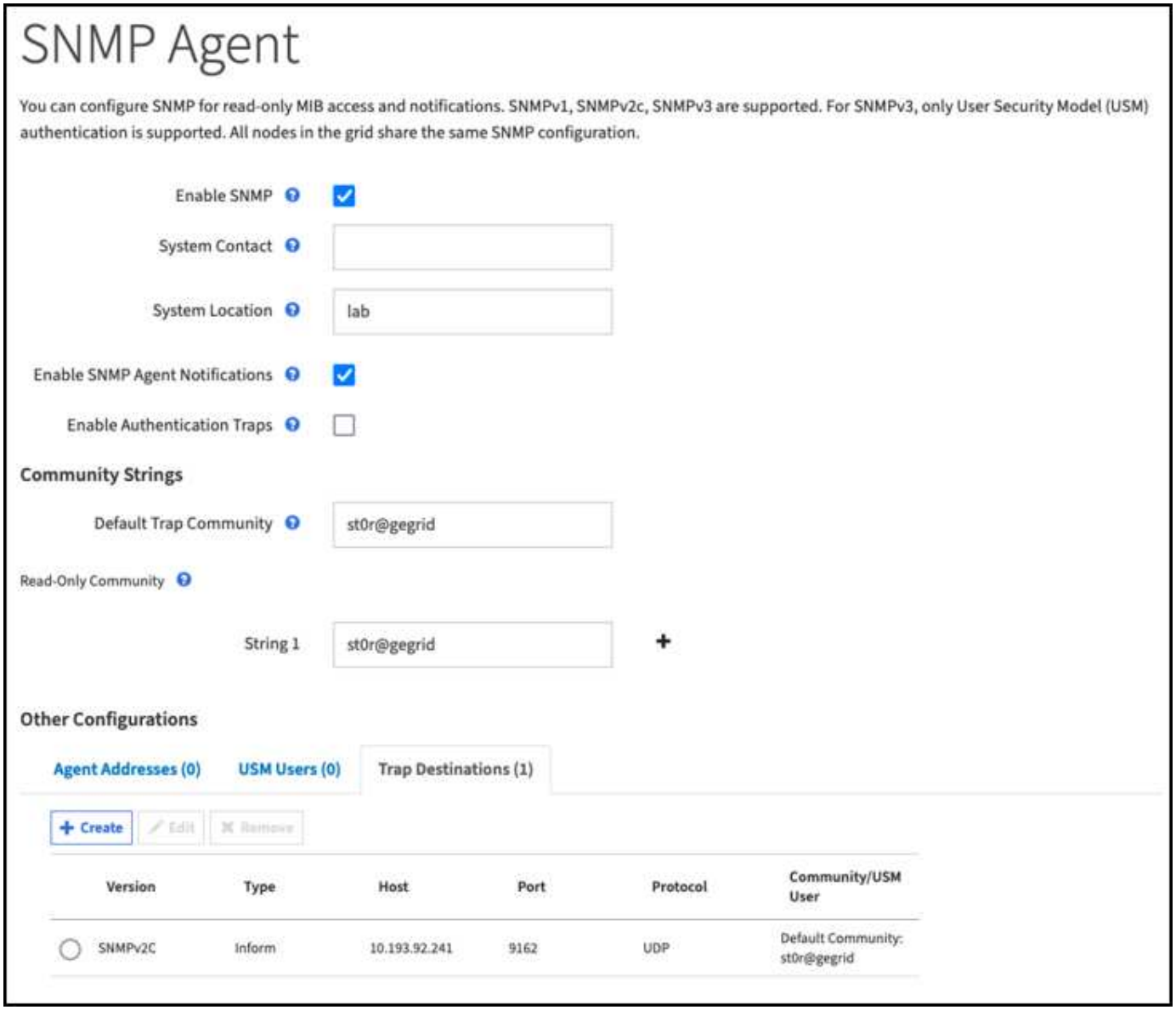
### Configuração SNMP para traps

A configuração para traps SNMP é definida no arquivo yaml de configuração de dados /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
```

**Exemplo de configuração StorageGRID SNMP**

O agente SNMP no seu sistema StorageGRID está localizado na guia configuração, coluna Monitoramento. Ative o SNMP e introduza as informações pretendidas. Se você deseja configurar traps, selecione "traps Destinations" e crie um destino para o host do agente Datadog que contém a configuração de traps.



# Use rclone para migrar, COLOCAR e EXCLUIR objetos no StorageGRID

*Por Siegfried Hepp e Aron Klein*

Rclone é uma ferramenta de linha de comando gratuita e cliente para operações S3. Você pode usar o rclone para migrar, copiar e excluir dados de objetos no StorageGRID. O rclone inclui a capacidade de excluir buckets mesmo quando não estiver vazio com uma função de "purga", como visto em um exemplo abaixo.

## Instalar e configurar o rclone

Para instalar o rclone em uma estação de trabalho ou servidor, baixe-o em "rclone.org".

### Etapas iniciais de configuração

1. Crie o arquivo de configuração rclone executando o script de configuração ou criando manualmente o arquivo.
2. Para este exemplo, vou usar o sgdemo para o nome do endpoint StorageGRID S3 remoto na configuração rclone.
   a. Crie o arquivo de configuração -/.config/rclone/rclone.conf

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

   b. Execute o rclone config

n 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
```

## Exemplos básicos de comandos

- **Crie um bucket:**

  `rclone mkdir remote:bucket`

  ```
  rclone mkdir sgdemo:test01
  ```

  Use --no-check-certificate se você precisar ignorar certificados SSL.

- **Liste todos os baldes:**

```
rclone lsd remote:
```

n rclone lsd sgdemo:

- **Liste objetos em um bucket específico:**

```
rclone ls remote:bucket
```

n 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
```

- **Excluir um balde:**

```
rclone rmdir remote:bucket
```

rclone rmdir sgdemo:test02

- **Coloque um objeto:**

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

cópia rclone/test/testfile.txt sgdemo:test01

- **Obter um objeto:**

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

Cópia rclone sgdemo:TEST01/testfile.txt/test/testfileS3.txt

- **Excluir um objeto:**

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

n rclone delete sgdemo:test01/testfile.txt

- **Migrar objetos em um bucket**

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

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

n 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
```

Use --progress ou -P para exibir o progresso da tarefa. Caso contrário, não há saída.

- **Excluir um bucket e todo o conteúdo do objeto**

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

n 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
```

n rclone ls sgdemo:test01

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

# Práticas recomendadas do StorageGRID para implantação com o Veeam Backup and Replication

*Por Oliver Haensel e Aron Klein*

Este guia concentra-se na configuração do NetApp StorageGRID e, em parte, no Veeam Backup and Replication. Este documento foi criado para administradores de storage e rede que estão familiarizados com os sistemas Linux e têm a tarefa de manter ou implementar um sistema NetApp StorageGRID em combinação com o Veeam Backup and Replication.

## Visão geral

Os administradores de storage buscam gerenciar o crescimento de seus dados com soluções que atendam às metas de disponibilidade, recuperação rápida, escalabilidade para atender às suas necessidades e automatizar suas políticas de retenção de dados a longo prazo. Essas soluções também devem fornecer proteção contra perdas ou ataques mal-intencionados. Juntas, a Veeam e a NetApp fizeram uma parceria para criar uma solução de proteção de dados que combina o Veeam Backup & Recovery com o NetApp StorageGRID para storage de objetos no local.

A Veeam e a NetApp StorageGRID fornecem uma solução fácil de usar que trabalham juntas para ajudar a atender às demandas do rápido crescimento de dados e do aumento das regulamentações em todo o mundo. O storage de objetos baseado em nuvem é conhecido por sua resiliência, capacidade de escala, eficiências operacionais e de custo que o tornam uma escolha natural como destino para seus backups. Este documento fornecerá orientações e recomendações para a configuração de sua solução Veeam Backup e do sistema StorageGRID.

A carga de trabalho de objetos da Veeam cria um grande número de operações simultâneas DE PUT, DELETE e LIST DE objetos pequenos. A ativação da imutabilidade será adicionada ao número de solicitações ao armazenamento de objetos para definir versões de retenção e listagem. O processo de uma tarefa de backup inclui a gravação de objetos para a alteração diária, então depois que as novas gravações forem concluídas, a tarefa excluirá quaisquer objetos com base na política de retenção do backup. O agendamento de trabalhos de cópia de segurança quase sempre se sobrepõe. Essa sobreposição resultará em uma grande parte da janela de backup que consiste em 50/50 COLOCAR/EXCLUIR carga de trabalho no armazenamento de objetos. Fazer ajustes na Veeam ao número de operações simultâneas com a configuração de slot de tarefa, aumentar o tamanho do objeto aumentando o tamanho do bloco de tarefas de backup, reduzir o número de objetos nas solicitações de exclusão de vários objetos e escolher a janela de tempo máximo para

as tarefas serem concluídas otimizará a solução para desempenho e custo.

Certifique-se de ler a documentação do produto para "Veeam Backup and Replication" e "StorageGRID" antes de começar. A Veeam fornece calculadoras para entender o dimensionamento da infraestrutura da Veeam e os requisitos de capacidade que devem ser usados antes de dimensionar sua solução StorageGRID . Por favor, verifique sempre as configurações validadas pelo Veeam- NetApp no site do Veeam Ready Program para "Veeam Ready Object, Object imutabilidade e Repository" .

## Configuração da Veeam

### Versão recomendada

É sempre recomendável manter-se atualizado e aplicar os hotfixes mais recentes para o seu sistema Veeam Backup & Replication 12 ou 12,1. Atualmente, recomendamos, no mínimo, a instalação do Veeam 12 patch P20230718.

### S3 Configuração do repositório

Um repositório de backup com escalabilidade horizontal (SOBR) é a camada de capacidade do storage de objetos S3. A camada de capacidade é uma extensão do repositório principal que fornece períodos de retenção de dados mais longos e uma solução de storage de baixo custo. A Veeam oferece a capacidade de fornecer imutabilidade por meio da API S3 Object Lock. O Veeam 12 pode usar vários buckets em um repositório com escalabilidade horizontal. O StorageGRID não tem um limite para o número de objetos ou capacidade em um único bucket. O uso de vários buckets pode melhorar o desempenho ao fazer backup de conjuntos de dados muito grandes, onde os dados de backup podem chegar à escala de petabytes em objetos.

A limitação de tarefas simultâneas pode ser necessária dependendo do dimensionamento de sua solução e requisitos específicos. As configurações padrão especificam um slot de tarefa do repositório para cada núcleo da CPU e para cada slot de tarefa um limite de slot de tarefa concorrente de 64. Por exemplo, se o servidor tiver 2 núcleos de CPU, um total de 128 threads simultâneos será usado para o armazenamento de objetos. Isso inclui o PUT, GET e Batch Delete. É recomendável selecionar um limite conservador para os slots de tarefa para começar e ajustar esse valor depois que os backups da Veeam atingirem um estado estável de novos backups e expirarem os dados de backup. Trabalhe com sua equipe de conta do NetApp para dimensionar o sistema StorageGRID de forma adequada para atender às janelas de tempo e desempenho desejados. Ajustar o número de slots de tarefa e o limite de tarefas por slot pode ser necessário para fornecer a solução ideal.

### Configuração do trabalho de cópia de segurança

As tarefas de backup da Veeam podem ser configuradas com diferentes opções de tamanho de bloco que devem ser consideradas cuidadosamente. O tamanho padrão do bloco é 1MB e, com as eficiências de storage oferecidas pela Veeam, a deduplicação e a compactação criam tamanhos de objetos de aproximadamente 500KB TB para o backup completo inicial e objetos 100-200kB TB para as tarefas incrementais. Podemos aumentar bastante o desempenho e reduzir os requisitos do armazenamento de objetos escolhendo um tamanho maior de bloco de backup. Embora o tamanho de bloco maior faça grandes melhorias no desempenho de armazenamento de objetos, ele vem com o custo do requisito de capacidade de storage primário potencialmente maior devido à performance de eficiência de storage reduzida. Recomenda-se que as tarefas de backup sejam configuradas com um tamanho de bloco 4MB que cria aproximadamente 2MB objetos para os backups completos e tamanhos de objetos 700kB-1MB para incrementos. Os clientes podem até mesmo configurar tarefas de backup usando tamanho de bloco de 8 MB, que podem ser habilitadas com a ajuda do suporte da Veeam.

A implementação de backups imutáveis faz uso do bloqueio de objetos S3 no armazenamento de objetos. A

opção imutabilidade gera um número maior de solicitações para o armazenamento de objetos para listar e reter atualizações nos objetos.

À medida que as retenções de cópia de segurança expiram, os trabalhos de cópia de segurança processarão a eliminação de objetos. A Veeam envia as solicitações de exclusão para o armazenamento de objetos em solicitações de exclusão multiobjetos de 1000 objetos por solicitação. Para soluções pequenas, isso pode precisar ser ajustado para reduzir o número de objetos por solicitação. A redução desse valor terá o benefício adicional de distribuir mais uniformemente as solicitações de exclusão entre os nós no sistema StorageGRID. Recomenda-se usar os valores na tabela abaixo como ponto de partida para configurar o limite de exclusão de vários objetos. Multiplique o valor na tabela pelo número de nós para o tipo de dispositivo escolhido para obter o valor para a configuração no Veeam. Se este valor for igual ou superior a 1000, não será necessário ajustar o valor predefinido. Se esse valor precisar ser ajustado, trabalhe com o suporte da Veeam para fazer a mudança.

| Modelo do aparelho | S3MultiObjectDeleteLimit PB por nó |
|---|---|
| SG5712 | 34 |
| SG5760 | 75 |
| SG6060 | 200 |

Trabalhe com sua equipe de conta do NetApp para obter a configuração recomendada com base em suas necessidades específicas. As recomendações de configurações da Veeam incluem:

- Tamanho do bloco de trabalho de backup: 4MB
- Limite de slot de tarefa SOBR 2-16
- Limite de exclusão de objetos múltiplos: 34-1000

## Configuração do StorageGRID

### Versão recomendada

O NetApp StorageGRID 11.9 ou 12.0 com o hotfix mais recente são as versões recomendadas para implantações do Veeam. É sempre recomendável manter-se atualizado e aplicar os últimos hotfixes para seu sistema StorageGRID .

### Balanceador de carga e configuração de endpoint S3

A Veeam exige que o endpoint seja conetado somente via HTTPS. Uma conexão não criptografada não é suportada pela Veeam. O certificado SSL pode ser um certificado auto-assinado, uma autoridade de certificação privada confiável ou uma autoridade de certificação pública confiável. Para garantir o acesso contínuo ao repositório S3, é recomendável usar pelo menos dois balanceadores de carga em uma configuração de HA. Os balanceadores de carga podem ser um serviço de balanceador de carga integrado fornecido pela StorageGRID localizado em cada nó de administrador e nó de gateway ou solução de terceiros, como F5, Kemp, HAproxy, Loadbalancer.org, etc. o uso de um balanceador de carga StorageGRID fornecerá a capacidade de definir classificadores de tráfego (regras de QoS) que podem priorizar a carga de trabalho da Veeam ou limitar a não impactar cargas de trabalho de alta prioridade no sistema StorageGRID.

### S3 balde

StorageGRID é um sistema de armazenamento multilocatário seguro. É recomendável criar um locatário dedicado para a carga de trabalho do Veeam. Uma cota de armazenamento pode ser atribuída

opcionalmente. Como prática recomendada, habilite “usar fonte de identidade própria”. Proteja o usuário de gerenciamento raiz do locatário com uma senha apropriada. O Veeam Backup 12 exige consistência forte para buckets S3. O StorageGRID oferece várias opções de consistência configuradas no nível do bucket. Para implantações em vários locais com o Veeam acessando os dados de vários locais, selecione “strong-global”. Se os backups e restaurações do Veeam ocorrerem apenas em um único site, o nível de consistência deverá ser definido como “strong-site”. Para obter mais informações sobre os níveis de consistência do bucket, consulte o "documentação" . Para usar o StorageGRID para backups de imutabilidade do Veeam, o S3 Object Lock deve ser habilitado globalmente e configurado no bucket durante a criação do bucket.

**Gerenciamento de ciclo de vida**

O StorageGRID é compatível com replicação e codificação de apagamento para proteção no nível de objeto em nós e sites da StorageGRID. A codificação de apagamento requer pelo menos um tamanho de objeto 200kB. O tamanho padrão do bloco para Veeam de 1MB produz tamanhos de objetos que geralmente podem estar abaixo desse tamanho mínimo recomendado de 200kB MB após as eficiências de storage da Veeam. Para o desempenho da solução, não é recomendável usar um perfil de codificação de apagamento abrangendo vários sites, a menos que a conetividade entre os sites seja suficiente para não adicionar latência ou restringir a largura de banda do sistema StorageGRID. Em um sistema StorageGRID multi-site, a regra ILM pode ser configurada para armazenar uma única cópia em cada local. Para uma durabilidade máxima, uma regra poderia ser configurada para armazenar uma cópia codificada de apagamento em cada local. O uso de duas cópias locais para os servidores do Veeam Backup é a implementação mais recomendada para essa carga de trabalho.

**Excluir desempenho**

O Veeam fornece ajuste de taxa de solicitação de exclusão e agendamento do processo de exclusão de backup. Para ajustar ainda mais o desempenho da exclusão, você pode desabilitar as exclusões síncronas e deixar que o scanner ILM gerencie a eventual exclusão de objetos.

**Etapas para desabilitar exclusões síncronas**

1. Abra o StorageGRID Grid Manager.
2. No canto superior direito, selecione o ponto de interrogação e depois Documentação da API.
3. No canto superior direito, clique no link da página Documentação da API privada.
4. Expanda ilm-advanced.
5. Selecione OBTER ilm-advanced.
6. Selecione Experimentar e depois Executar.
7. Verifique o resultado da resposta.

    a. Se os valores forem nulos, significa que os valores padrão do ilm-advanced estão em uso.

    b. Se os valores não forem nulos, significa que os valores avançados do ILM personalizado estão em uso. Copie toda a saída depois de "data" :, começando com { até o segundo ao último }.

        i. Salve-o em algum editor de texto.

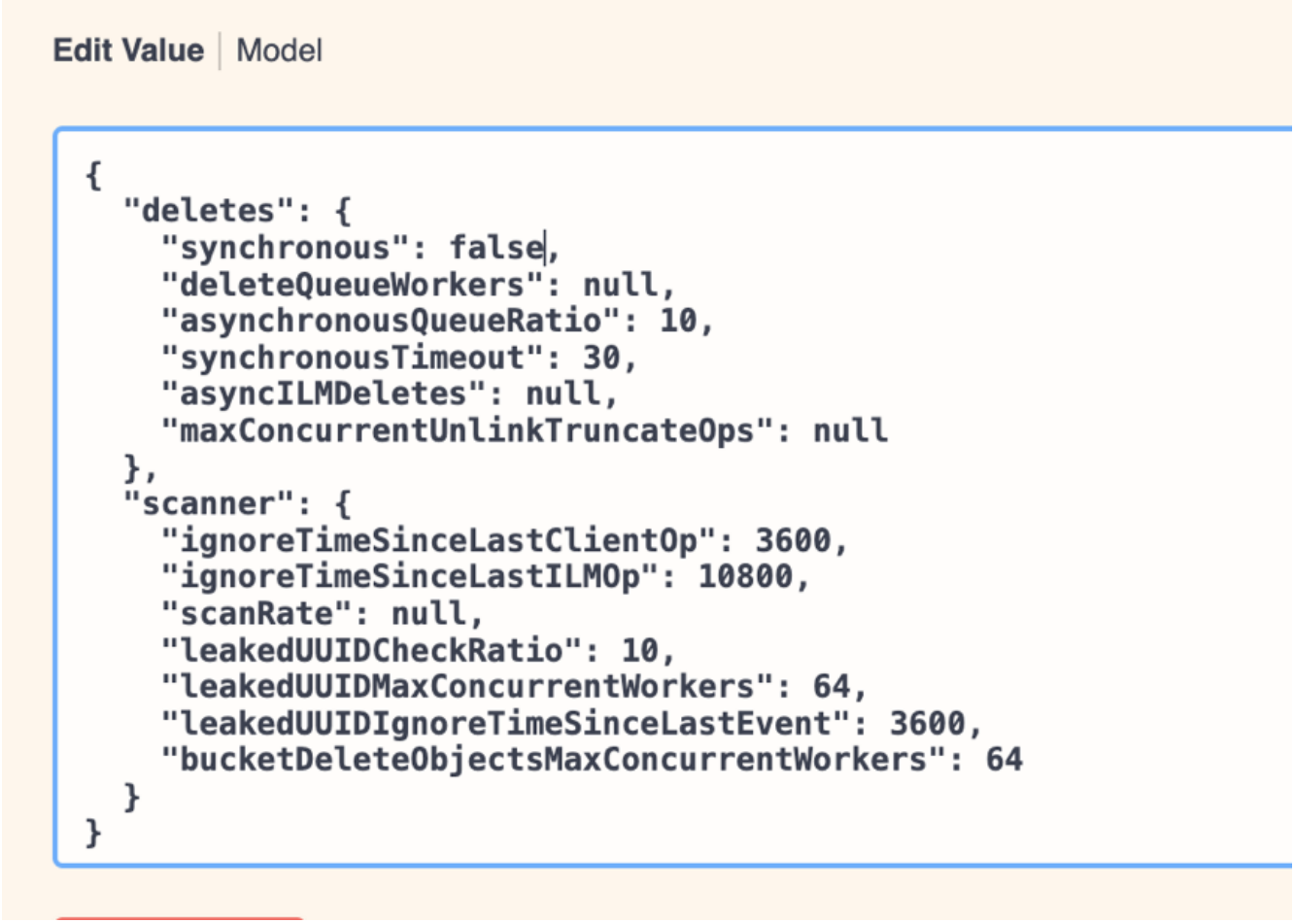
        Exemplo de resposta:

**Response body**


```
{
  "responseTime": "2025-09-19T15:01:28.142Z",
  "status": "success",
  "apiVersion": "4.2",
  "data": {
    "deletes": {
      "synchronous": null,
      "deleteQueueWorkers": null,
      "asynchronousQueueRatio": null,
      "synchronousTimeout": null,
      "asyncILMDeletes": null,
      "maxConcurrentUnlinkTruncateOps": null
    },
    "scanner": {
      "ignoreTimeSinceLastClientOp": null,
      "ignoreTimeSinceLastILMOp": null,
      "scanRate": null,
      "leakedUUIDCheckRatio": null,
      "leakedUUIDMaxConcurrentWorkers": null,
      "leakedUUIDIgnoreTimeSinceLastEvent": null,
      "bucketDeleteObjectsMaxConcurrentWorkers": null
    }
  }
}
```


8. Selecione PUT ilm-advanced.
9. Selecione Experimentar para começar a editar o corpo da API.
    a. Por padrão, o corpo da API conterá valores padrão e não quaisquer valores personalizados que tenham sido configurados anteriormente. É por isso que é MUITO importante executar os passos 5 a 7.
10. Se valores não padrão forem encontrados nas etapas 5 a 7, substitua o corpo da API pela saída salva na etapa 7. . Caso contrário, se os valores forem nulos nas etapas 5 a 7, deixe o corpo da API como está.
11. Ajuste os seguintes parâmetros na caixa do corpo da API:
    a. Defina o valor síncrono como falso.

       Exemplo de texto do corpo da API:

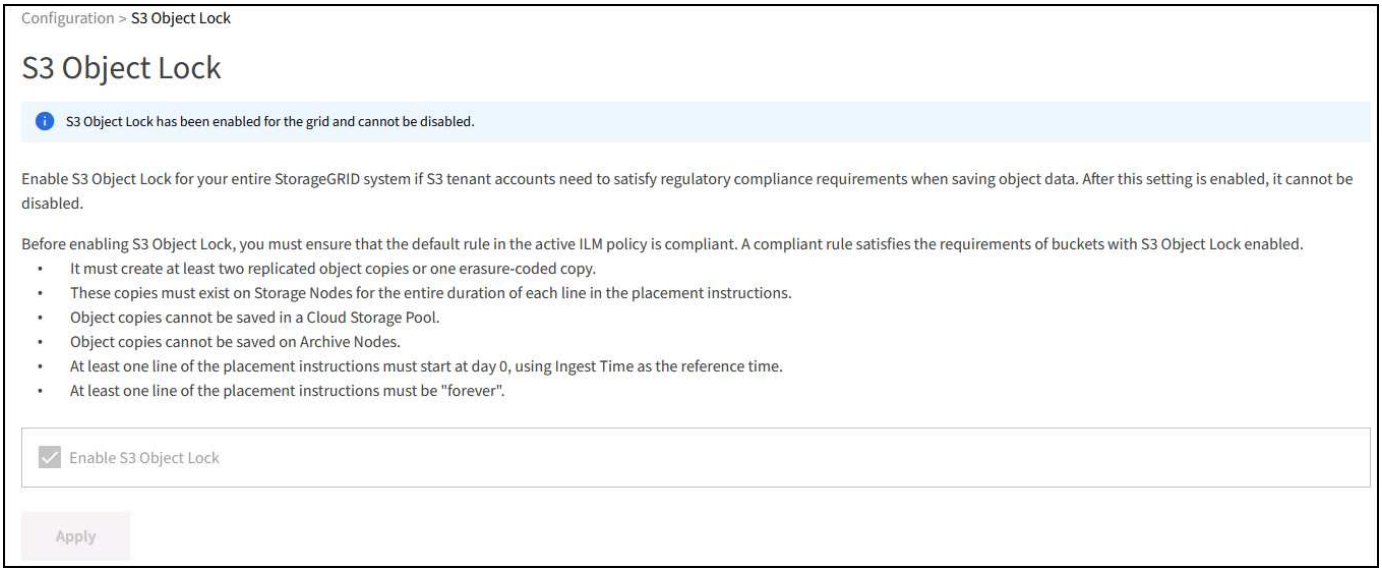



12. Após a conclusão, selecione Executar

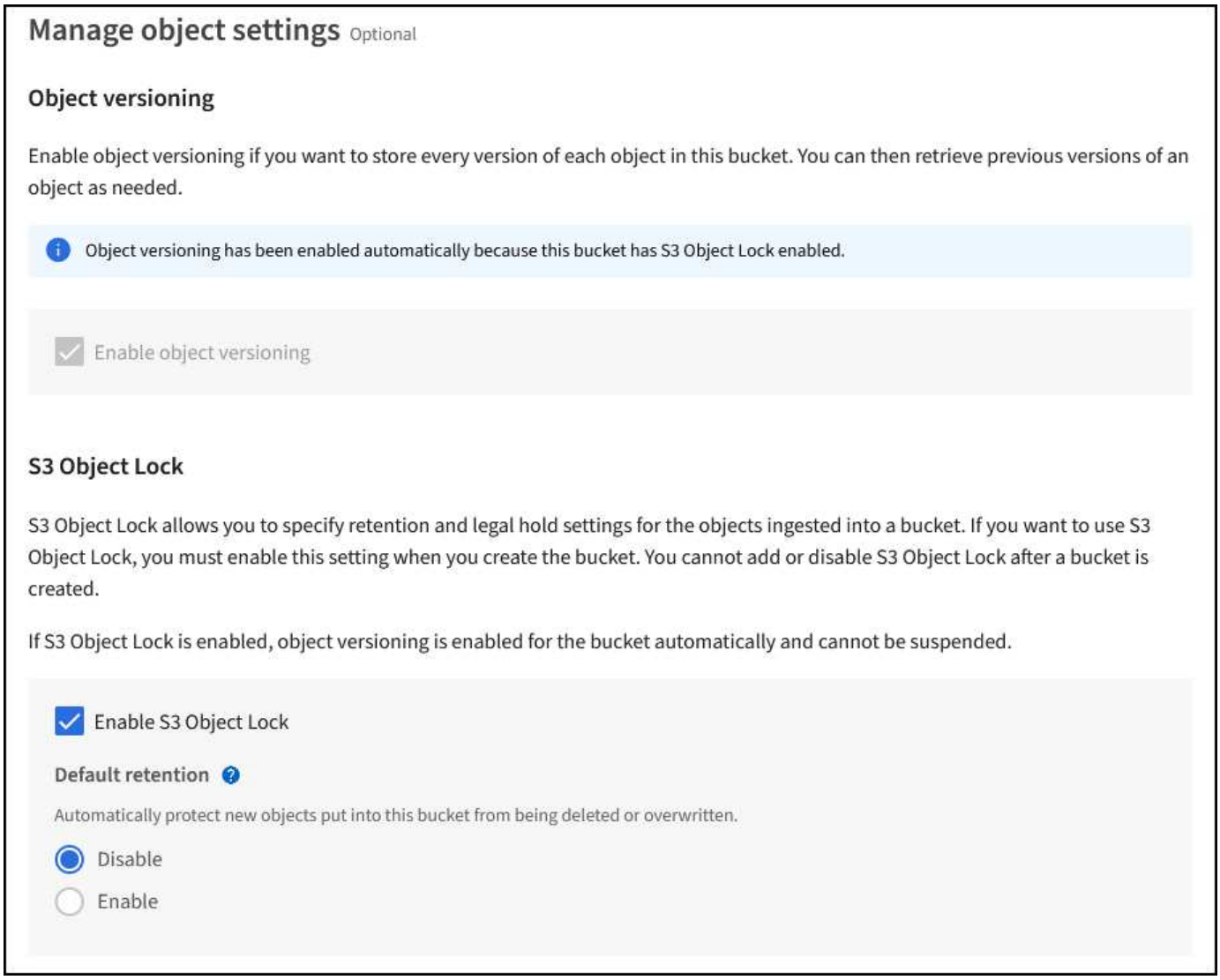
## Pontos-chave de implementação

### StorageGRID

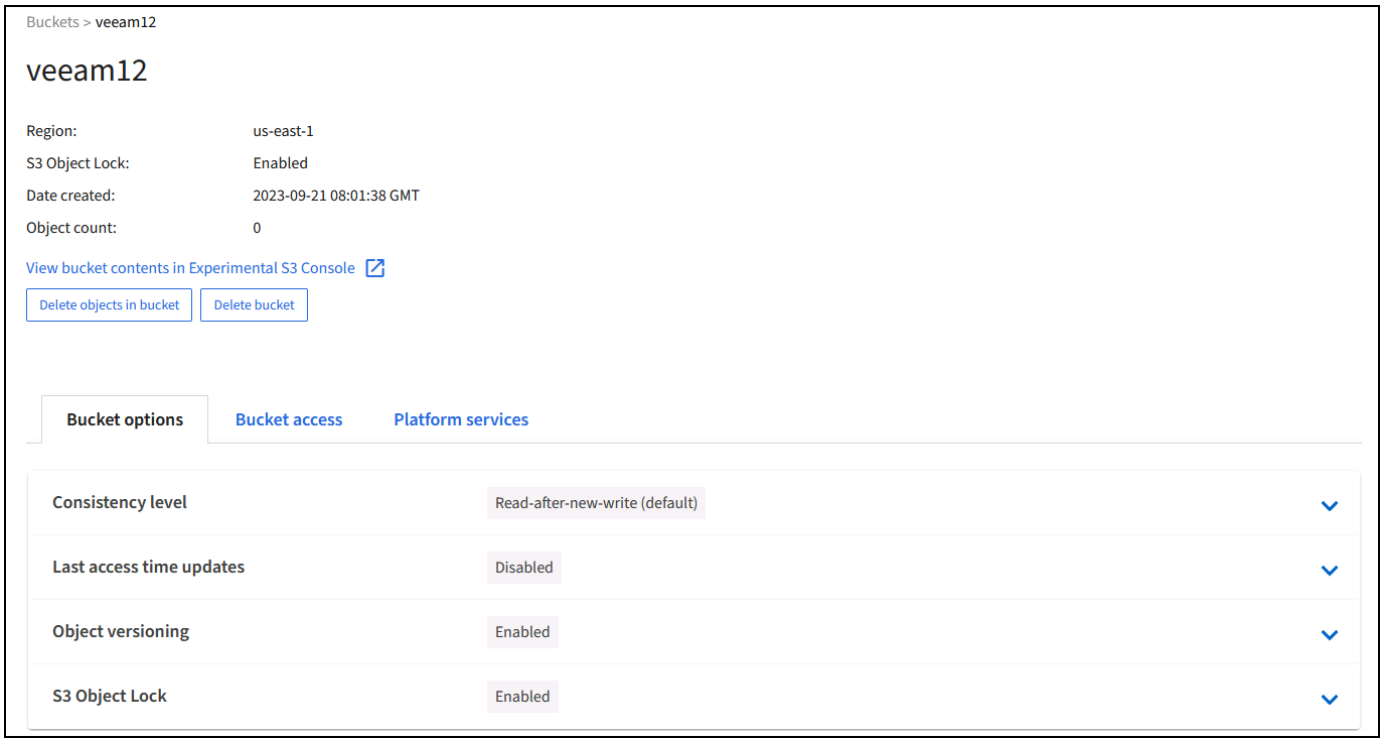
Certifique-se de que o bloqueio de objetos está ativado no sistema StorageGRID se a imutabilidade for necessária. Encontre a opção na IU de gerenciamento em Configuration/S3 Object Lock.



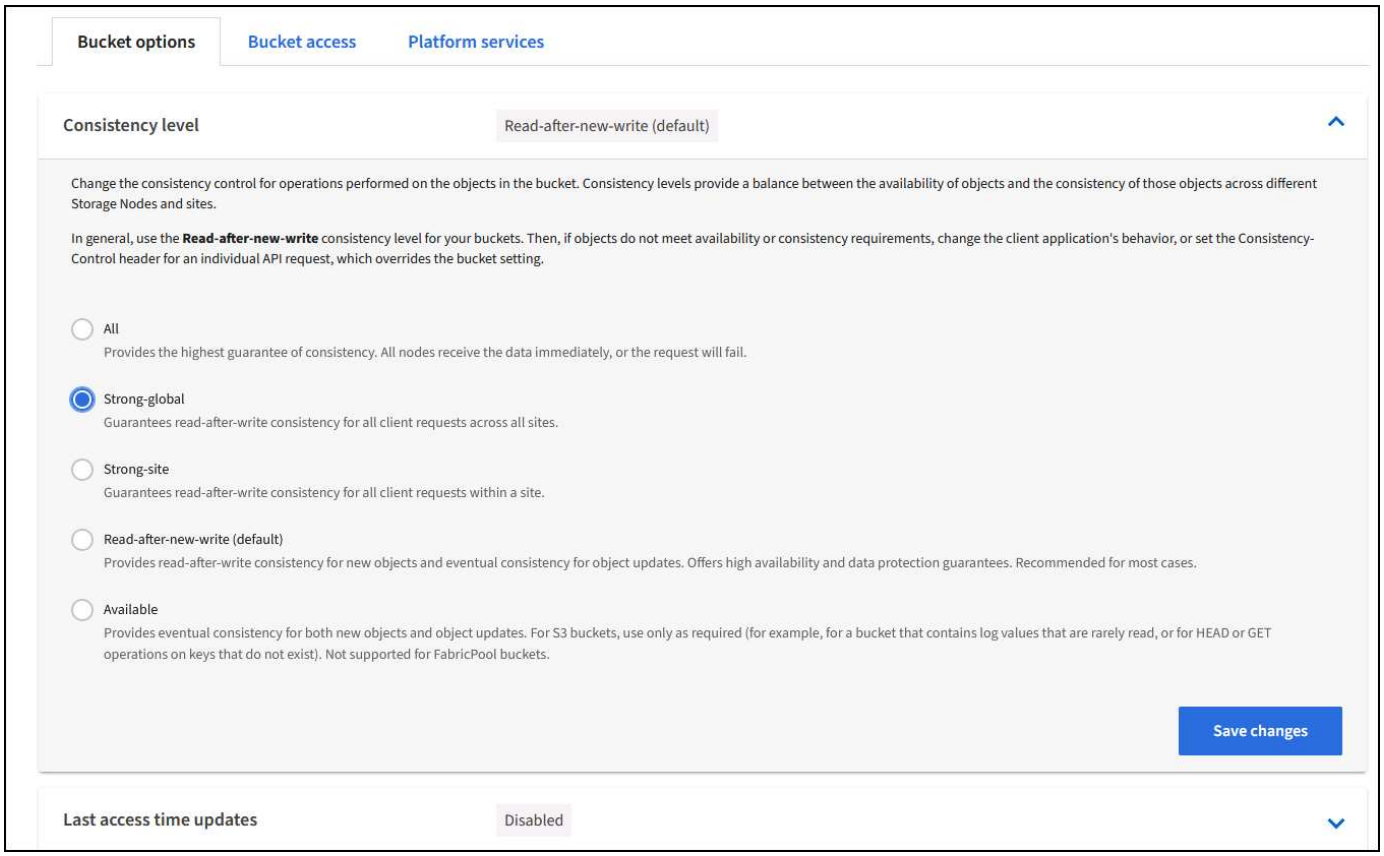
Ao criar o bucket, selecione "Ativar bloqueio de objetos S3" se esse bucket for usado para backups de imutabilidade. Isso habilitará automaticamente o controle de versão do bucket. Deixe a retenção padrão desativada, pois a Veeam definirá a retenção de objetos explicitamente. Controle de versão e bloqueio de objetos S3 não devem ser selecionados se a Veeam não estiver criando backups imutáveis.




Quando o bucket for criado, vá para a página de detalhes do bucket criado. Selecione o nível de consistência.



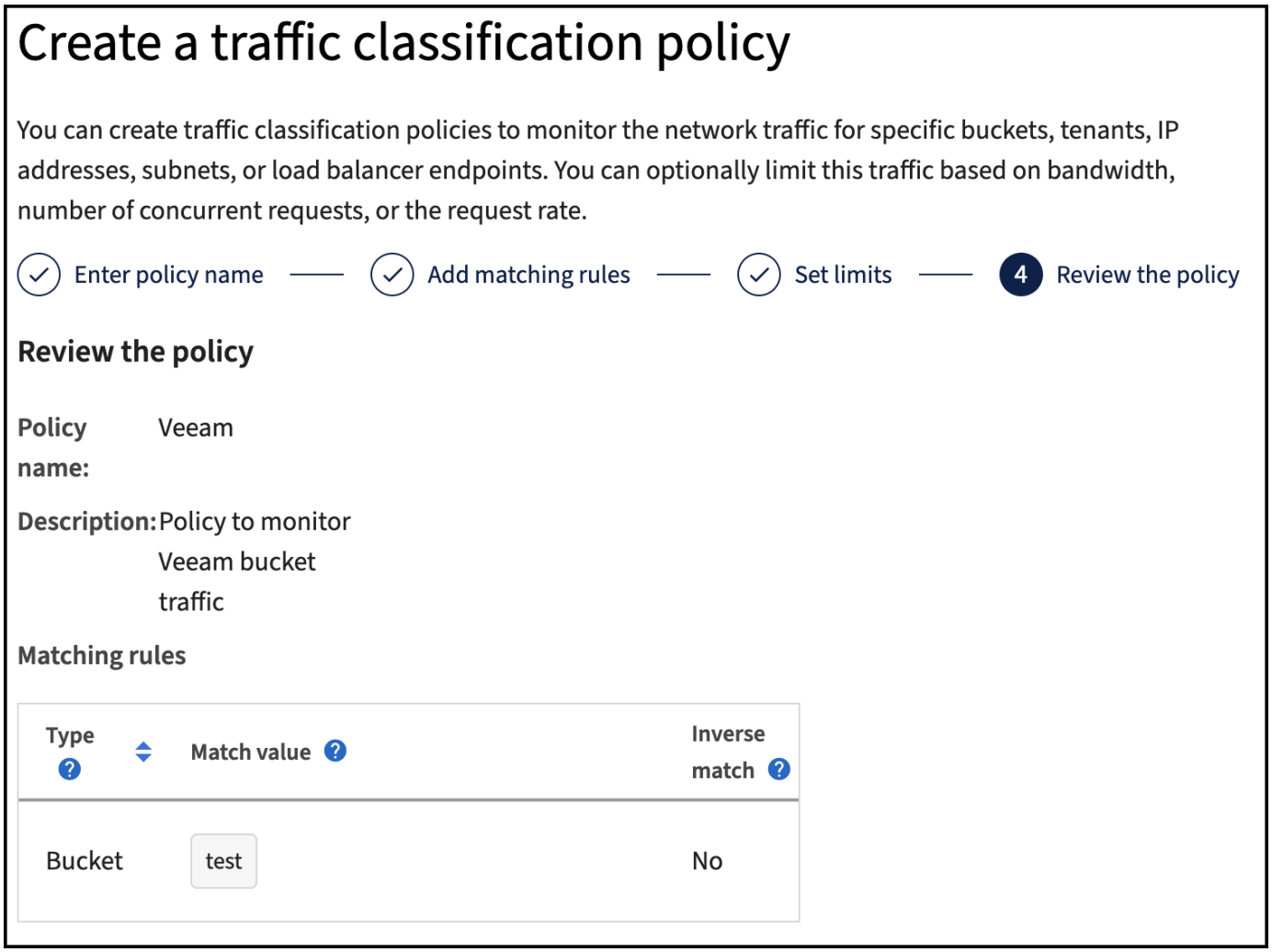

A Veeam requer uma forte consistência para buckets do S3. Então, para implantações em vários locais com a Veeam acessando os dados de vários locais, selecione "forte global". Se os backups e restaurações da Veeam acontecerem apenas em um único local, o nível de consistência deve ser definido como "local forte". Salve as alterações.




O StorageGRID fornece um serviço de balanceador de carga integrado em todos os nós de administração e

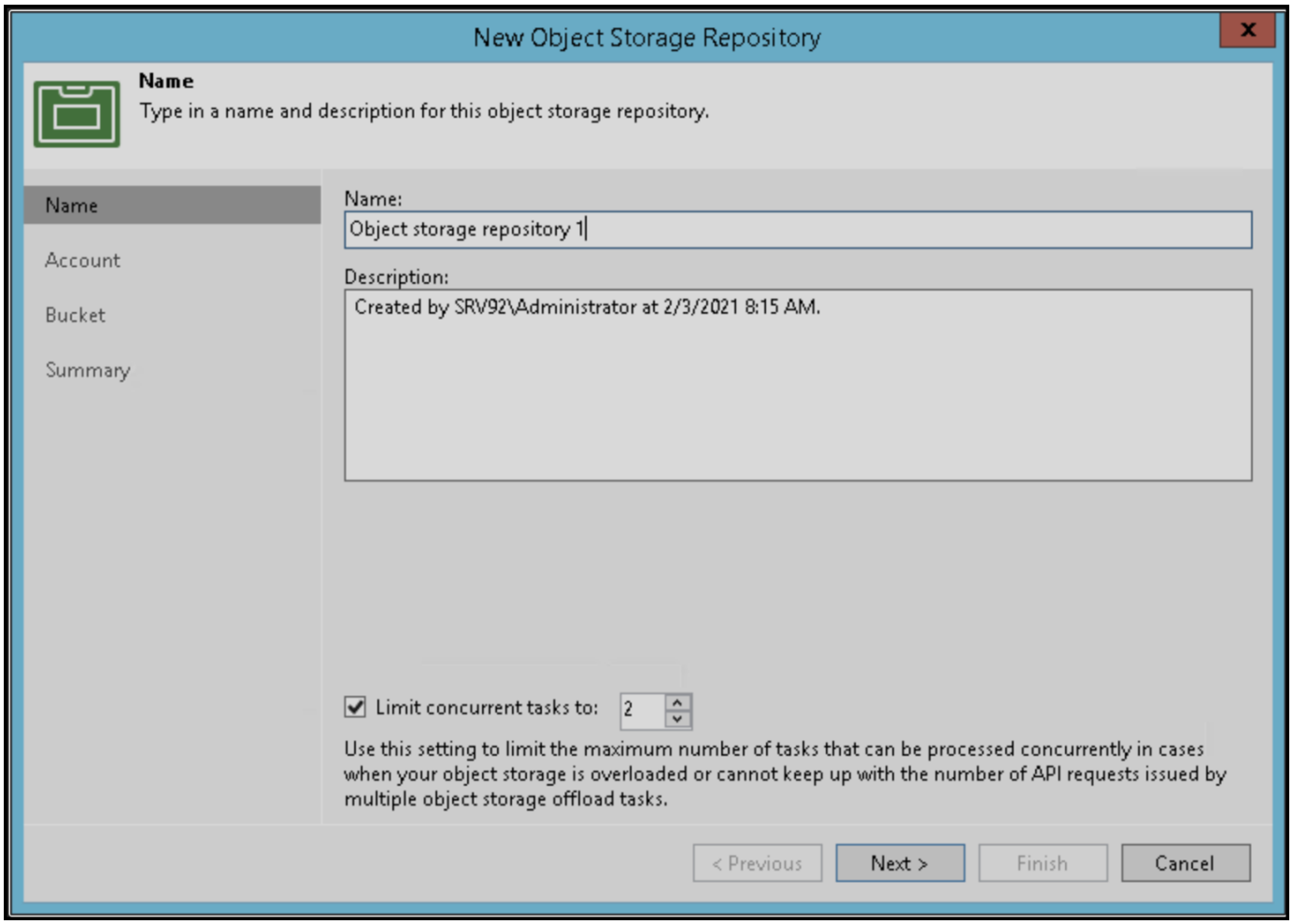
de gateway dedicados. Uma das muitas vantagens de usar este balanceador de carga é a capacidade de configurar as políticas de classificação de tráfego (QoS). Embora eles sejam usados principalmente para limitar o impactos de aplicativos em outras cargas de trabalho de clientes ou priorizar uma carga de trabalho sobre outras, eles também fornecem um bônus de coleta de métricas adicionais para ajudar no monitoramento.

No separador de configuração, selecione "classificação de tráfego" e crie uma nova política. Nomeie a regra e selecione o(s) intervalo(s) ou o locatário como o tipo. Introduza o(s) nome(s) do(s) bucket(s) ou inquilino(s). Se a QoS for necessária, defina um limite, mas para a maioria das implementações, queremos apenas adicionar os benefícios de monitoramento que isso proporciona, portanto, não defina um limite.

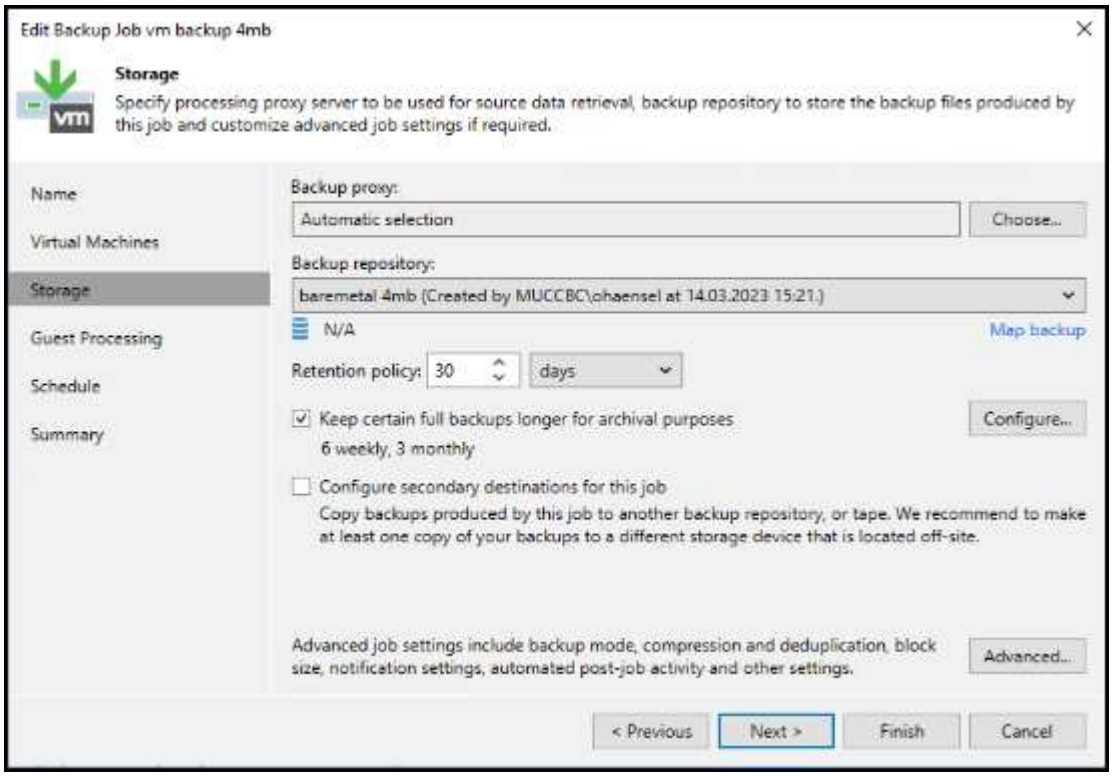



### Veeam

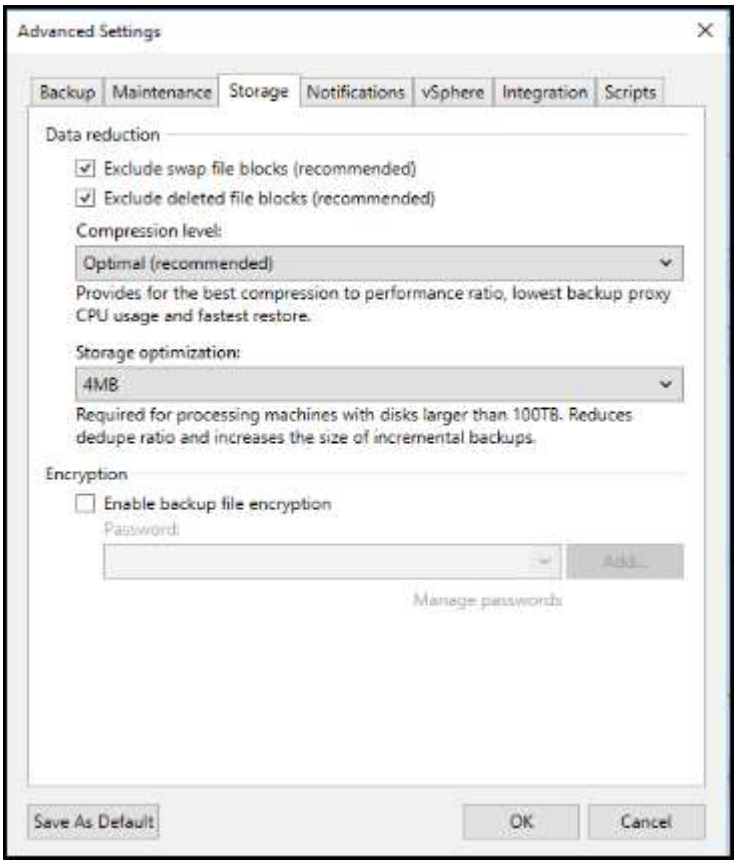
Dependendo do modelo e da quantidade de dispositivos StorageGRID, pode ser necessário selecionar e configurar um limite para o número de operações simultâneas no bucket.




Siga a documentação da Veeam sobre a configuração da tarefa de backup no console da Veeam para iniciar o assistente. Depois de adicionar VMs, selecione o repositório SOBR.



Clique em Configurações avançadas e altere as configurações de otimização de armazenamento para 4 MB ou mais. A compactação e a deduplicação devem ser habilitadas. Altere as configurações do convidado de acordo com seus requisitos e configure o agendamento do trabalho de backup.




## Monitorização do StorageGRID

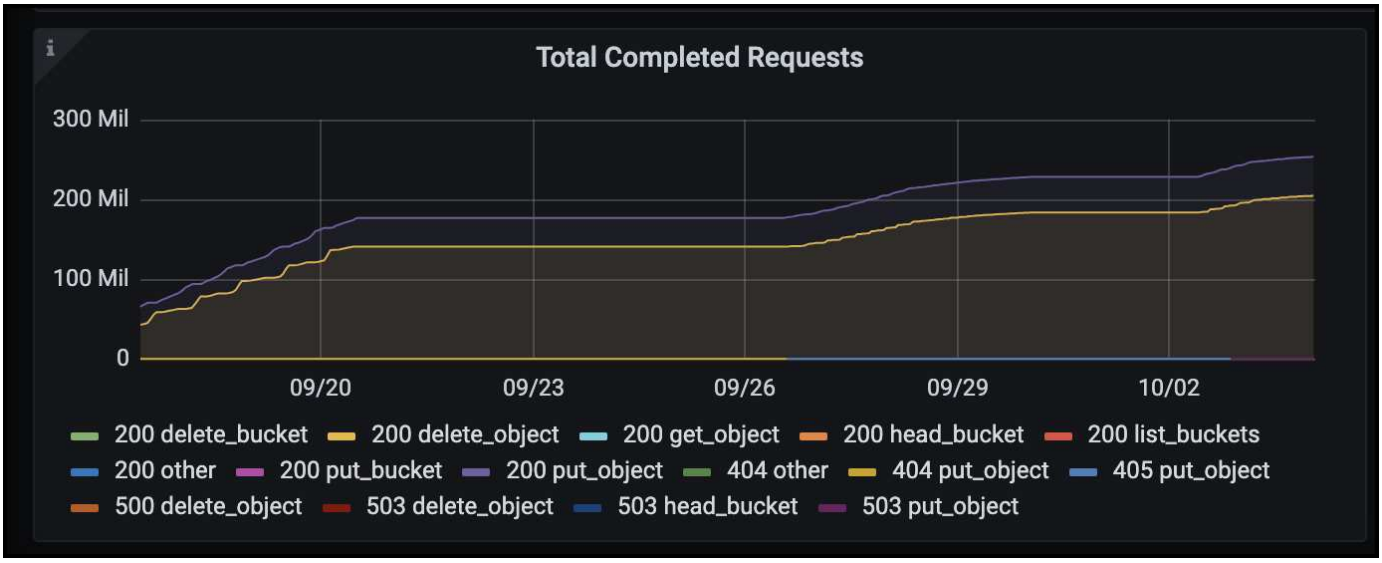
Para ter uma visão completa de como a Veeam e o StorageGRID estão funcionando juntos, você precisará esperar até que o tempo de retenção dos primeiros backups expire. Até esse ponto, a carga de trabalho da Veeam consiste principalmente em operações PUT e não ocorreram exclusões. Uma vez que os dados de backup expiram e as limpezas estão ocorrendo, você pode agora ver o uso consistente completo no armazenamento de objetos e ajustar as configurações no Veeam, se necessário.

O StorageGRID fornece gráficos convenientes para monitorar o funcionamento do sistema localizado na página métricas do separador suporte. Os principais painéis a serem analisados serão a Visão geral do S3, ILM e a Política de classificação de tráfego, se uma política foi criada. No painel Visão geral do S3, você encontrará informações sobre as taxas de operação, latências e respostas de solicitações do S3.

Olhando para as taxas do S3 e as solicitações ativas, você pode ver quanto da carga cada nó está lidando e o número total de solicitações por tipo.


O gráfico de duração média mostra o tempo médio que cada nó está tomando para cada tipo de solicitação. Esta é a latência média da solicitação e pode ser um bom indicador de que ajustes adicionais podem ser necessários, ou há espaço para o sistema StorageGRID assumir mais carga.


No gráfico Total de solicitações concluídas, você pode ver as solicitações por tipo e códigos de resposta. Se você vir respostas diferentes de 200 (OK) para as respostas, isso pode indicar um problema como o sistema StorageGRID está recebendo fortemente carregado enviando respostas 503 (lento) e alguma sintonização adicional pode ser necessária, ou chegou a hora de expandir o sistema para a carga aumentada.




No Painel ILM, você pode monitorar o desempenho de exclusão do seu sistema StorageGRID. O StorageGRID usa uma combinação de exclusões síncronas e assíncronas em cada nó para tentar otimizar o desempenho geral de todas as solicitações.


Com uma Política de classificação de tráfego, podemos visualizar métricas sobre a taxa de transferência de solicitação do balanceador de carga, taxas, duração, bem como os tamanhos de objeto que a Veeam está enviando e recebendo.


## Onde encontrar informações adicionais

Para saber mais sobre as informações descritas neste documento, consulte os seguintes documentos e/ou sites:

- "Documentação do produto NetApp StorageGRID"
- "Veeam Backup and Replication"

# Configure a fonte de dados do Dremio com o StorageGRID

*Por Angela Cheng*

O Dremio dá suporte a uma variedade de fontes de dados, incluindo armazenamento de objetos baseado na nuvem ou no local. Você pode configurar o Dremio para usar o StorageGRID como fonte de dados de armazenamento de objetos.

## Configurar a fonte de dados do Dremio

### Pré-requisitos

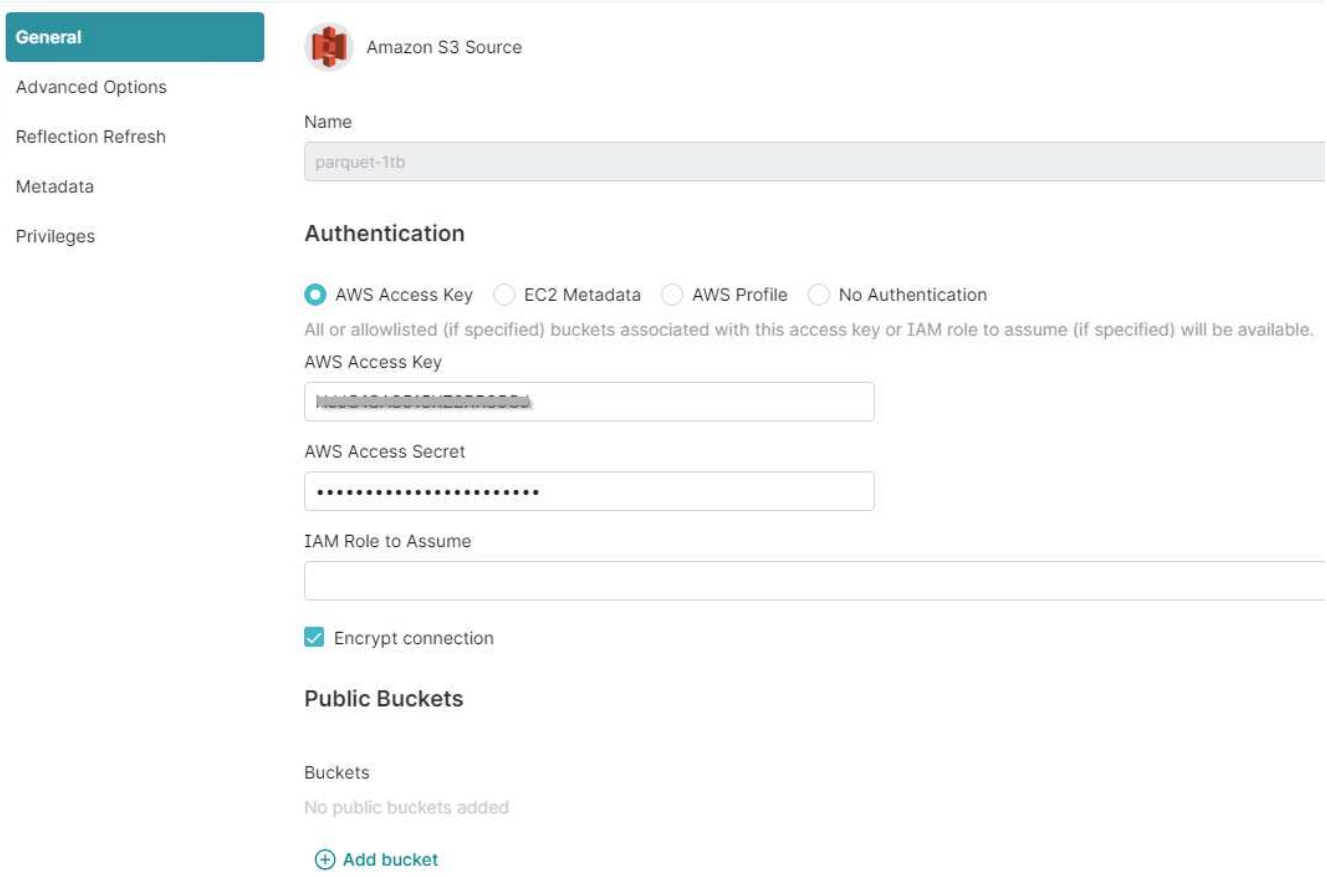
- Um URL de endpoint do StorageGRID S3, um ID de chave de acesso do locatário S3 e chave de acesso secreta.
- Recomendação de configuração do StorageGRID: Desativar a compactação (desativada por padrão). Dremio usa o intervalo de bytes get para buscar diferentes intervalos de bytes dentro do mesmo objeto simultaneamente durante a consulta. O tamanho típico para solicitações de intervalo de bytes é 1MB. O objeto comprimido degrada a gama de bytes OBTENHA desempenho.

### Guia de Dremio

"Conetando ao Amazon S3 - Configurando o armazenamento compatível com S3".

## Instrução

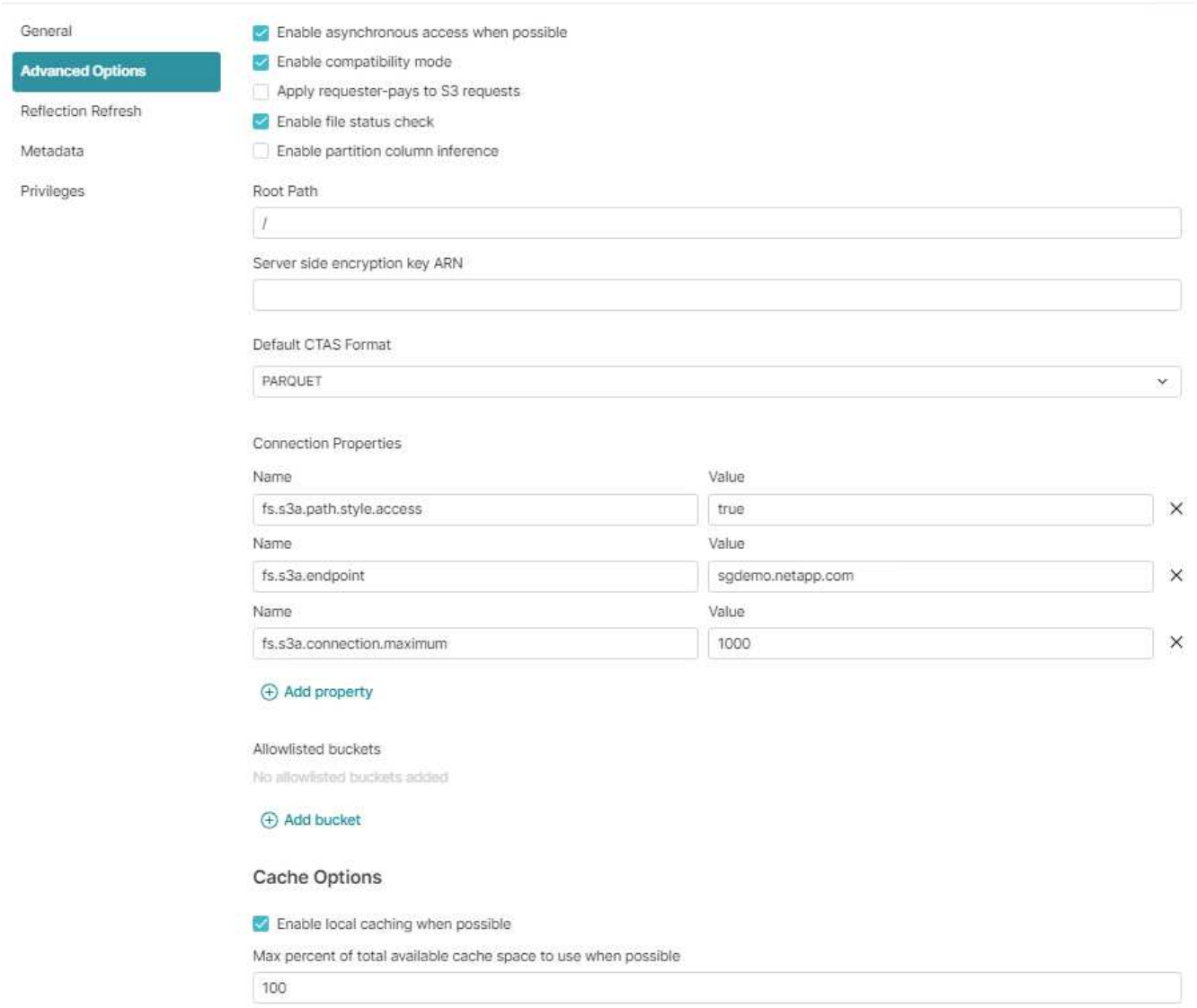
1. Na página Datasets do Dremio, clique em assinar para adicionar uma fonte, selecione 'Amazon S3'.
2. Insira um nome para esta nova fonte de dados, ID da chave de acesso ao locatário do StorageGRID S3 e chave de acesso secreto.
3. Marque a caixa 'criptografar conexão' se estiver usando https para conexão com o endpoint StorageGRID S3. Se estiver usando CA cert autoassinado para este endpoint S3, siga a instrução de guia Dremio para adicionar este CA cert no servidor Dremio <JAVA_HOME>/jre/lib/security **sample screenshot**



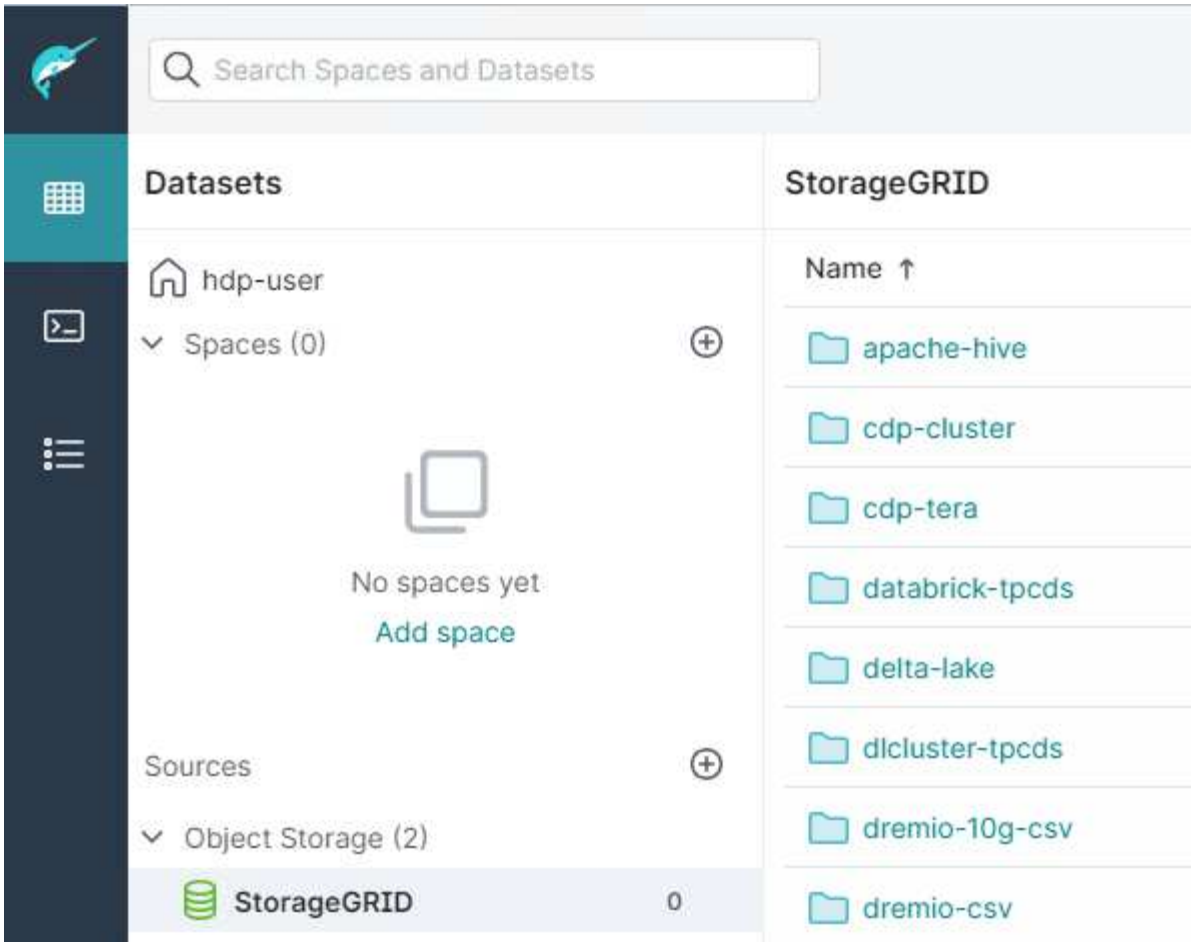

4. Clique em "Opções avançadas", verifique "Ativar modo de compatibilidade"
5. Em Propriedades de conexão, clique em Adicionar propriedades e adicione essas s3a propriedades.
6. fs.s3a.connection.o padrão máximo é 100. Se os conjuntos de dados do S3 incluírem ficheiros Parquet grandes com 100 ou mais colunas, tem de introduzir um valor superior a 100. Consulte o guia Dremio para obter esta definição.

| Nome | Valor |
| --- | --- |
| fs.s3a.endpoint | _ Endpoint do cliente StorageGRID S3:port>_ |
| fs.s3a.path.style.access | verdadeiro |
| fs.s3a.connection.maximum | _ valor de cliente maior que 100>_ |

**Captura de tela de amostra**




7. Configure outras opções do Dremio de acordo com os requisitos da sua organização ou aplicação.
8. Clique no botão Salvar para criar esta nova fonte de dados.
9. Depois que a fonte de dados StorageGRID for adicionada com sucesso, uma lista de buckets será exibida no painel esquerdo. * Captura de tela de amostra*




# NetApp StorageGRID com GitLab

*Por Angela Cheng*

A NetApp testou o StorageGRID com o GitLab. Veja exemplo de configuração do GitLab abaixo. "Guia de configuração de armazenamento de objetos GitLab"Consulte para obter detalhes.

## Exemplo de conexão de armazenamento de objetos

Para instalações do pacote Linux, este é um exemplo `connection` da configuração na forma consolidada. Edite `/etc/gitlab/gitlab.rb` e adicione as seguintes linhas, substituindo os valores desejados:

```
# 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'
```

## Informações sobre direitos autorais



## Informações sobre marcas comerciais