Guide agli strumenti e alle applicazioni

StorageGRID solutions and resources

NetApp
December 10, 2025

Sommario

Guide agli strumenti e alle applicazioni

USA il connettore S3A di Cloudera Hadoop con StorageGRID

Di Angela Cheng

Hadoop è da tempo la preferita dai data scientist. Hadoop consente l'elaborazione distribuita di grandi set di dati tra cluster di computer utilizzando semplici framework di programmazione. Hadoop è stato progettato per scalare da singoli server a migliaia di macchine, con ogni macchina in possesso di calcolo e storage locali.

Perché utilizzare S3A per i flussi di lavoro Hadoop?

Con la crescita del volume di dati nel tempo, l'approccio all'aggiunta di nuove macchine con il proprio calcolo e storage è diventato inefficiente. La scalabilità lineare crea delle sfide per l'utilizzo efficiente delle risorse e la gestione dell'infrastruttura.

Per affrontare queste sfide, il client Hadoop S3A offre i/o dalle performance elevate rispetto allo storage a oggetti S3. L'implementazione di un workflow Hadoop con S3A consente di sfruttare lo storage a oggetti come repository di dati e consente di separare calcolo e storage, il che consente di scalare calcolo e storage in modo indipendente. Il disaccoppiamento di calcolo e storage consente inoltre di dedicare la giusta quantità di risorse per i processi di calcolo e di fornire capacità in base alle dimensioni del set di dati. Pertanto, è possibile ridurre il TCO complessivo per i flussi di lavoro Hadoop.

Configurare S3A Connector per l'utilizzo di StorageGRID

Prerequisiti

- Un URL endpoint StorageGRID S3, una chiave di accesso s3 tenant e una chiave segreta per il test di connessione Hadoop S3A.
- Un cluster Cloudera e un'autorizzazione root o sudo a ciascun host del cluster per installare il pacchetto Java.

Ad aprile 2022, Java 11.0.14 con Cloudera 7.1.7 è stato testato rispetto a StorageGRID 11.5 e 11.6. Tuttavia, il numero di versione di Java potrebbe essere diverso al momento di una nuova installazione.

Installare il pacchetto Java

1. Controllare "Matrice di supporto di Cloudera" Per la versione JDK supportata.
2. Scaricare il "Pacchetto Java 11.x." Che corrisponde al sistema operativo del cluster Cloudera. Copiare questo pacchetto su ciascun host del cluster. In questo esempio, il pacchetto rpm viene utilizzato per CentOS.
3. Accedere a ciascun host come root o utilizzando un account con autorizzazione sudo. Eseguire le seguenti operazioni su ciascun host:
   a. Installare il pacchetto:

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

b. Controllare dove è installato Java. Se sono installate più versioni, impostare la nuova versione installata come predefinita:

```
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. Aggiungere questa riga alla fine di `/etc/profile`. Il percorso deve corrispondere al percorso della selezione precedente:

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

d. Eseguire il seguente comando per rendere effettivo il profilo:

```
source /etc/profile
```

Configurazione di Cloudera HDFS S3A

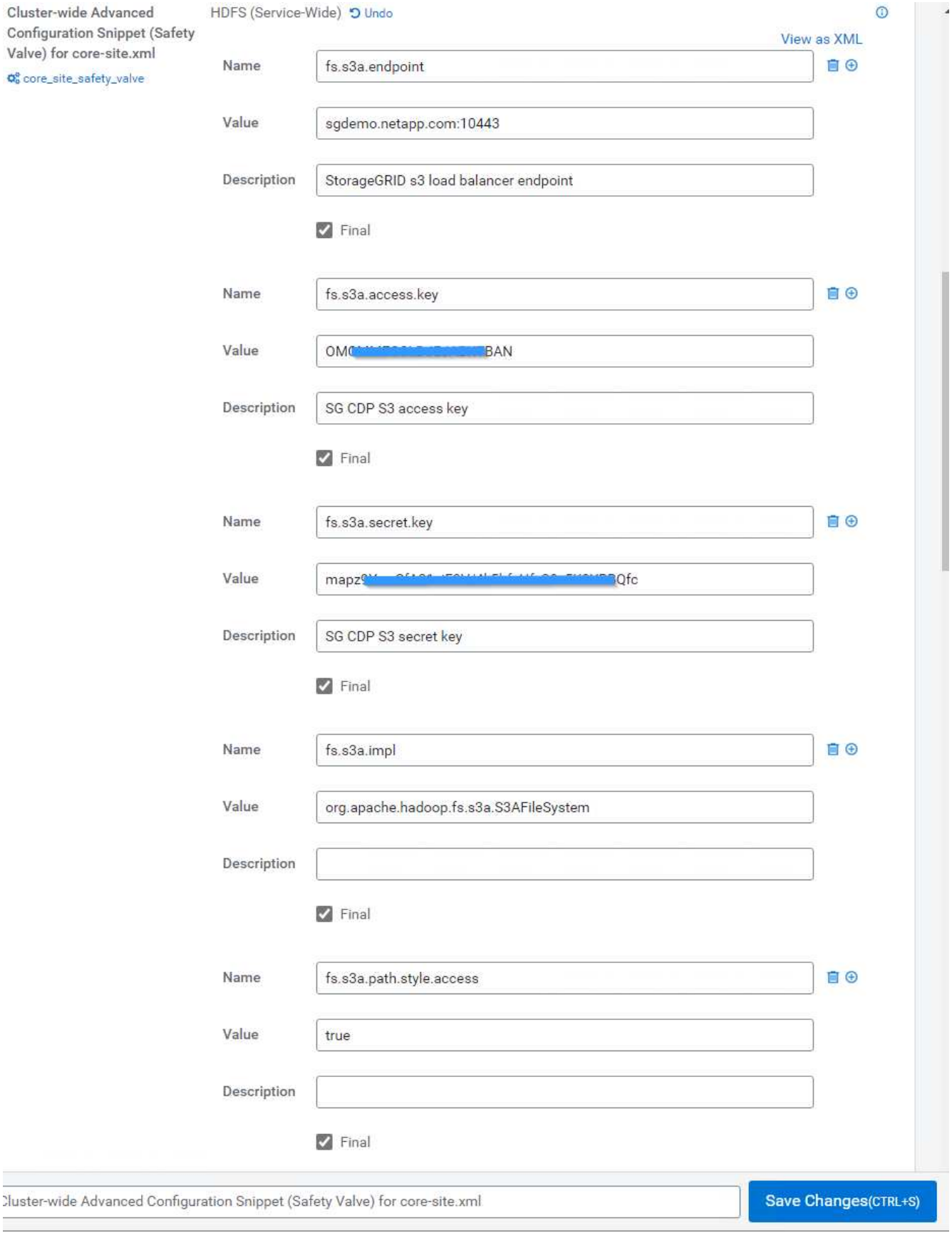
Fasi

1. Dalla GUI di Cloudera Manager, selezionare Clusters > HDFS e selezionare Configuration (Configurazione).
2. In CATEGORY (CATEGORIA), selezionare Advanced (Avanzate) e scorrere verso il basso per individuare `Cluster-wide Advanced Configuration Snippet (Safety Valve) for core-site.xml`.
3. Fare clic sul segno (+) e aggiungere le seguenti coppie di valori.

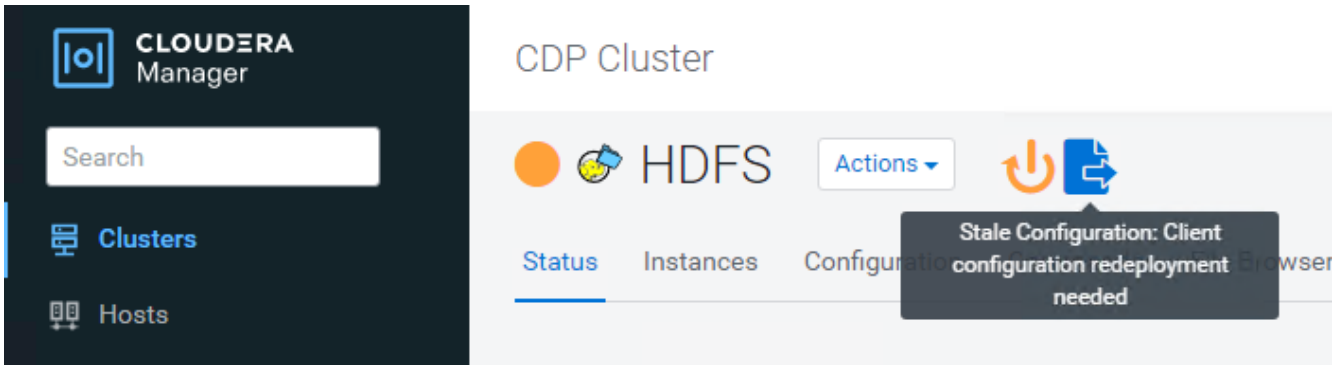
Nome	Valore
fs.s3a.access.key	*<chiave di accesso s3 tenant da StorageGRID>*
fs.s3a.secret.key	*<chiave segreta s3 tenant da StorageGRID>*
fs.s3a.connection.ssl.enabled	[true o false] (l'impostazione predefinita è https se questa voce non è presente)
fs.s3a.endpoint	*<StorageGRID S3 endpoint:porta>*

Nome	Valore
fs.s3a.impl	Org.apache.hadoop.fs.s3a.S3AFileSystem
fs.s3a.path.style.access	[true o false] (l'impostazione predefinita è lo stile dell'host virtuale se questa voce non è presente)

Esempio di screenshot



4. Fare clic sul pulsante Save Changes (Salva modifiche). Selezionare l'icona di configurazione obsoleta dalla barra dei menu di HDFS, selezionare Restart stale Services (Riavvia servizi obsoleti) nella pagina successiva e selezionare Restart Now (Riavvia ora).



Verificare la connessione S3A a StorageGRID

Eseguire un test di connessione di base

Accedere a uno degli host nel cluster Cloudera e immettere `hadoop fs -ls s3a://<bucket-name>/`.

Nell'esempio seguente viene utilizzato il path syle con un bucket hdfs-test preesistente e un oggetto test.

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

Risoluzione dei problemi

Scenario 1

Utilizzare una connessione HTTPS a StorageGRID e ottenere un `handshake_failure` errore dopo un timeout di 15 minuti.

Motivo: versione precedente di JRE/JDK che utilizza una suite di crittografia TLS obsoleta o non supportata per la connessione a StorageGRID.

Esempio di messaggio di errore

```
[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
```

Risoluzione: assicurarsi che JDK 11.x o versione successiva sia installato e impostare la libreria Java predefinita. Fare riferimento a. Installare il pacchetto Java per ulteriori informazioni.

Scenario 2:

Impossibile connettersi a StorageGRID con messaggio di errore `Unable to find valid certification path to requested target`.

Motivo: il certificato del server endpoint StorageGRID S3 non è attendibile dal programma Java.

Esempio di messaggio di errore:

```
[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
```

Risoluzione: NetApp consiglia di utilizzare un certificato server emesso da un'autorità pubblica nota per la firma del certificato per garantire che l'autenticazione sia sicura. In alternativa, aggiungere un certificato CA o server personalizzato all'archivio di trust Java.

Completare i seguenti passaggi per aggiungere un certificato CA o server personalizzato StorageGRID all'archivio di trust Java.

1. Eseguire il backup del file caacerts Java predefinito esistente.

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

2. Importare il certificato dell'endpoint StorageGRID S3 nell'archivio di trust 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>
```

Suggerimenti per la risoluzione dei problemi

1. Aumentare il livello di log di hadoop per ESEGUIRE IL DEBUG.

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

2. Eseguire il comando e indirizzare i messaggi di log a error.log.

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

Di Angela Cheng

Utilizzare S3cmd per testare e dimostrare l'accesso S3 su StorageGRID

Di Aron Klein

S3cmd è un tool e client a riga di comando gratuito per le operazioni S3. È possibile utilizzare s3cmd per testare e dimostrare l'accesso s3 su StorageGRID.

Installare e configurare S3cmd

Per installare S3cmd su una workstation o su un server, scaricarlo da "Client S3 della riga di comando". S3cmd è preinstallato su ciascun nodo StorageGRID come strumento per facilitare la risoluzione dei problemi.

Fasi iniziali della configurazione

1. s3cmd --configure
2. Fornire solo access_key e secret_key, per il resto mantenere le impostazioni predefinite.
3. Verificare l'accesso con le credenziali fornite? [Y/n]: n (ignora il test perché non riesce)
4. Salvare le impostazioni? [s/N] e
   a. Configurazione salvata in '/root/.s3cfg'
5. In .s3cfg svuotare i campi host_base e host_bucket dopo il segno "=" :
   a. host_base =
   b. bucket_host =

> Se si specifica host_base e host_bucket nel passaggio 4, non è necessario specificare un endpoint con --host nella CLI. Esempio:

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

Esempi di comandi di base

- **Creare un bucket:**

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

- **Elenca tutti i bucket:**

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

- **Elenca tutti i bucket e il loro contenuto:**

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

- **Elenca oggetti in un bucket specifico:**

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

- **Eliminare un bucket:**

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

- **Mettere un oggetto:**

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

- **Ottenere un oggetto:**

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

- **Elimina un oggetto:**

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

Database in modalità Vertica Eon che utilizza NetApp StorageGRID come storage comune

Di Angela Cheng

Questa guida descrive la procedura per creare un database Vertica Eon Mode con storage comune su NetApp StorageGRID.

Introduzione

Vertica è un software per la gestione di database analitici. Si tratta di una piattaforma di storage colonnare progettata per gestire grandi volumi di dati, che consente performance di query molto veloci in uno scenario tradizionalmente intensivo. Un database Vertica viene eseguito in una delle due modalità: EON o Enterprise. Puoi implementare entrambe le modalità on-premise o nel cloud.

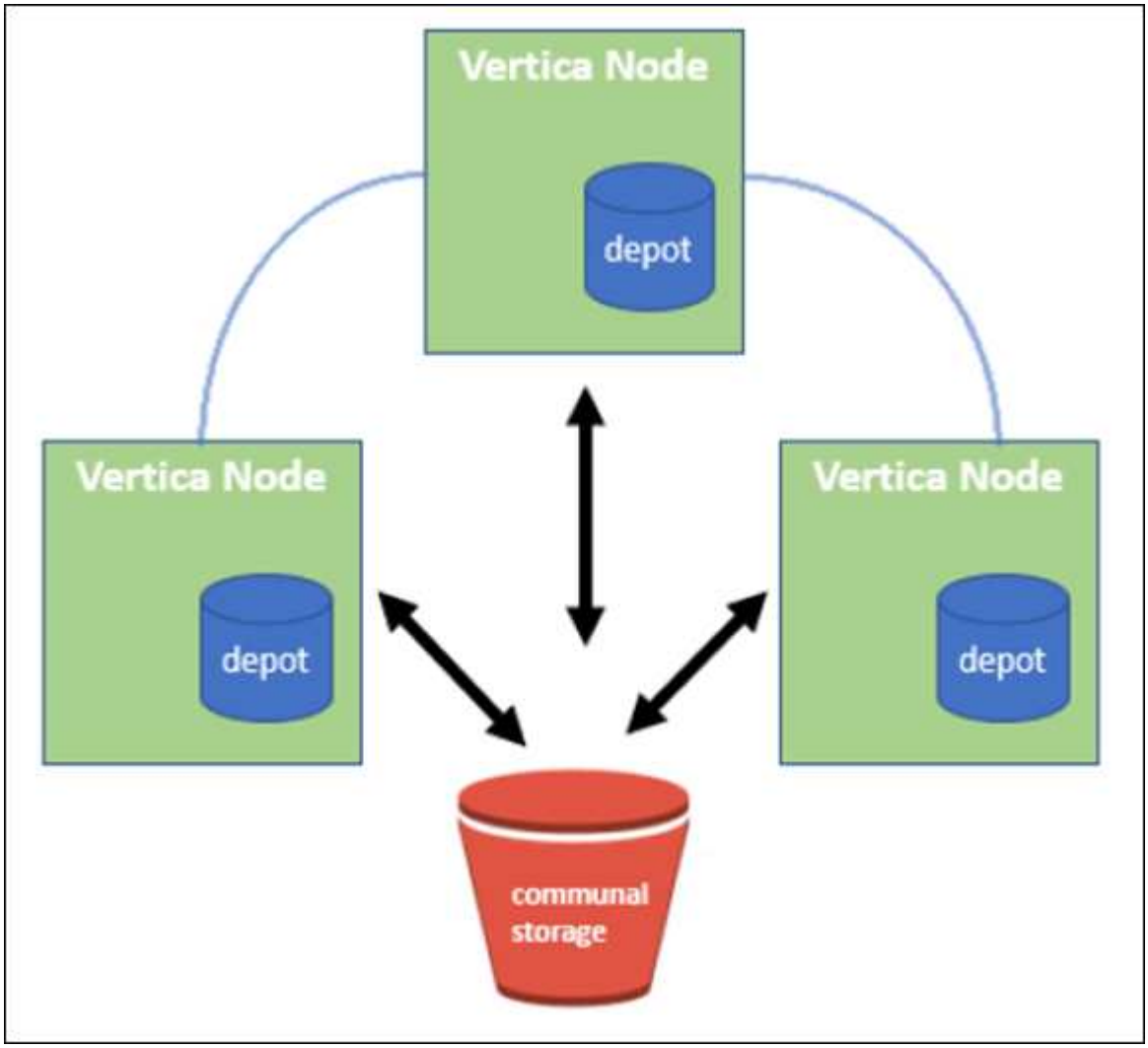
Le modalità EON ed Enterprise si differenziano principalmente per la posizione in cui memorizzano i dati:

- I database EON Mode utilizzano lo storage comune per i propri dati. Questo è consigliato da Vertica.
- I database in modalità Enterprise memorizzano i dati localmente nel file system dei nodi che compongono il database.

Architettura EON Mode

La modalità EON separa le risorse di calcolo dal livello di storage comune del database, consentendo la scalabilità separata di calcolo e storage. Vertica in Eon Mode è ottimizzato per gestire carichi di lavoro variabili e isolarli l'uno dall'altro utilizzando risorse di calcolo e storage separate.

La modalità EON memorizza i dati in un archivio di oggetti condiviso chiamato storage comune, un bucket S3, ospitato on-premise o su Amazon S3.



Storage in comune

Invece di memorizzare i dati in locale, Eon Mode utilizza una singola posizione di storage comune per tutti i dati e il catalogo (metadati). Lo storage comune è la posizione di storage centralizzata del database, condivisa tra i nodi del database.

Lo storage comune ha le seguenti proprietà:

- Lo storage comune nel cloud o in sede è più resiliente e meno suscettibile alla perdita di dati dovuta a

guasti dello storage rispetto allo storage su disco su singoli computer.

- Tutti i dati possono essere letti da qualsiasi nodo utilizzando lo stesso percorso.
- La capacità non è limitata dallo spazio su disco sui nodi.
- Poiché i dati vengono memorizzati in maniera comune, è possibile scalare in modo elastico il cluster per soddisfare le esigenze in continua evoluzione. Se i dati fossero memorizzati localmente sui nodi, l'aggiunta o la rimozione di nodi richiederebbe lo spostamento di quantità significative di dati tra i nodi per spostarli dai nodi che vengono rimossi o nei nodi appena creati.

Il deposito

Uno svantaggio dello storage comune è la sua velocità. L'accesso ai dati da una posizione cloud condivisa è più lento rispetto alla lettura dal disco locale. Inoltre, la connessione allo storage comune può diventare un collo di bottiglia se molti nodi stanno leggendo i dati da esso contemporaneamente. Per migliorare la velocità di accesso ai dati, i nodi in un database Eon Mode mantengono una cache locale su disco di dati chiamata depot. Durante l'esecuzione di una query, i nodi verificano innanzitutto se i dati necessari si trovano nel deposito. In tal caso, la query viene completata utilizzando la copia locale dei dati. Se i dati non si trovano nel deposito, il nodo recupera i dati dallo storage comune e ne salva una copia nel deposito.

Consigli di NetApp StorageGRID

Vertica archivia i dati del database nello storage a oggetti sotto forma di migliaia (o milioni) di oggetti compressi (le dimensioni osservate sono da 200 a 500 MB per oggetto. Quando un utente esegue query di database, Vertica recupera l'intervallo di dati selezionato da questi oggetti compressi in parallelo utilizzando la chiamata get dell'intervallo di byte. Ogni byte-range GET è di circa 8 KB.

Durante il test delle query utente di depot del database da 10 TB, sono state inviate alla griglia da 4,000 a 10,000 richieste GET (byte-range GET) al secondo. Quando si esegue questo test utilizzando appliance SG6060, sebbene la percentuale di utilizzo della CPU per nodo appliance sia bassa (dal 20% al 30% circa), 2/3 CPU sono in attesa di i/O. Una percentuale molto piccola (da 0% a 0.5%) di attesa i/o viene osservata su SGF6024.

A causa dell'elevata richiesta di IOPS di piccole dimensioni con requisiti di latenza molto bassi (la media dovrebbe essere inferiore a 0.01 secondi), NetApp consiglia di utilizzare SFG6024 per i servizi di storage a oggetti. Se SG6060 è necessario per database di dimensioni molto grandi, il cliente deve collaborare con il team account Vertica per il dimensionamento dei depositi al fine di supportare il set di dati attivamente interrogato.

Per il nodo di amministrazione e il nodo di gateway API, il cliente può utilizzare SG100 o SG1000. La scelta dipende dal numero di richieste di query degli utenti in parallelo e dalle dimensioni del database. Se il cliente preferisce utilizzare un bilanciatore di carico di terze parti, NetApp consiglia un bilanciatore di carico dedicato per carichi di lavoro con domanda di performance elevate. Per il dimensionamento di StorageGRID, consulta l'account team di NetApp.

Altri consigli per la configurazione di StorageGRID includono:

- **Topologia della griglia**. Non mischiare SGF6024 con altri modelli di appliance di storage sullo stesso sito di grid. Se si preferisce utilizzare SG6060 per la protezione dell'archivio a lungo termine, mantenere SGF6024 con un sistema di bilanciamento del carico di rete dedicato nel proprio sito di rete (sito fisico o logico) per un database attivo al fine di migliorare le performance. La combinazione di diversi modelli di appliance sullo stesso sito riduce le performance complessive del sito.
- **Protezione dei dati**. Utilizzare copie replicate per la protezione. Non utilizzare la codifica di cancellazione per un database attivo. Il cliente può utilizzare l'erasure coding per una protezione a lungo termine dei database inattivi.

- **Non attivare la compressione della griglia**. Vertica comprime gli oggetti prima di memorizzarli nello storage a oggetti. L'abilitazione della compressione grid non consente di risparmiare ulteriormente l'utilizzo dello storage e riduce significativamente le performance di BYTE-range GET.
- **HTTP rispetto alla connessione endpoint HTTPS S3**. Durante il test di benchmark, abbiamo osservato un miglioramento delle performance pari a circa il 5% quando si utilizza una connessione HTTP S3 dal cluster Vertica all'endpoint del bilanciamento del carico di StorageGRID. Questa scelta deve essere basata sui requisiti di sicurezza del cliente.

I consigli per una configurazione Vertica includono:

- **Le impostazioni predefinite del depot del database Vertica sono attivate (valore = 1) per le operazioni di lettura e scrittura**. NetApp consiglia vivamente di mantenere abilitate queste impostazioni di deposito per migliorare le performance.
- **Disattiva le limitazioni dello streaming**. Per informazioni dettagliate sulla configurazione, consultare la sezione Disattivazione delle limitazioni dello streaming.

Installazione della modalità Eon on on on-premise con storage comune su StorageGRID

Nelle sezioni seguenti viene descritta la procedura per installare la modalità Eon on on on-premise con lo storage comune su StorageGRID. La procedura per configurare lo storage a oggetti compatibile con S3 (Simple Storage Service) on-premise è simile alla procedura della guida Vertica, "Installare un database in modalità Eon on on on-premise".

Per il test funzionale è stata utilizzata la seguente configurazione:

- StorageGRID 11.4.0.4
- Verticale 10.1.0
- Tre macchine virtuali (VM) con sistema operativo CentOS 7.x per i nodi Vertica per formare un cluster. Questa configurazione è solo per il test funzionale, non per il cluster di database di produzione Vertica.

Questi tre nodi sono configurati con una chiave Secure Shell (SSH) per consentire SSH senza una password tra i nodi all'interno del cluster.

Informazioni richieste da NetApp StorageGRID

Per installare la modalità Eon on on on-premise con lo storage comune su StorageGRID, è necessario disporre delle seguenti informazioni sui prerequisiti.

- Indirizzo IP o FQDN (Fully Qualified Domain Name) e numero di porta dell'endpoint StorageGRID S3. Se si utilizza HTTPS, utilizzare un'autorità di certificazione personalizzata (CA) o un certificato SSL autofirmato implementato sull'endpoint StorageGRID S3.
- Nome bucket. Deve essere pre-esistente e vuoto.
- Access key ID (ID chiave di accesso) e secret access key (chiave di accesso segreta) con accesso in lettura e scrittura al bucket.

Creazione di un file di autorizzazione per accedere all'endpoint S3

I seguenti prerequisiti si applicano quando si crea un file di autorizzazione per accedere all'endpoint S3:

- Vertica è installato.

- Un cluster viene configurato, configurato e pronto per la creazione del database.

Per creare un file di autorizzazione per accedere all'endpoint S3, attenersi alla seguente procedura:

1. Accedere al nodo Vertica in cui si desidera eseguire `admintools` Per creare il database Eon Mode.

   L'utente predefinito è `dbadmin`, Creato durante l'installazione del cluster Vertica.

2. Utilizzare un editor di testo per creare un file in `/home/dbadmin` directory. Il nome del file può essere qualsiasi cosa si desideri, ad esempio `sg_auth.conf`.
3. Se l'endpoint S3 utilizza una porta HTTP standard 80 o una porta HTTPS 443, ignorare il numero della porta. Per utilizzare HTTPS, impostare i seguenti valori:

   ◦ `awsenablehttps = 1`, altrimenti impostare il valore su `0`.

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

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

   Per utilizzare una CA personalizzata o un certificato SSL autofirmato per la connessione HTTPS dell'endpoint StorageGRID S3, specificare il percorso completo del file e il nome del file del certificato. Questo file deve trovarsi nella stessa posizione su ciascun nodo Vertica e disporre dell'autorizzazione di lettura per tutti gli utenti. Saltare questo passaggio se il certificato SSL StorageGRID S3 Endpoint è firmato da una CA pubblicamente conosciuta.

   - `awscafile = <filepath/filename>`

   Ad esempio, vedere il seguente file di esempio:

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

+

> In un ambiente di produzione, il cliente deve implementare un certificato server firmato da una CA pubblicamente conosciuta su un endpoint di bilanciamento del carico StorageGRID S3.

Scelta di un percorso di deposito su tutti i nodi Vertica

Scegliere o creare una directory su ciascun nodo per il percorso di storage del deposito. La directory fornita per il parametro del percorso di storage del deposito deve essere la seguente:

- Lo stesso percorso su tutti i nodi del cluster (ad esempio, `/home/dbadmin/depot`)
- Essere leggibile e scrivibile dall'utente dbadmin
- Storage sufficiente

  Per impostazione predefinita, Vertica utilizza il 60% dello spazio del file system contenente la directory per

lo storage del depot. È possibile limitare le dimensioni del deposito utilizzando `--depot-size` argomento in `create_db` comando. Vedere "Dimensionamento del cluster Vertica per un database in modalità Eon" articolo per le linee guida generali sul dimensionamento di Vertica o consulta il tuo account manager Vertica.

Il `admintools create_db` lo strumento tenta di creare il percorso del deposito se non ne esiste uno.

Creazione del database Eon on on on-premise

Per creare il database Eon on on on-premise, attenersi alla seguente procedura:

1. Per creare il database, utilizzare `admintools create_db` tool.

   L'elenco seguente fornisce una breve spiegazione degli argomenti utilizzati in questo esempio. Consultare il documento Vertica per una spiegazione dettagliata di tutti gli argomenti richiesti e facoltativi.

   - -x <path/filename of authorization file created in "Creazione di un file di autorizzazione per accedere all'endpoint S3" >.

     I dettagli dell'autorizzazione vengono memorizzati all'interno del database dopo la creazione. È possibile rimuovere questo file per evitare di esporre la chiave segreta S3.

   - --communal-storage-location <s3://storagegrid bucketname>
   - -S <comma-separated list of Vertica nodes to be used for this database>
   - -d <name of database to be created>
   - -p <password to be set for this new database>. Ad esempio, vedere il seguente comando di esempio:

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

     La creazione di un nuovo database richiede diversi minuti a seconda del numero di nodi del database. Quando si crea un database per la prima volta, viene richiesto di accettare il Contratto di licenza.

Ad esempio, vedere il seguente file di autorizzazione di esempio 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.
```

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Dimensione oggetto (byte)	Percorso completo della chiave bucket/oggetto
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

Disattivazione delle limitazioni dello streaming

Questa procedura si basa sulla guida Vertica per altri storage a oggetti on-premise e deve essere applicabile a

StorageGRID.

1. Dopo aver creato il database, disattivare `AWSStreamingConnectionPercentage` parametro di configurazione impostandolo su `0`. Questa impostazione non è necessaria per un'installazione on Mode on-premise con storage comune. Questo parametro di configurazione controlla il numero di connessioni all'archivio di oggetti utilizzate da Vertica per le letture in streaming. In un ambiente cloud, questa impostazione consente di evitare che i dati in streaming dall'archivio di oggetti utilizzino tutti gli handle di file disponibili. In questo modo, alcuni handle di file sono disponibili per altre operazioni di archiviazione di oggetti. A causa della bassa latenza degli archivi di oggetti on-premise, questa opzione non è necessaria.
2. Utilizzare un `vsql` per aggiornare il valore del parametro. La password è la password del database impostata in "creazione del database Eon on on on-premise". Ad esempio, vedere il seguente esempio di output:

```
[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
```

Verifica delle impostazioni del deposito in corso

Le impostazioni predefinite del depot del database Vertica sono attivate (valore = 1) per le operazioni di lettura e scrittura. NetApp consiglia vivamente di mantenere abilitate queste impostazioni di deposito per migliorare le performance.

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

Caricamento dei dati di esempio (opzionale)

Se questo database è destinato al test e verrà rimosso, è possibile caricare i dati campione in questo database per il test. Vertica viene fornito con un set di dati di esempio, VMart, disponibile in `/opt/vertica/examples/VMart_Schema/` Su ogni nodo Vertica. Sono disponibili ulteriori informazioni su questo set di dati di esempio "qui".

Per caricare i dati di esempio, procedere come segue:

1. Accedere come dbadmin a uno dei nodi Vertica: cd /opt/vertica/exemes/VMart_Schema/
2. Caricare i dati di esempio nel database e inserire la password del database quando richiesto nelle fasi 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. Esistono più query SQL predefinite, alcune delle quali possono essere eseguite per confermare che i dati di test sono stati caricati correttamente nel database. Ad esempio: `vsql < vmart_queries1.sql`

Dove trovare ulteriori informazioni

Per ulteriori informazioni sulle informazioni descritte in questo documento, consultare i seguenti documenti e/o siti Web:

- "Documentazione del prodotto NetApp StorageGRID 11,7"
- "Scheda tecnica di StorageGRID"
- "Documentazione del prodotto Vertica 10.1"

Cronologia delle versioni

Versione	Data	Cronologia delle versioni del documento
Versione 1.0	Settembre 2021	Release iniziale.

Di Angela Cheng

Analisi dei log StorageGRID con stack ELK

Di Angela Cheng

Con la funzione di inoltro syslog di StorageGRID, è possibile configurare un server syslog esterno per raccogliere e analizzare i messaggi di registro StorageGRID. ELK (Elasticsearch, Logstash, Kibana) è diventata una delle soluzioni di analisi dei log più diffuse. Osservare la "Analisi del log StorageGRID con video ELK" per visualizzare una configurazione ELK di esempio e come può essere utilizzata per identificare e risolvere i problemi relativi alle richieste S3 non riuscite. StorageGRID 11,9 supporta l'esportazione del log di accesso agli endpoint del bilanciamento del carico nel server syslog esterno. Guarda questo "Video di YouTube" articolo per saperne di più su questa nuova funzionalità. Questo articolo fornisce file di esempio di configurazione di Logstash, query Kibana, grafici e dashboard per fornire un rapido avvio per la gestione dei log e l'analisi di StorageGRID.

Requisiti

- StorageGRID 11.6.0.2 o superiore
- ELK (Elasticsearch, Logstash e Kibana) 7.1x o superiore installato e in funzione

File di esempio

- "Scarica il pacchetto di file di esempio Logstash 7.x." + **checksum md5** 148c23d0021d9a4bb4a6c0287464deab + **checksum sha256** f51ec9e2e3f842d5a7861566b167a561beb4373038b4e7bb3c8be3d522adf2d6
- "Scarica il pacchetto di file di esempio Logstash 8.x." + **checksum md5** e11bae3a662f87c310ef363d0fe06835 + **checksum sha256** 5c670755742cfdfd5aa723a596ba087e0153a65bcaef3934afdb682f61cd278d

- "Scaricare il pacchetto file di esempio Logstash 8.x per StorageGRID 11,9" + **md5 checksum** 41272857c4a54600f95995f6ed74800d + **sha256 checksum** 67048e8661052719990851e1ad960d4902fe537a6e135e8600177188da6779c9

Assunzione

I lettori conoscono la terminologia e le operazioni di StorageGRID ed ELK.

Istruzioni

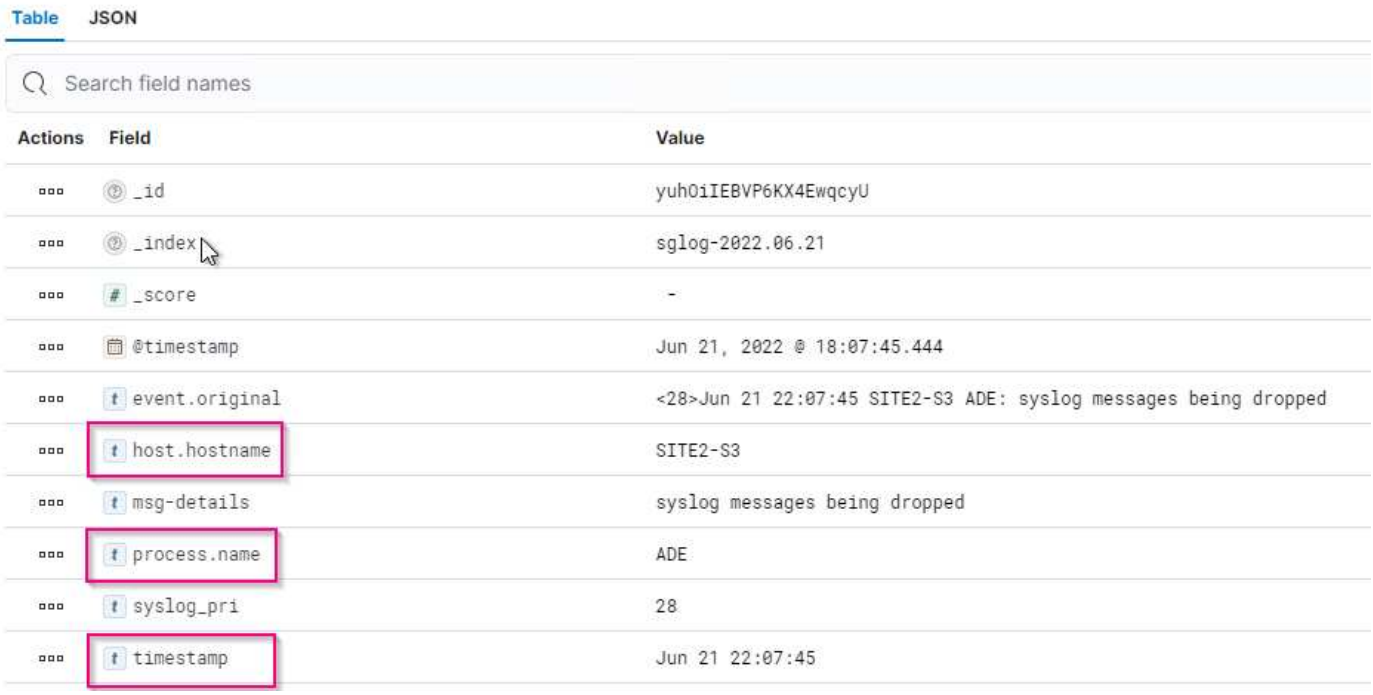
Due versioni di esempio sono fornite a causa delle differenze nei nomi definiti dai modelli grok. Ad esempio, il modello SYSLOGBASE grok nel file di configurazione di Logstash definisce i nomi dei campi in modo diverso a seconda della versione di Logstash installata.

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

Esempio di Logstash 7.17

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

Esempio di Logstash 8.23



Fasi

1. Decomprimere l'esempio fornito in base alla versione ELK installata. La cartella di esempio include due esempi di configurazione di Logstash: + **sglog-2-file.conf:** questo file di configurazione genera messaggi di log StorageGRID in un file su Logstash senza trasformazione dei dati. È possibile utilizzare questa opzione per confermare che Logstash riceve messaggi StorageGRID o per comprendere meglio i modelli di log di StorageGRID. + **sglog-2-es.conf:** questo file di configurazione trasforma i messaggi di log di StorageGRID utilizzando vari modelli e filtri. Include istruzioni drop di esempio, che consentono di eliminare i messaggi in base a modelli o filtri. L'output viene inviato a Elasticsearch per l'indicizzazione. + personalizzare il file di configurazione selezionato in base alle istruzioni contenute nel file.
2. Verificare il file di configurazione personalizzato:

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

   Se l'ultima riga restituita è simile alla riga seguente, il file di configurazione non presenta errori di sintassi:

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

3. Copiare il file di configurazione personalizzato nella configurazione del server Logstash: /Etc/logstash/conf.d + se non si è abilitato config.reload.automatic in /etc/logstash/logstash.yml, riavviare il servizio Logstash. In caso contrario, attendere lo scadere dell'intervallo di ricarica della configurazione.

```
 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
```

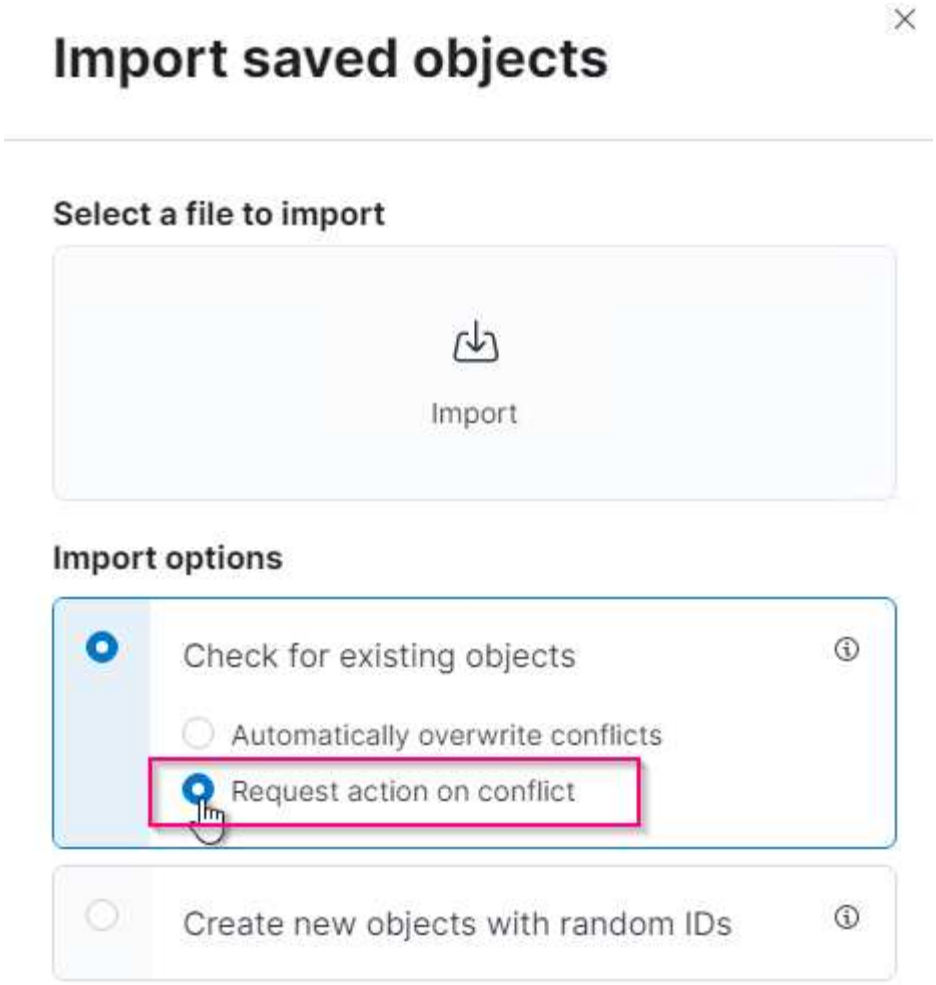
4. Controllare /var/log/logstash/logstash-plain.log e verificare che non ci siano errori durante l'avvio di Logstash con il nuovo file di configurazione.
5. Verificare che la porta TCP sia stata avviata e in attesa. + in questo esempio, viene utilizzata la porta TCP 5000.

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

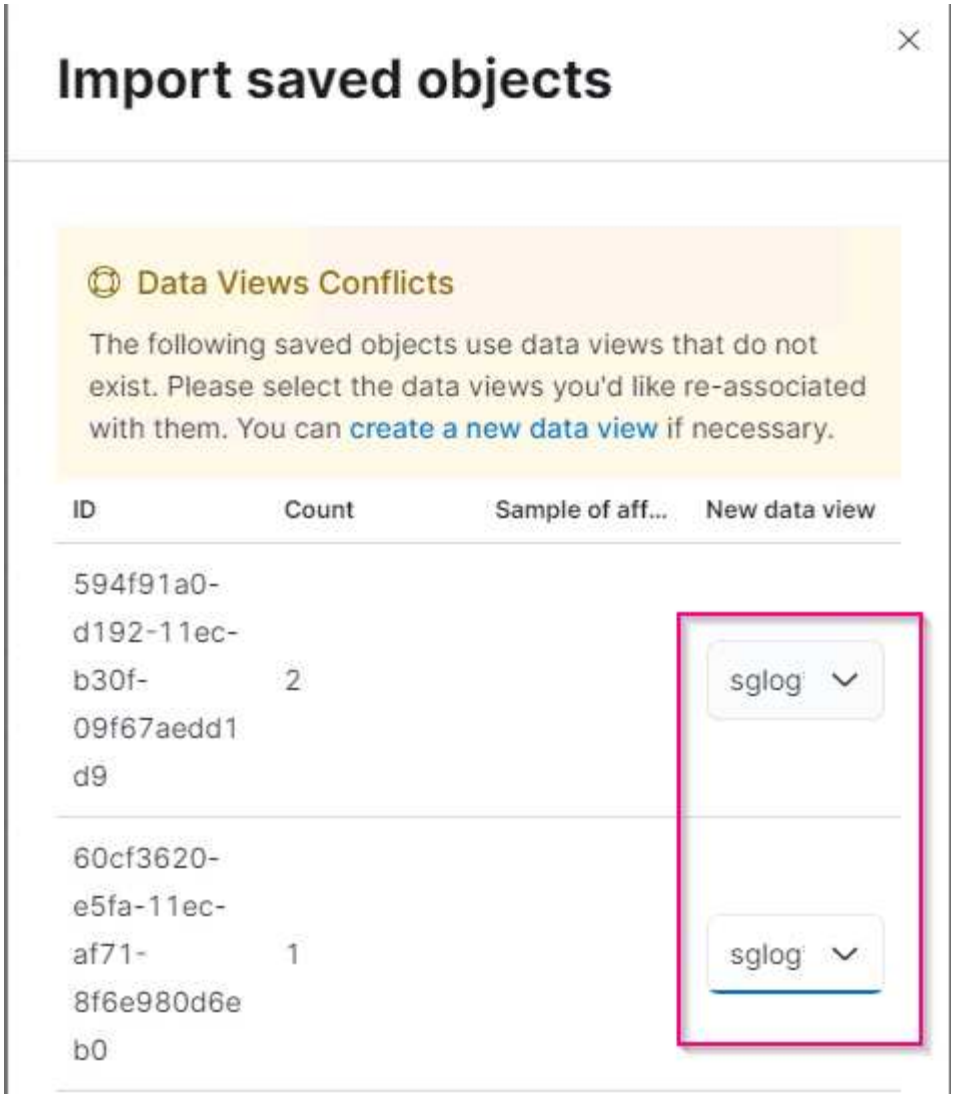
6. Dalla GUI di StorageGRID Manager, configurare il server syslog esterno per l'invio dei messaggi di log a Logstash. Per ulteriori informazioni, fare riferimento alla "video dimostrativo" .
7. È necessario configurare o disattivare il firewall sul server Logstash per consentire la connessione dei nodi StorageGRID alla porta TCP definita.
8. Dalla GUI di Kibana, selezionare Management (Gestione) → Dev Tools (Strumenti di sviluppo). Nella pagina Console, eseguire questo comando GET per confermare la creazione di nuovi indici in Elasticsearch.

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

9. Dalla GUI di Kibana, creare un modello di indice (ELK 7.x) o una vista dati (ELK 8.x).
10. Dalla GUI di Kibana, inserire "oggetti memorizzati" nella casella di ricerca situata in alto al centro. + nella pagina Saved Objects (oggetti salvati), selezionare Import (Importa). In Opzioni di importazione, selezionare "Richiedi azione in caso di conflitto"



Importa elk <version>-query-chart-sample.ndjson. + quando viene richiesto di risolvere il conflitto, selezionare il modello di indice o la vista dati creata al punto 8.



Vengono importati i seguenti oggetti Kibana: + **Query** + * audit-msg-s3rq-orlm + * bycast log S3 messaggi correlati + * avviso livello di accesso o superiore + * evento di sicurezza non riuscito + * nginx-gw log di accesso endpoint (disponibile solo in elk8-sample-for-sg119.zip) + **grafico** + * S3 conteggio richieste basato su bycast.log + * Codice di stato HTTP + * tipo di controllo + tempo di risposta medio del dashboard * S3 msg del dashboard * dati di analisi del dashboard * dati del tipo di analisi del dashboard * S3.

A questo punto, è possibile eseguire l'analisi del registro StorageGRID utilizzando Kibana.

Risorse aggiuntive

- "syslog101"
- "Cos'è lo stack ELK"
- "Elenco dei modelli di grok"
- "Guida per principianti a Logstash: Grok"
- "Una guida pratica a Logstash: Approfondimento di Syslog"
- "Guida di Kibana – Esplora il documento"
- "Riferimento ai messaggi del registro di controllo di StorageGRID"

Utilizza Prometheus e Grafana per estendere la conservazione delle metriche

Di Aron Klein

Questo report tecnico fornisce istruzioni dettagliate per la configurazione NetApp StorageGRID con servizi esterni Prometheus e Grafana.

Introduzione

StorageGRID memorizza le metriche utilizzando Prometheus e fornisce visualizzazioni di queste metriche attraverso dashboard Grafana integrate. È possibile accedere in modo sicuro alle metriche Prometheus da StorageGRID configurando i certificati di accesso client e abilitando l'accesso prometheus per il client specificato. Oggi, la conservazione di questi dati metrici è limitata dalla capacità di storage del nodo di amministrazione. Per ottenere una durata maggiore e la possibilità di creare visualizzazioni personalizzate di queste metriche, implementeremo un nuovo server Prometheus e Grafana, configureremo il nostro nuovo server per scartare le metriche dall'istanza StorageGRID e costruiremo una dashboard con le metriche che sono importanti per noi. È possibile ottenere ulteriori informazioni sulle metriche Prometheus raccolte in "Documentazione StorageGRID".

Federare Prometheus

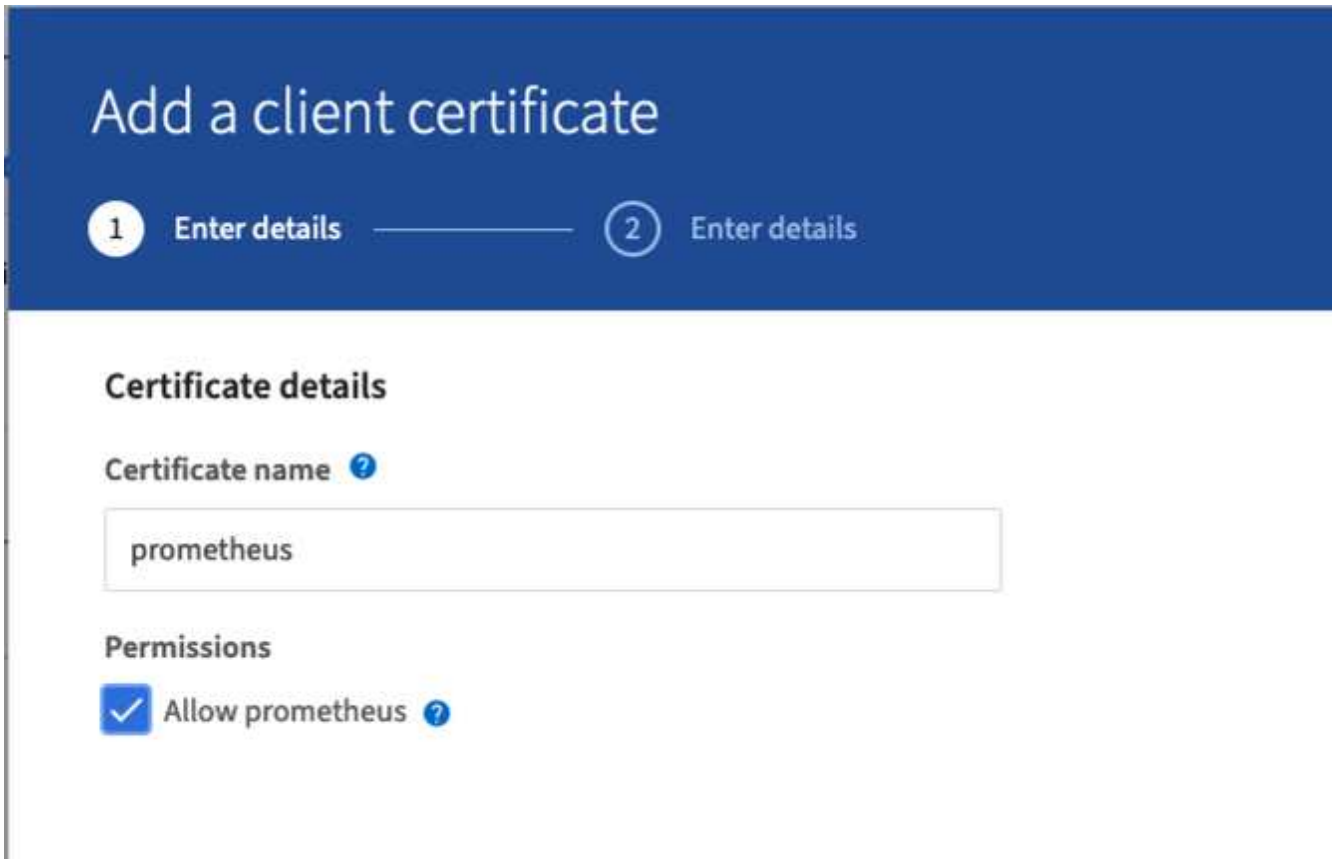
Dettagli del laboratorio

Ai fini di questo esempio, userò tutte le macchine virtuali per i nodi StorageGRID 11.6 e un server Debian 11. L'interfaccia di gestione di StorageGRID è configurata con un certificato CA pubblicamente attendibile. Questo esempio non riguarda l'installazione e la configurazione del sistema StorageGRID o dell'installazione di Debian linux. Puoi utilizzare qualsiasi versione di Linux supportata da Prometheus e Grafana. Prometheus e Grafana possono essere installati come container docker, build from source o binari pre-compilati. In questo esempio installerò entrambi i binari Prometheus e Grafana direttamente sullo stesso server Debian. Scaricare e seguire le istruzioni di installazione di base da https://prometheus.io e. https://grafana.com/grafana/ rispettivamente.

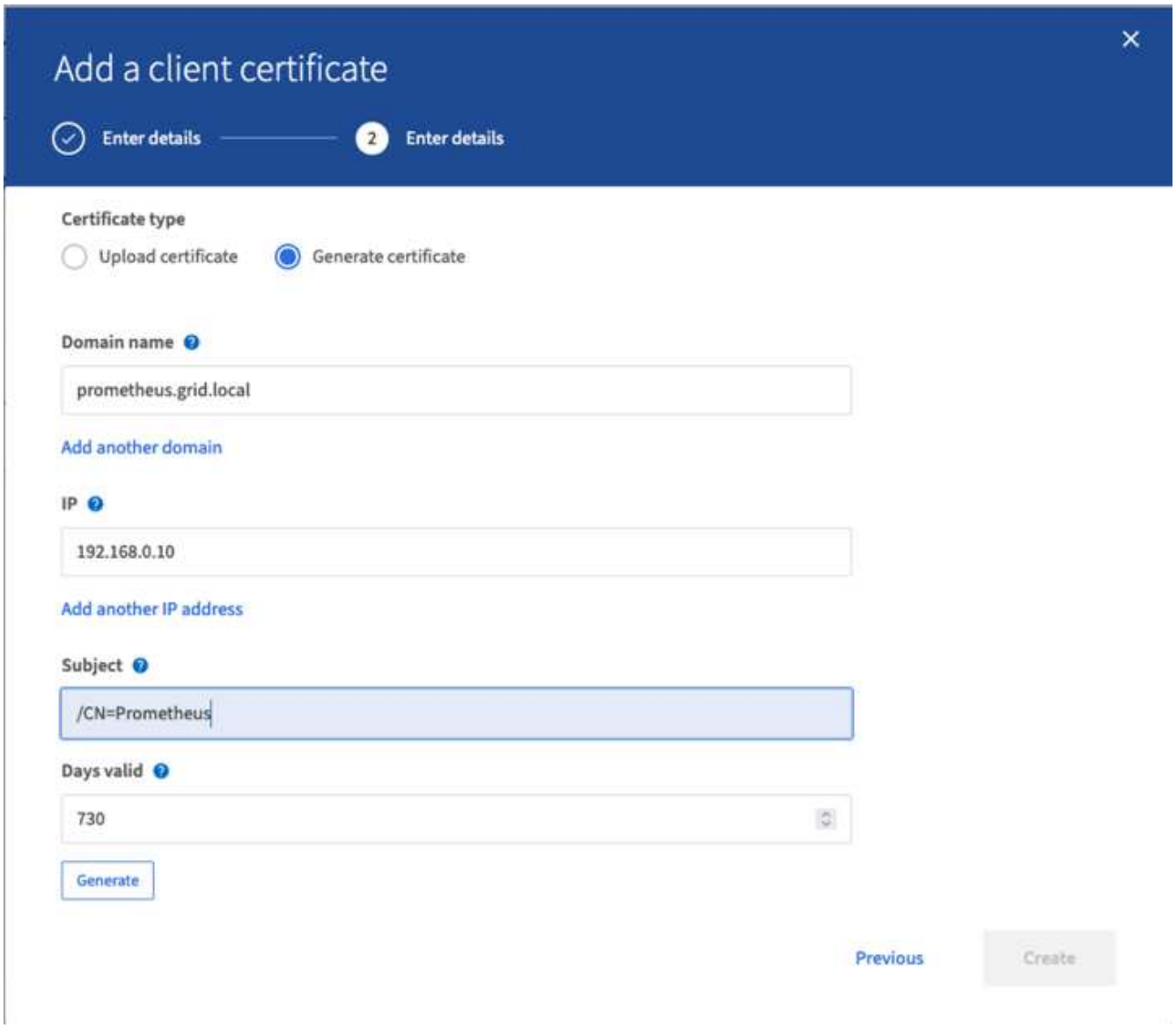
Configurare StorageGRID per l'accesso al client Prometheus

Per ottenere l'accesso alle metriche StorageGRID Stored prometheus, è necessario generare o caricare un certificato client con chiave privata e abilitare l'autorizzazione per il client. L'interfaccia di gestione StorageGRID deve disporre di un certificato SSL. Il certificato deve essere attendibile dal server prometheus da una CA attendibile o manualmente se autofirmato. Per ulteriori informazioni, visitare il "Documentazione StorageGRID".

1. Nell'interfaccia di gestione di StorageGRID, selezionare "CONFIGURATION" (CONFIGURAZIONE) in basso a sinistra e nella seconda colonna sotto "Security" (sicurezza) fare clic su Certificates (certificati).
2. Nella pagina certificati, selezionare la scheda "Client" e fare clic sul pulsante "Aggiungi".
3. Specificare un nome per il client a cui verrà concesso l'accesso e utilizzare questo certificato. Fare clic sulla casella sotto "permessi", davanti a "Consenti Prometheus" e fare clic sul pulsante continua.

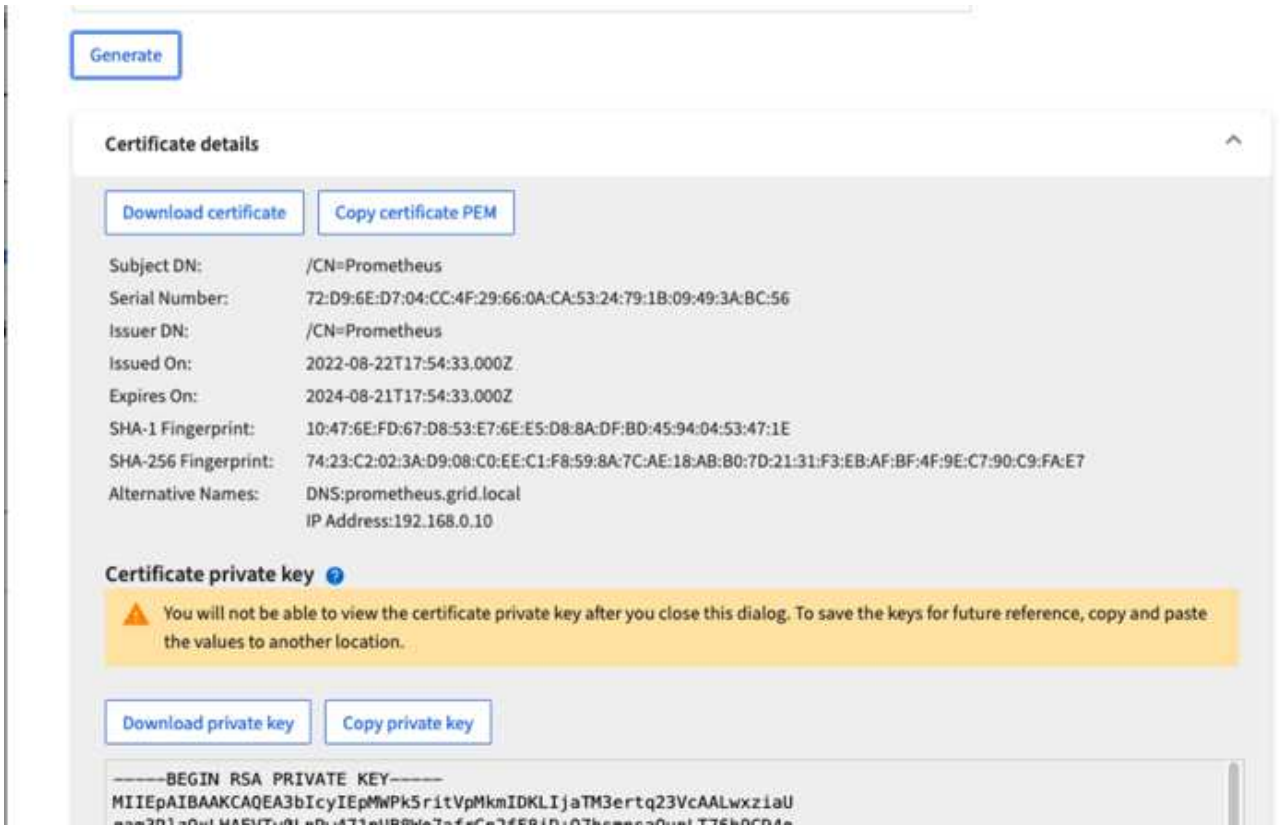


4. Se si dispone di un certificato firmato dalla CA, è possibile selezionare il pulsante di opzione "carica certificato", ma in questo caso StorageGRID genererà il certificato client selezionando il pulsante di opzione "genera certificato". Vengono visualizzati i campi obbligatori da compilare. Inserire l'FQDN per il server client, l'IP del server, l'oggetto e i giorni validi. Quindi fare clic sul pulsante "generate" (genera).



> 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. Scaricare il file pem del certificato e il file pem della chiave privata.



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

Preparare il server Linux per l'installazione di Prometheus

Prima di installare Prometheus, desidero preparare il mio ambiente con un utente Prometheus, la struttura di directory e configurare la capacità per la posizione di storage delle metriche.

1. Creare l'utente Prometheus.

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

2. Creare le directory per Prometheus, certificato client e dati di metriche.

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

3. Ho formattato il disco che sto usando per la conservazione delle metriche con un filesystem ext4.

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

4. Ho quindi montato il file system nella directory Prometheus metrics.

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

5. Ottenere l'uuid del disco utilizzato per i dati delle metriche.

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

6. Aggiungere una voce in /etc/fstab/ rendere il mount persistente durante i riavvii usando l'uuid di /dev/sdb.

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

Installare e configurare Prometheus

Ora che il server è pronto, posso iniziare l'installazione di Prometheus e configurare il servizio.

1. Estrarre il pacchetto di installazione di Prometheus

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

2. Copiare i file binari in /usr/local/bin e modificare la proprietà dell'utente prometheus creato in precedenza

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

3. Copiare le console e le librerie in /etc/prometheus

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

4. Copiare i file PEM del certificato client e della chiave privata scaricati in precedenza da StorageGRID in /etc/prometheus/certs
5. Creare il file yaml di configurazione prometheus

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

6. Inserire la seguente configurazione. Il nome del lavoro può essere qualsiasi cosa si desideri. Modificare "-

targets: [""] in FQDN del nodo admin e, se i nomi dei file dei certificati e delle chiavi private sono modificati, aggiornare la sezione tls_config in modo che corrisponda. quindi salvare il file. Se l'interfaccia di gestione della griglia utilizza un certificato autofirmato, scaricare il certificato e posizionarlo con il certificato client con un nome univoco, quindi nella sezione tls_config aggiungere ca_file: /Etc/prometheus/cert/UIcert.pem

a. In questo esempio, vengono raccolte tutte le metriche che iniziano con alertmanager, cassandra, Node e StorageGRID. Per ulteriori informazioni sulle metriche Prometheus, consultare la "Documentazione 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 l'interfaccia di gestione della griglia utilizza un certificato autofirmato, scaricare il certificato e posizionarlo con il certificato client con un nome univoco. Nella sezione tls_config aggiungere il certificato sopra le righe del certificato client e della chiave privata

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

1. Modificare la proprietà di tutti i file e le directory in /etc/prometheus e /var/lib/prometheus nell'utente prometheus

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

2. Creare un file di servizio prometheus in /etc/systemd/system

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

3. Inserire le seguenti righe, annotare il n.--storage.tsdb.retention.time=1y n. che imposta la conservazione dei dati metrici su 1 anno. In alternativa, è possibile utilizzare n.--storage.tsdb.retention.size=n. 300GiB per basare la conservazione sui limiti di storage. Questa è l'unica posizione in cui impostare la conservazione delle metriche.

```
[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. Ricaricare il servizio systemd per registrare il nuovo servizio prometheus. quindi avviare e attivare il servizio prometheus.

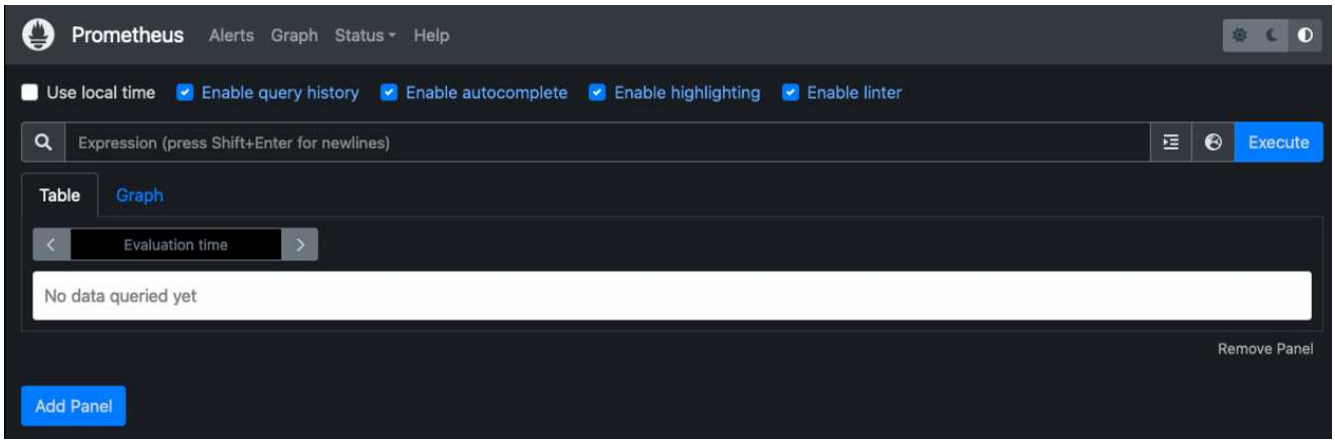
```
sudo systemctl daemon-reload
sudo systemctl start prometheus
sudo systemctl enable prometheus
```

5. Verificare che il servizio sia in funzione correttamente

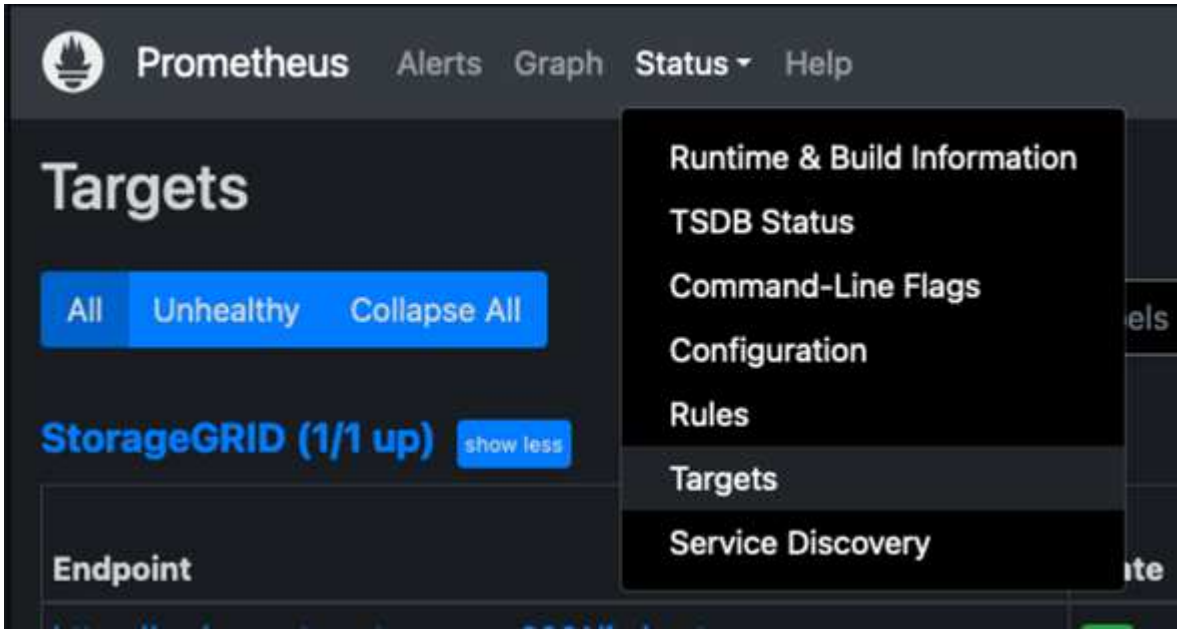
```
sudo systemctl status prometheus
```

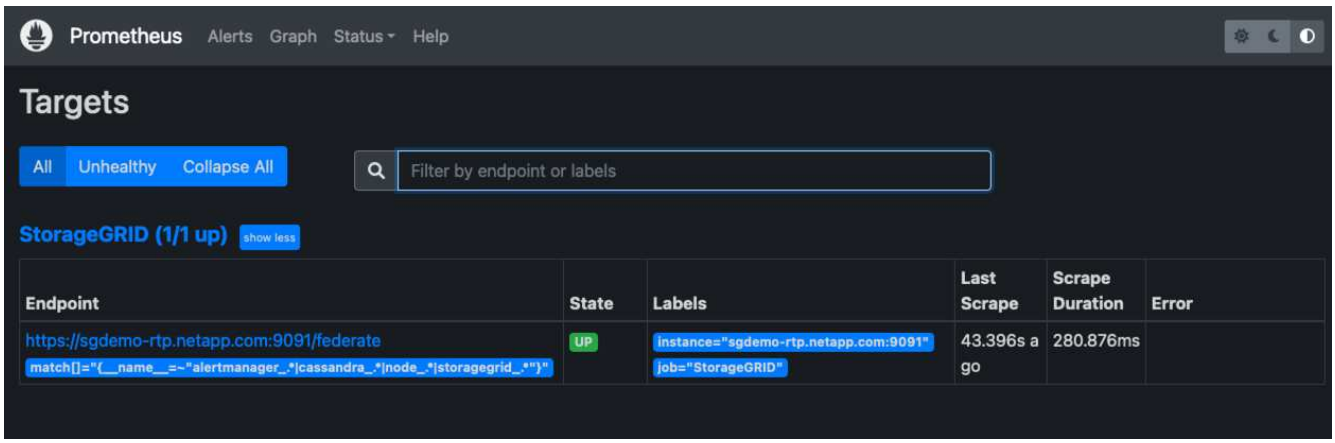
```
● prometheus.service - Prometheus Time Series Collection and Processing
Server
     Loaded: loaded (/etc/systemd/system/prometheus.service; enabled;
vendor preset: enabled)
     Active: active (running) since Mon 2022-08-22 15:14:24 EDT; 2s ago
   Main PID: 6498 (prometheus)
      Tasks: 13 (limit: 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. A questo punto, dovresti essere in grado di accedere all'interfaccia utente del tuo server prometheus http://Prometheus-server:9090 E consultare l'interfaccia utente

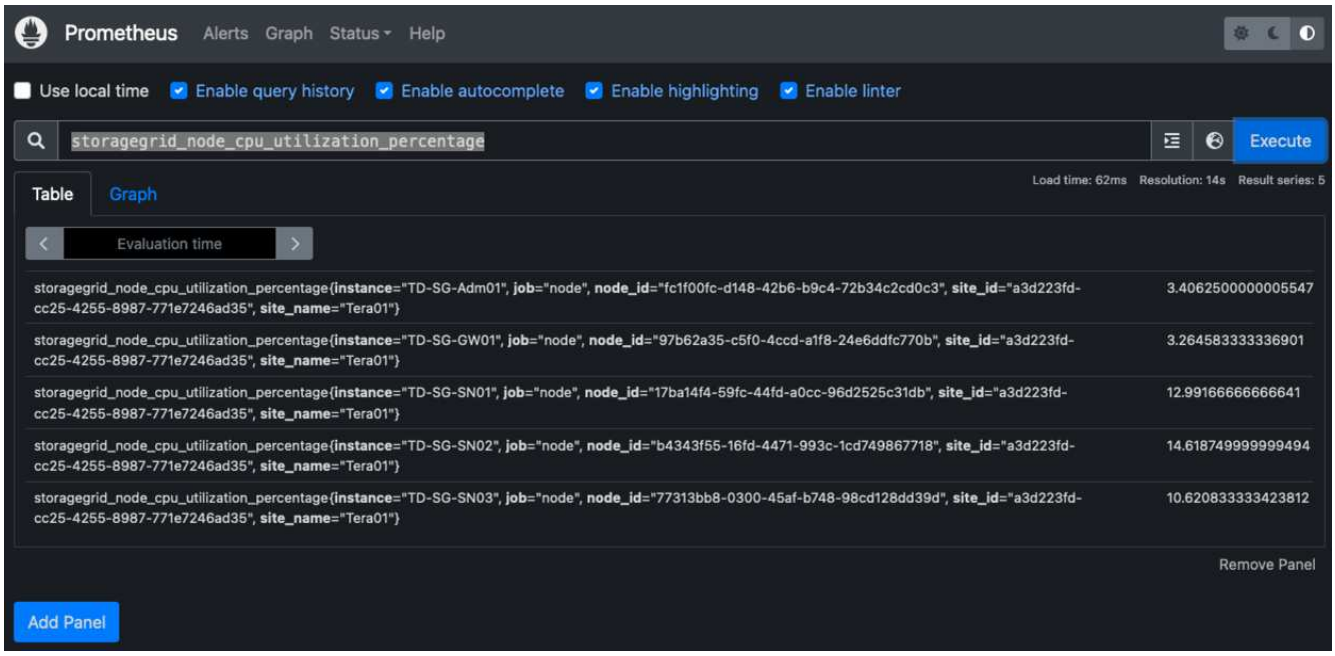


7. Sotto "Stato", è possibile visualizzare lo stato dell'endpoint StorageGRID configurato in prometheus.yml





8. Nella pagina Graph (grafico), è possibile eseguire una query di test e verificare che i dati siano stati scartati correttamente. Ad esempio, immettere "storagegrid_node_cpu_Utilization_percent" nella barra delle query e fare clic sul pulsante Execute.



Installare e configurare Grafana

Ora che prometheus è installato e funzionante, possiamo passare all'installazione di Grafana e alla configurazione di una dashboard

Installazione di Grafana

1. Installare l'ultima edizione Enterprise di 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. Aggiungi questo repository per le release stabili:

```
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. Dopo aver aggiunto il repository.

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

4. Ricaricare il servizio systemd per registrare il nuovo servizio Grafana. Quindi avviare e attivare il servizio Grafana.

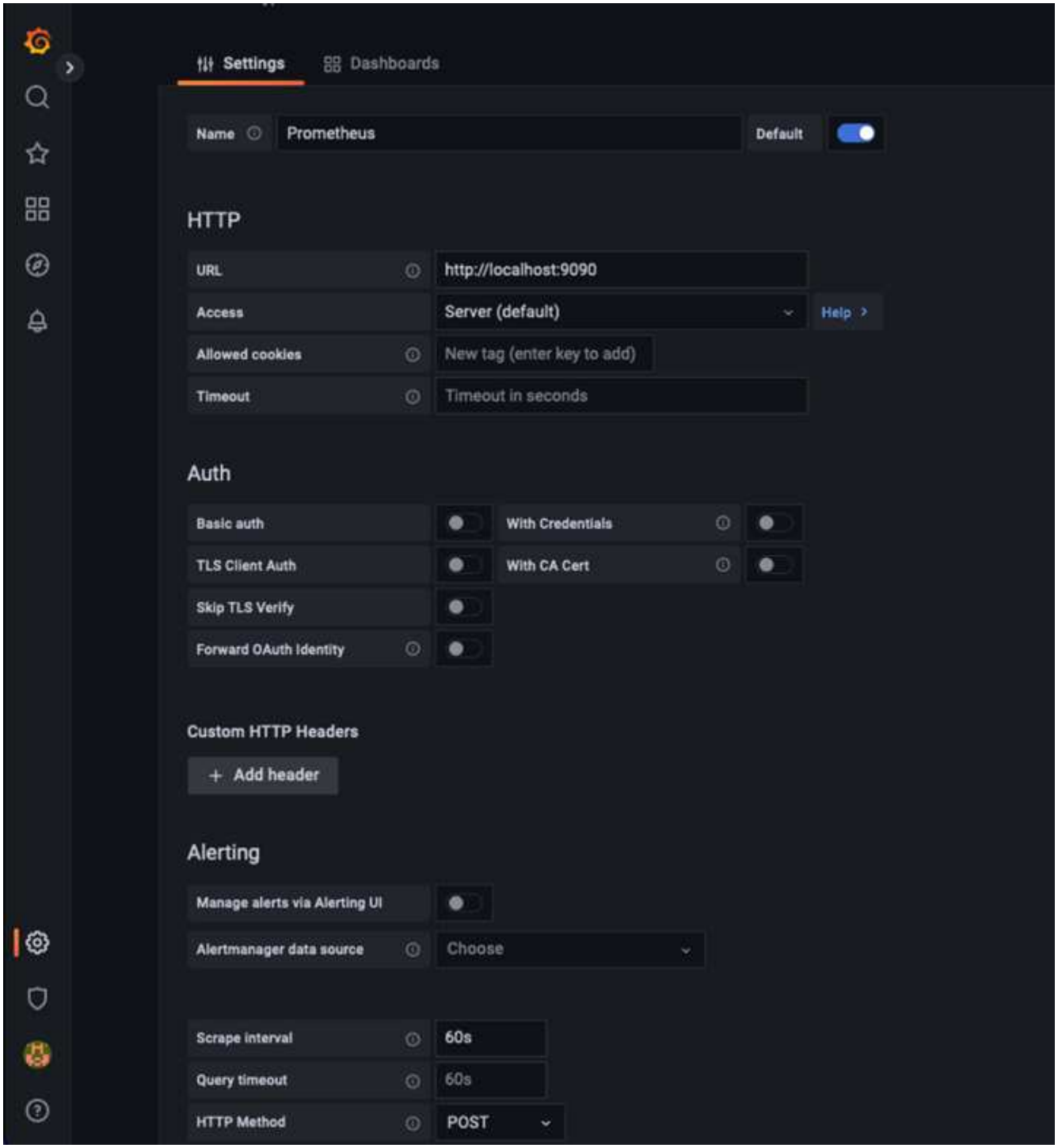
```
sudo systemctl daemon-reload
sudo systemctl start grafana-server
sudo systemctl enable grafana-server.service
```

5. Grafana è ora installato e in esecuzione. Quando si apre un browser per HTTP://Prometheus-server:3000 viene visualizzata la pagina di accesso Grafana.
6. Le credenziali di accesso predefinite sono admin/admin ed è necessario impostare una nuova password come richiesto.

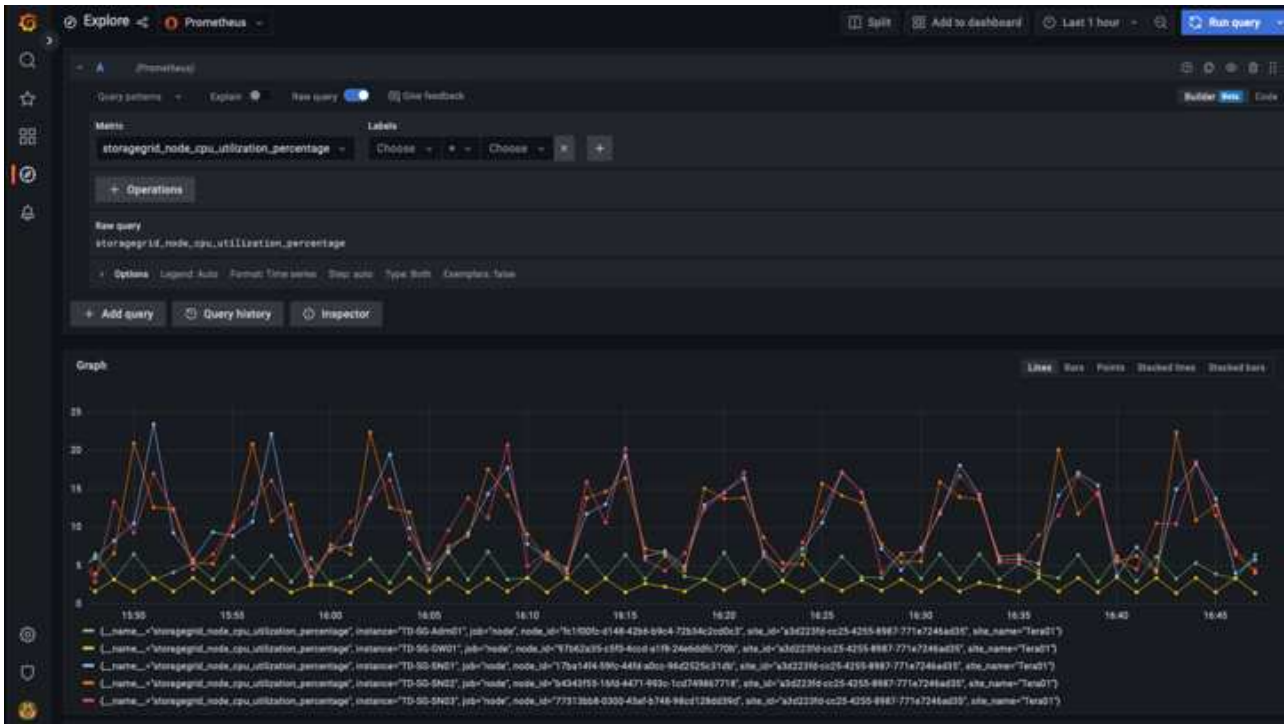
Creare una dashboard Grafana per StorageGRID

Con Grafana e Prometheus installati e in esecuzione, ora è il momento di collegare i due elementi creando un'origine dati e creando una dashboard

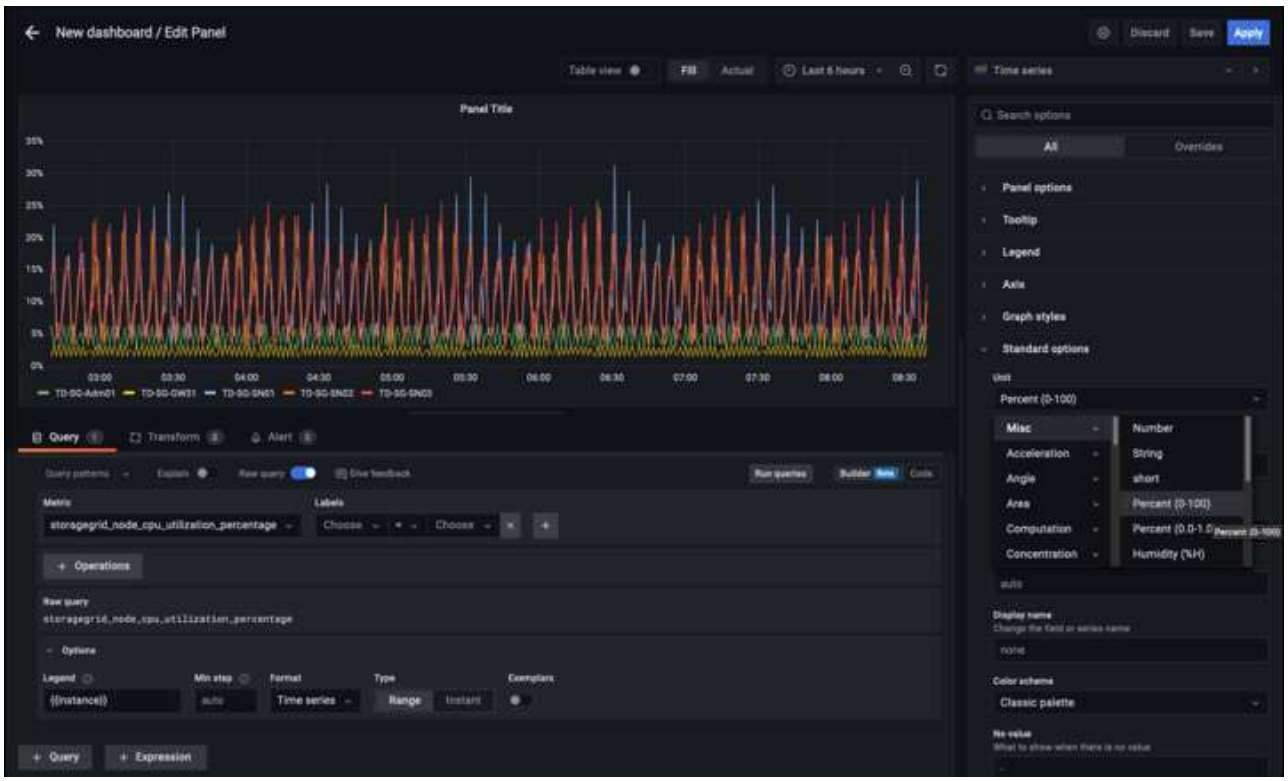
1. Nel riquadro di sinistra, espandere "Configuration" (Configurazione) e selezionare "Data Sources" (origini dati), quindi fare clic sul pulsante "Add Data Source" (Aggiungi origine dati)
2. Prometheus sarà una delle principali fonti di dati tra cui scegliere. In caso contrario, utilizzare la barra di ricerca per individuare "Prometheus"
3. Configurare l'origine Prometheus immettendo l'URL dell'istanza prometheus e l'intervallo di scrape in modo che corrisponda all'intervallo Prometheus. Ho anche disattivato la sezione degli avvisi perché non ho configurato il gestore degli avvisi su prometheus.



4. Una volta inserite le impostazioni desiderate, scorrere verso il basso e fare clic su "Save & test" (Salva e verifica).
5. Una volta completato il test di configurazione, fare clic sul pulsante Esplora.
    a. Nella finestra Esplora puoi utilizzare la stessa metrica che abbiamo testato Prometheus con "storagegrid_node_cpu_Utilization_percent" e fare clic sul pulsante "Esegui query"



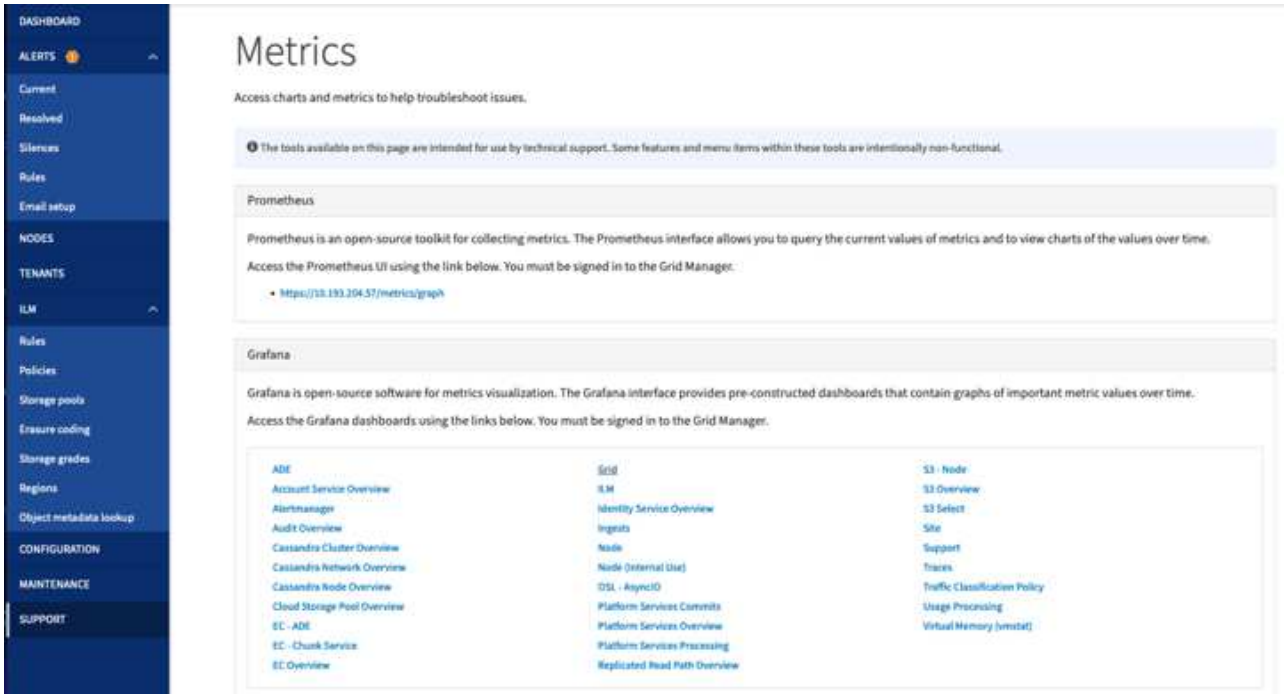
6. Ora che abbiamo configurato l'origine dati, possiamo creare una dashboard.
   a. Nel riquadro di sinistra, espandere "Dashboard" e selezionare "+ new Dashboard"
   b. Seleziona "Aggiungi un nuovo pannello"
   c. Configurare il nuovo pannello selezionando una metrica, di nuovo userò "storagegrid_node_cpu_Utilization_Percent", inserire un titolo per il pannello, espandere "Opzioni" in basso e per la modifica della legenda su custom e inserire "{{instance}}" per definire i nomi dei nodi", e nel pannello di destra in "Opzioni standard" impostare "unità" su "varie/percentuali(0-100)". Quindi fare clic su "Apply" (Applica) per salvare il pannello nella dashboard.



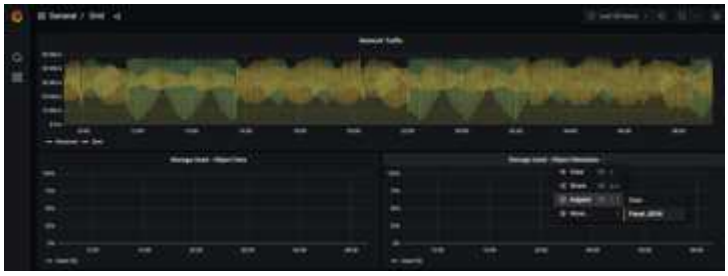
7. Potremmo continuare a costruire la nostra dashboard in questo modo per ogni metrica che vogliamo, ma fortunatamente StorageGRID dispone già di dashboard con pannelli che possiamo copiare nelle nostre dashboard personalizzate.

   a. Dal riquadro sinistro dell'interfaccia di gestione StorageGRID, selezionare "supporto", quindi fare clic su "metriche" nella parte inferiore della colonna "Strumenti".

   b. All'interno delle metriche, selezionerò il link "Grid" nella parte superiore della colonna centrale.

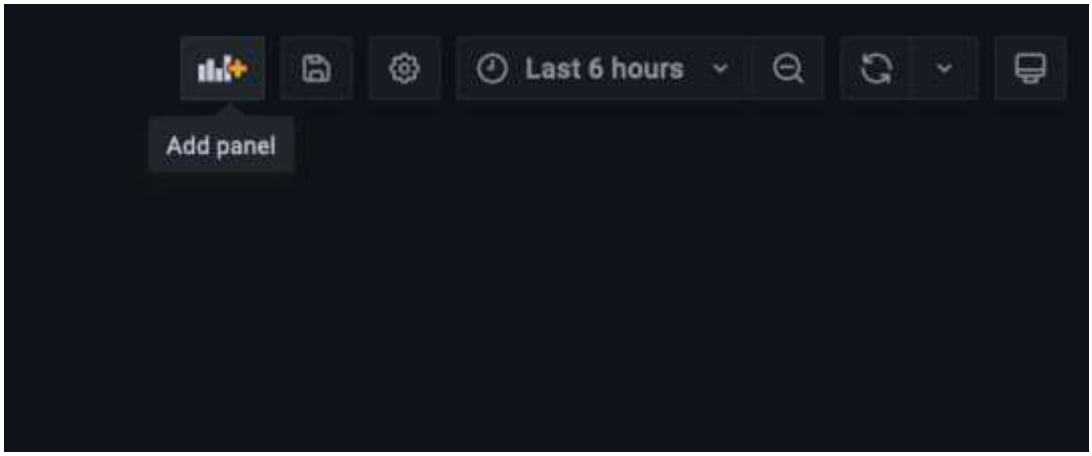


   c. Dalla dashboard della griglia, selezionare il pannello "Storage used - Object Metadata" (Storage utilizzato - metadati oggetto). Fare clic sulla piccola freccia verso il basso e sulla fine del titolo del pannello per visualizzare un menu a discesa. Da questo menu selezionare "Inspect" (ispezione) e "Panel JSON" (pannello JSON).



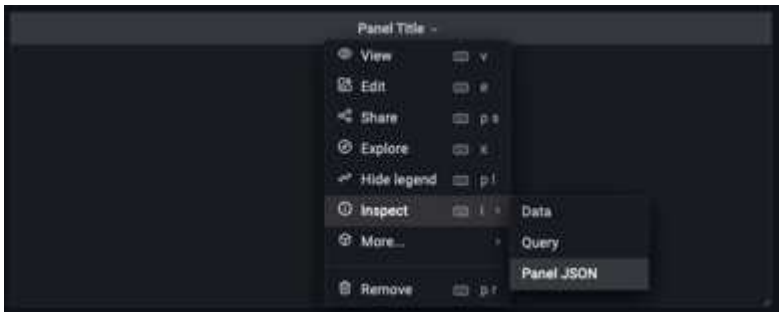
   d. Copiare il codice JSON e chiudere la finestra.

e. Nella nuova dashboard, fare clic sull'icona per aggiungere un nuovo pannello.

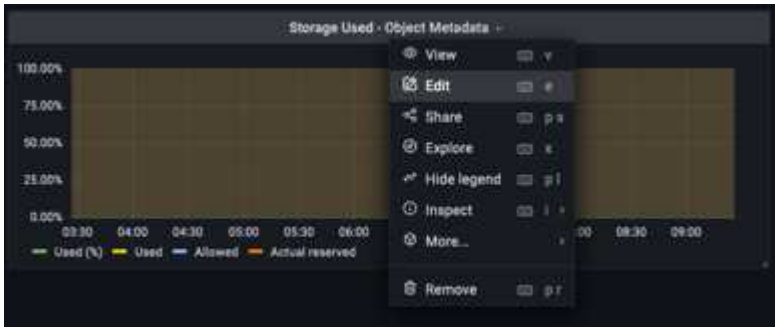


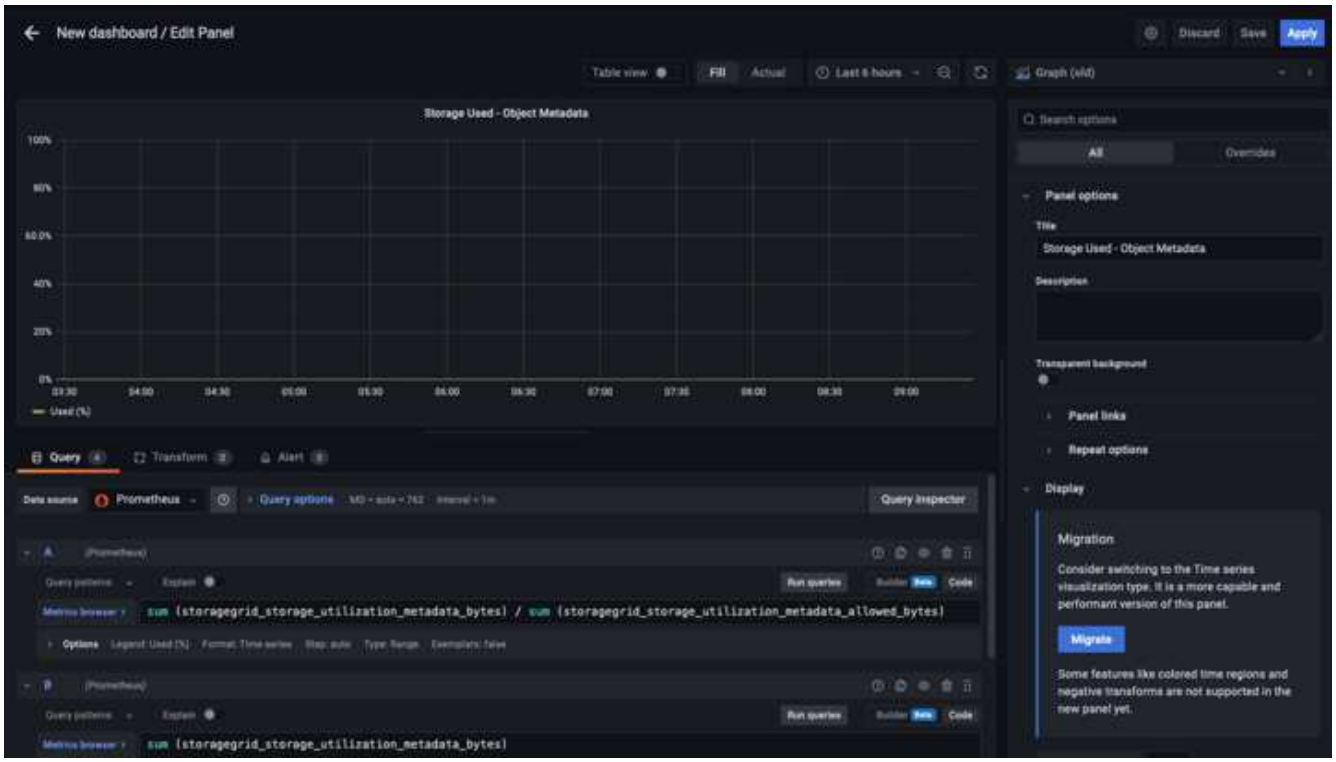
f. Applicare il nuovo pannello senza apportare modifiche

g. Proprio come per il pannello StorageGRID, controllare il JSON. Rimuovere tutto il codice JSON e sostituirlo con il codice copiato dal pannello StorageGRID.



h. Modificare il nuovo pannello e sul lato destro viene visualizzato un messaggio di migrazione con il pulsante "Migrate" (migrazione). Fare clic sul pulsante, quindi sul pulsante "Apply" (Applica).





8. Una volta che tutti i pannelli sono in posizione e configurati come si desidera. Salvare la dashboard facendo clic sull'icona del disco in alto a destra e assegnando un nome alla dashboard.

Conclusione

Ora disponiamo di un server Prometheus con capacità di storage e conservazione dei dati personalizzabili. Con questo possiamo continuare a costruire le nostre dashboard con le metriche più rilevanti per le nostre operazioni. È possibile ottenere ulteriori informazioni sulle metriche Prometheus raccolte in "Documentazione StorageGRID".

Utilizzare F5 DNS per bilanciare il carico globale StorageGRID

Di Steve Gorman (F5)

Questo report tecnico fornisce istruzioni dettagliate per la configurazione NetApp StorageGRID con i servizi DNS F5 per il bilanciamento del carico globale, al fine di garantire una migliore disponibilità dei dati, una maggiore coerenza dei dati e ottimizzare il routing delle transazioni S3 quando la griglia è distribuita su più siti e/o gruppi HA.

Introduzione

La soluzione DNS F5 BIG-IP, precedentemente denominata BIG-IP GTM (Global Traffic Manager) e informalmente GSLB (Global Server Load Balancing), consente di realizzare in modo efficace un accesso senza interruzioni tra più gruppi HA attivi-attivi e soluzioni StorageGRID multi-sito attive-attive.

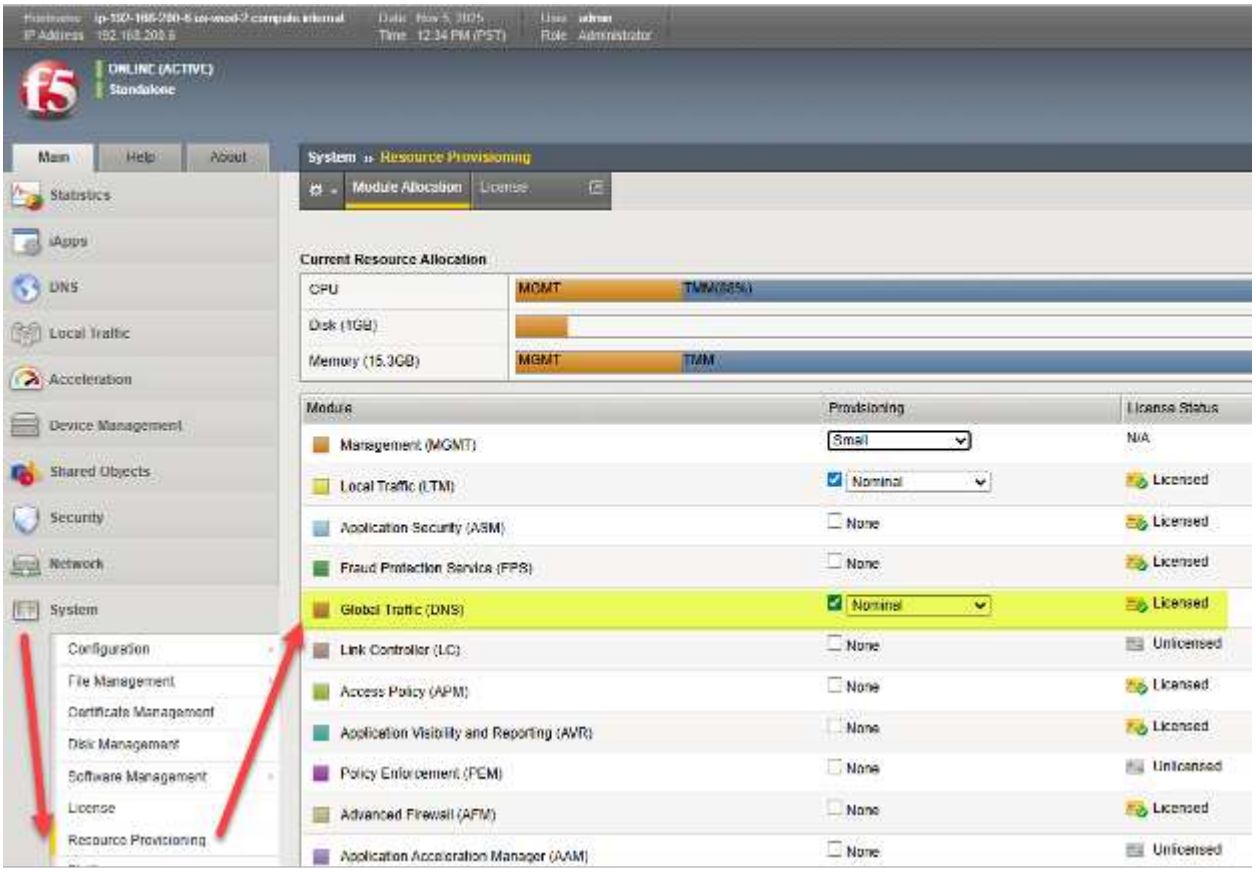
Configurazione StorageGRID multi-sito F5 BIG-IP

Indipendentemente dal numero di siti StorageGRID da supportare, almeno due appliance BIG-IP, fisiche o virtuali, devono avere il modulo DNS BIG-IP abilitato e configurato. Maggiore è il numero di dispositivi DNS,

maggiore sarà il grado di ridondanza di cui un'azienda potrà beneficiare.

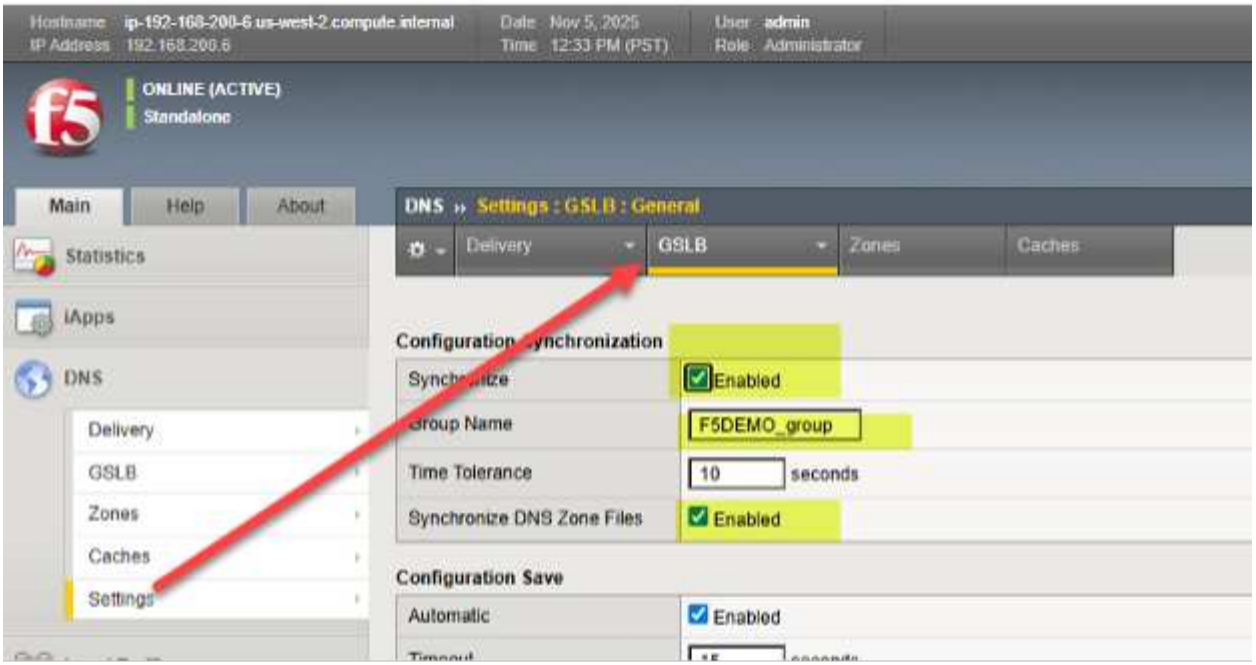
BIG-IP DNS - Primi passi nella configurazione iniziale

Una volta che l'appliance BIG-IP ha completato almeno il provisioning iniziale, utilizzare un browser Web per accedere all'interfaccia TMUI (BIG-IP GUI) e selezionare Sistema → Provisioning risorse. Come evidenziato, assicurati che il modulo "Traffico globale (DNS)" abbia un segno di spunta e che sia autorizzato. Si noti che, come nell'immagine, è comune che il "Traffico locale (LTM)" possa essere fornito sullo stesso dispositivo.

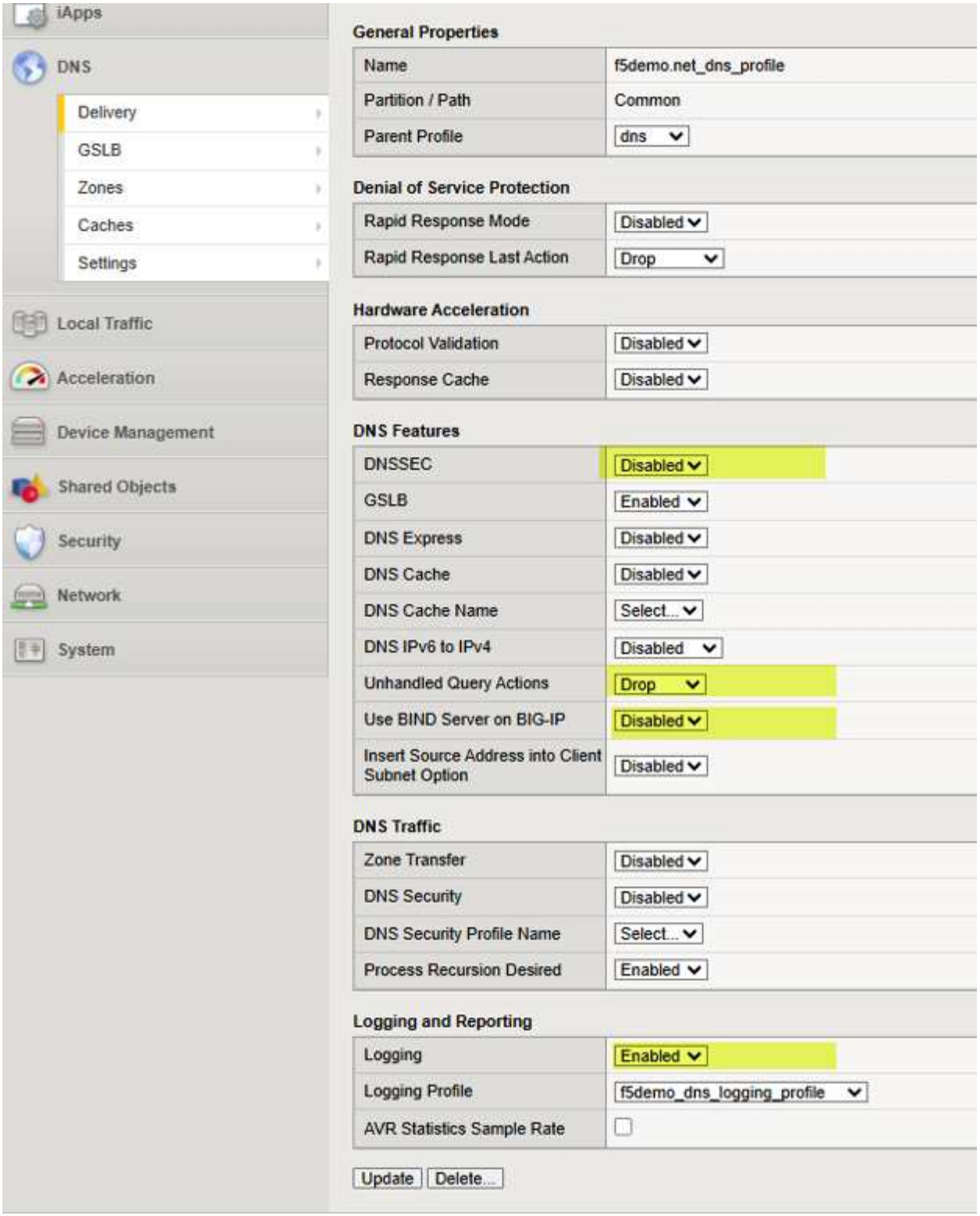


Configurare gli elementi fondamentali del protocollo DNS

Il primo passo verso la gestione del traffico globale per i siti StorageGRID è scegliere la scheda DNS, dove verrà configurato praticamente tutto il controllo del traffico globale, e scegliere Impostazioni→ GLSB. Abilita le due opzioni di sincronizzazione e scegli un nome di gruppo DNS che verrà condiviso tra gli appliance BIG-IP partecipanti.



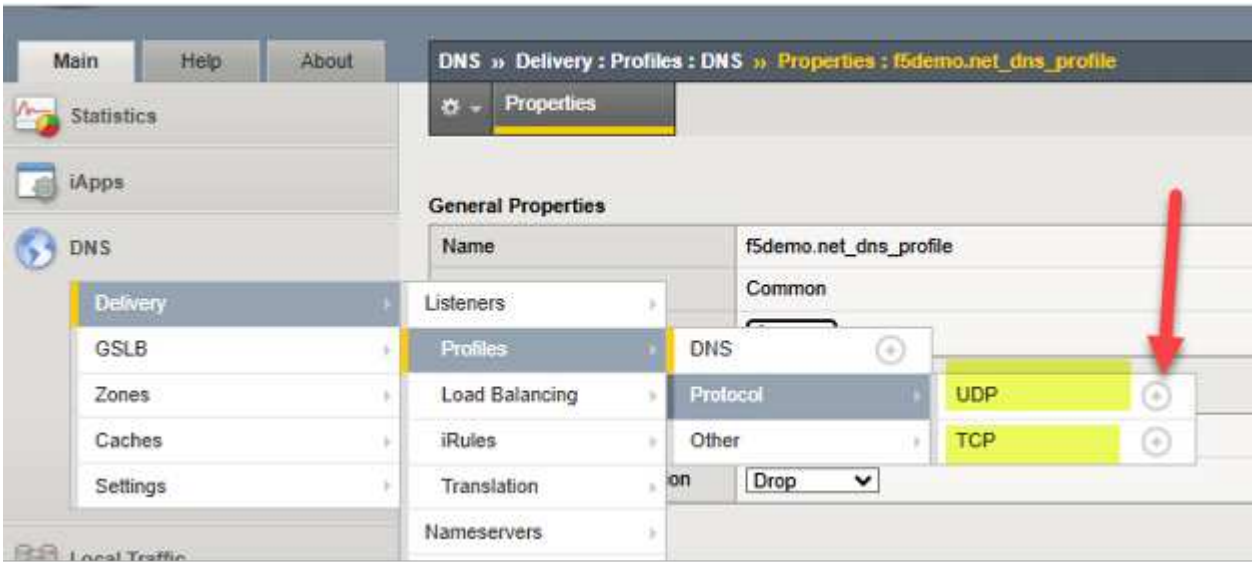
Successivamente, vai su DNS > Consegna > Profili > DNS: Crea e crea un profilo che regolerà le funzionalità DNS che desideri abilitare o disabilitare. Se si è interessati alla generazione di registri DNS specifici, consultare il collegamento precedente per la guida in aula DNS. Ecco un esempio di un profilo DNS funzionante; notare i quattro punti salienti che rappresentano impostazioni che sono valori importanti. Per maggiore consapevolezza, ogni possibile impostazione è spiegata nel seguente articolo F5 KB (Knowledge Base) "qui".



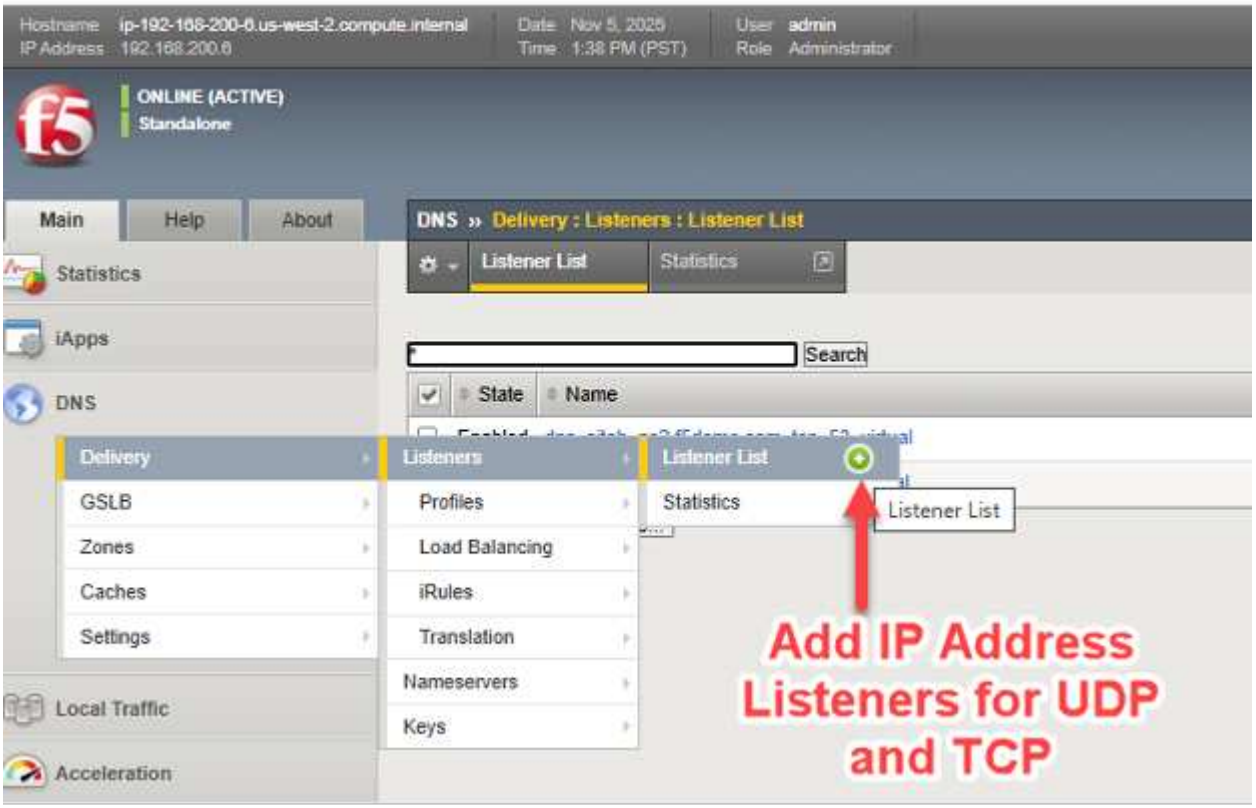
A questo punto possiamo regolare le caratteristiche dei protocolli UDP e TCP, attraverso dei “profili” creati, che possono entrambi trasportare traffico DNS che coinvolge BIG-IP. Basta creare un nuovo profilo per UDP e TCP. Supponendo che il traffico DNS attraverserà i collegamenti WAN, una buona pratica è semplicemente quella di ereditare le caratteristiche UDP e TCP note per funzionare bene negli ambienti WAN. Per aggiungerli, basta cliccare sull’icona "+" accanto a ciascun protocollo e impostare il profilo padre come segue:

UDP → usa il profilo “padre” “udp_gtm_dns”

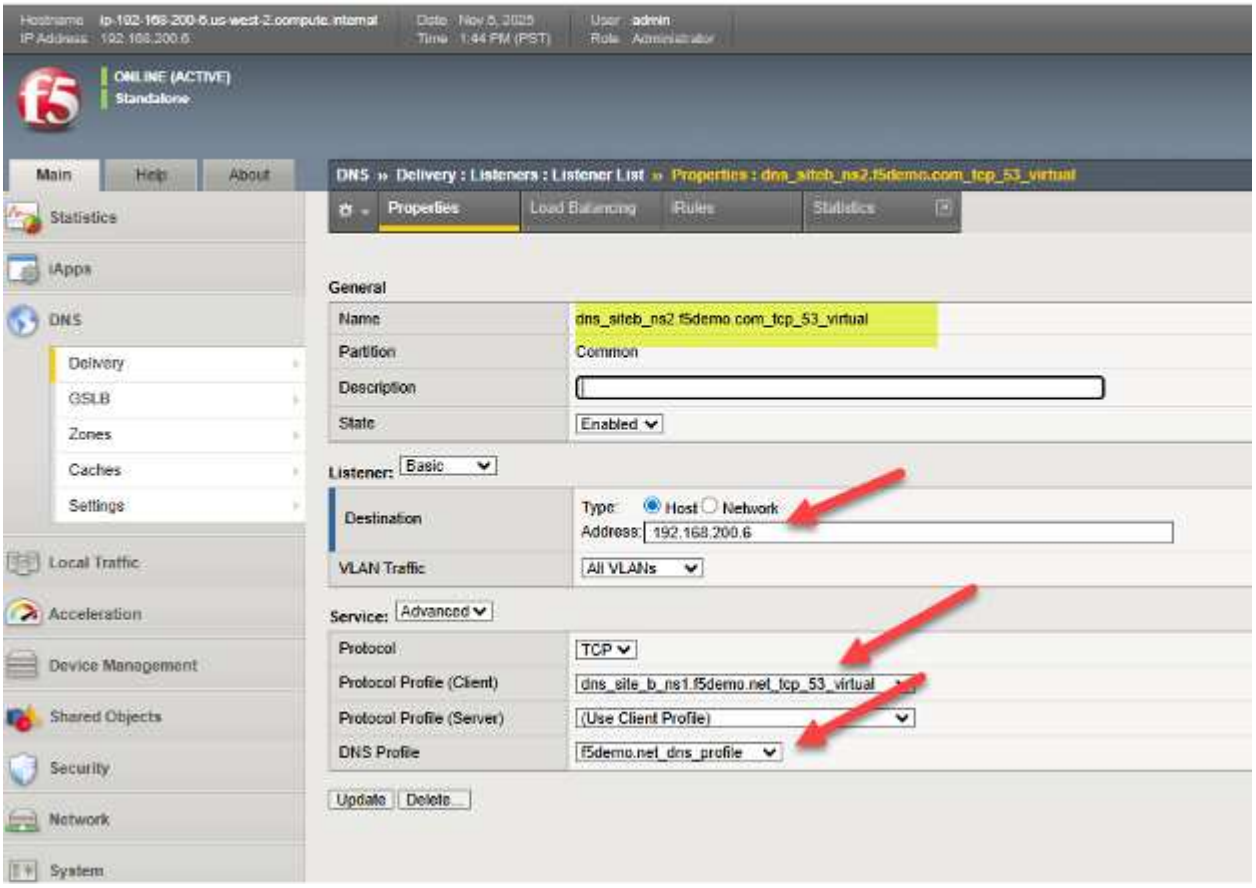
TCP → usa il profilo “padre” “f5-tcp-wan”



Ora dobbiamo semplicemente assegnare un indirizzo IP sia per il traffico UDP che per quello TCP che coinvolge il DNS BIG-IP. Per chi ha familiarità con BIG-IP LTM, si tratta essenzialmente della creazione di server virtuali DNS, e i server virtuali necessitano di indirizzi IP "in ascolto". Come nello screenshot, segui le frecce per creare server di ascolto/virtuali per DNS/UDP e DNS/TCP.



Di seguito è riportato un esempio di un DNS BIG-IP live, in cui vediamo le impostazioni dell’ascoltatore del server virtuale TCP e possiamo vedere come collega molti dei passaggi precedenti. Ciò include il riferimento al profilo DNS e al profilo del protocollo (TCP), nonché la configurazione di un indirizzo IP valido da utilizzare da parte del DNS. Come per tutti gli oggetti creati con BIG-IP, è utile utilizzare un nome significativo che serva a identificare automaticamente l’oggetto, come dns/siteb/TCP53 nell’esempio di nome assegnato.



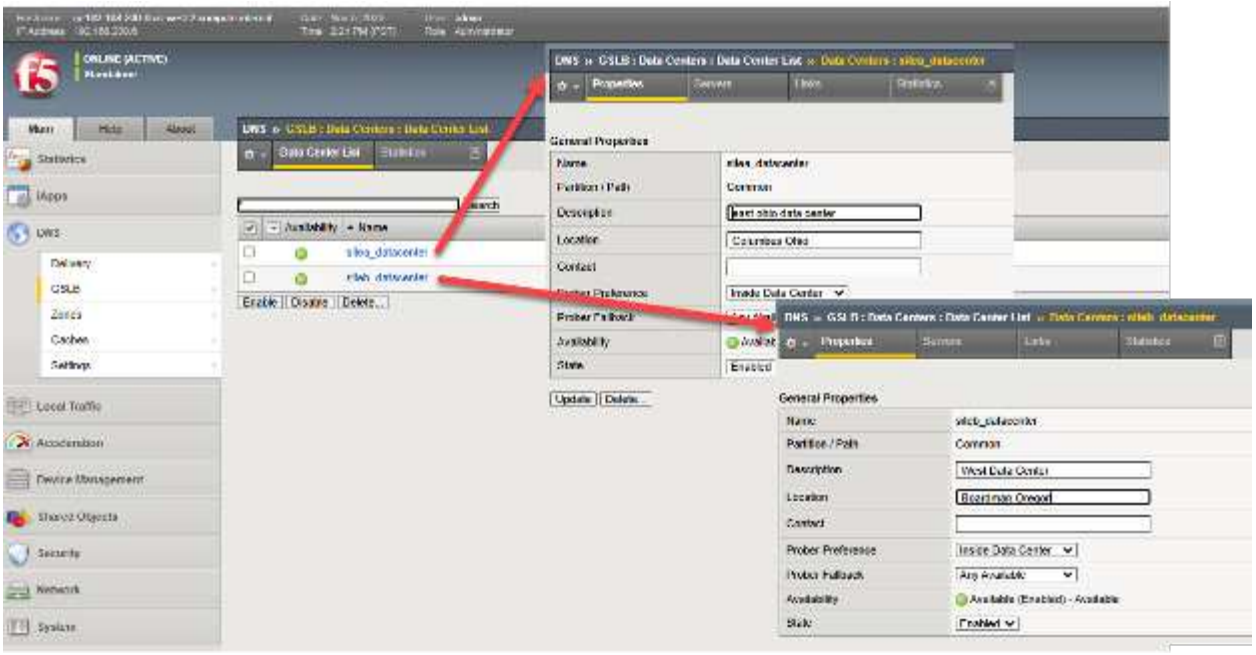
Questo conclude i passaggi preliminari, in genere "una tantum", di configurazione di un dispositivo BIG-IP con il modulo DNS abilitato. A questo punto siamo pronti per passare ai dettagli della configurazione di una soluzione di gestione del traffico globale con i nostri dispositivi, che sarà ovviamente collegata alle caratteristiche dei siti StorageGRID .

Configurazione dei siti dei data center e creazione delle comunicazioni inter-BIG-IP in quattro passaggi

Fase uno: creare data center

Ogni sito che ospiterà cluster di nodi da bilanciare localmente tramite BIG-IP LTM deve essere inserito nel DNS BIG-IP. Questa operazione deve essere eseguita su un solo DNS BIG-IP, poiché stiamo creando un gruppo DNS sincronizzato per supportare la gestione del traffico, pertanto questa configurazione sarà condivisa tra i membri DNS del gruppo.

Tramite l'interfaccia utente grafica TMUI, selezionare DNS > GSLB > Data Center > Elenco Data Center e creare una voce per ciascuno dei siti StorageGRID . Se si utilizza una configurazione di rete allineata con la Figura 1, dispositivo DNS situato in altri siti non StorageGRID , aggiungere i data center per questi siti oltre ai siti di archiviazione. In questo esempio i siti a e b sono creati in Ohio e Oregon, i BIG-IP sono dispositivi DNS e LTM doppi.



Passaggio due: creare server (elenco di tutti gli appliance BIG-IP nella soluzione)

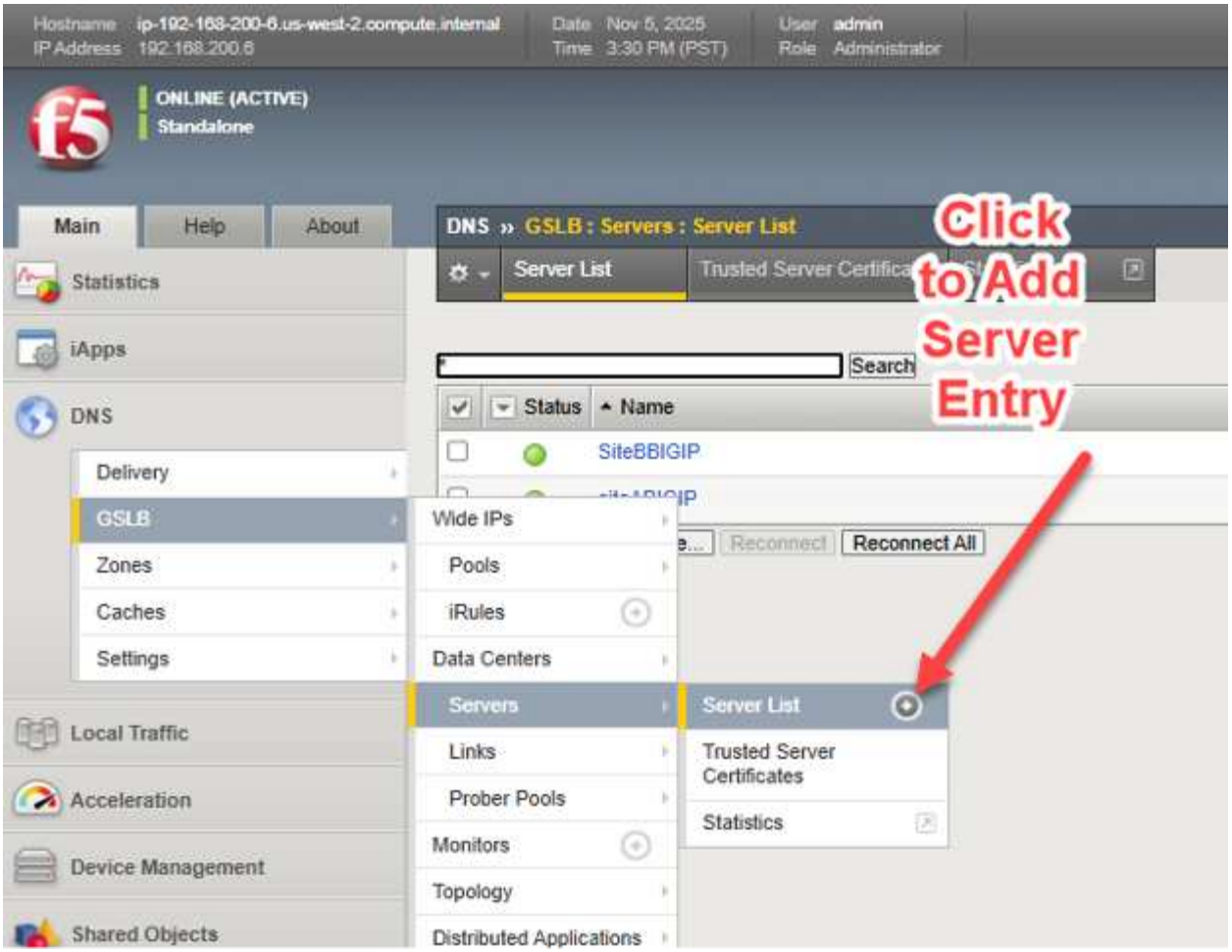
Ora siamo pronti per connettere i singoli cluster del sito StorageGRID alla configurazione DNS BIG-IP. Ricordiamo che l'appliance BIG-IP in ogni sito eseguirà l'effettivo bilanciamento del carico del traffico S3, tramite la configurazione di server virtuali che collegano un indirizzo IP/porta raggiungibile "front-end" a un set di "pool" back-end di appliance Storage Node, utilizzando indirizzi IP/porte "back-end".

Ad esempio, se tutti i nodi di archiviazione in un pool dovessero essere disconnessi a livello amministrativo, ad esempio per la dismissione di un sito o inaspettatamente a causa di controlli di integrità non riusciti in tempo reale, il traffico verrà indirizzato ad altri siti modificando le risposte alle query DNS.

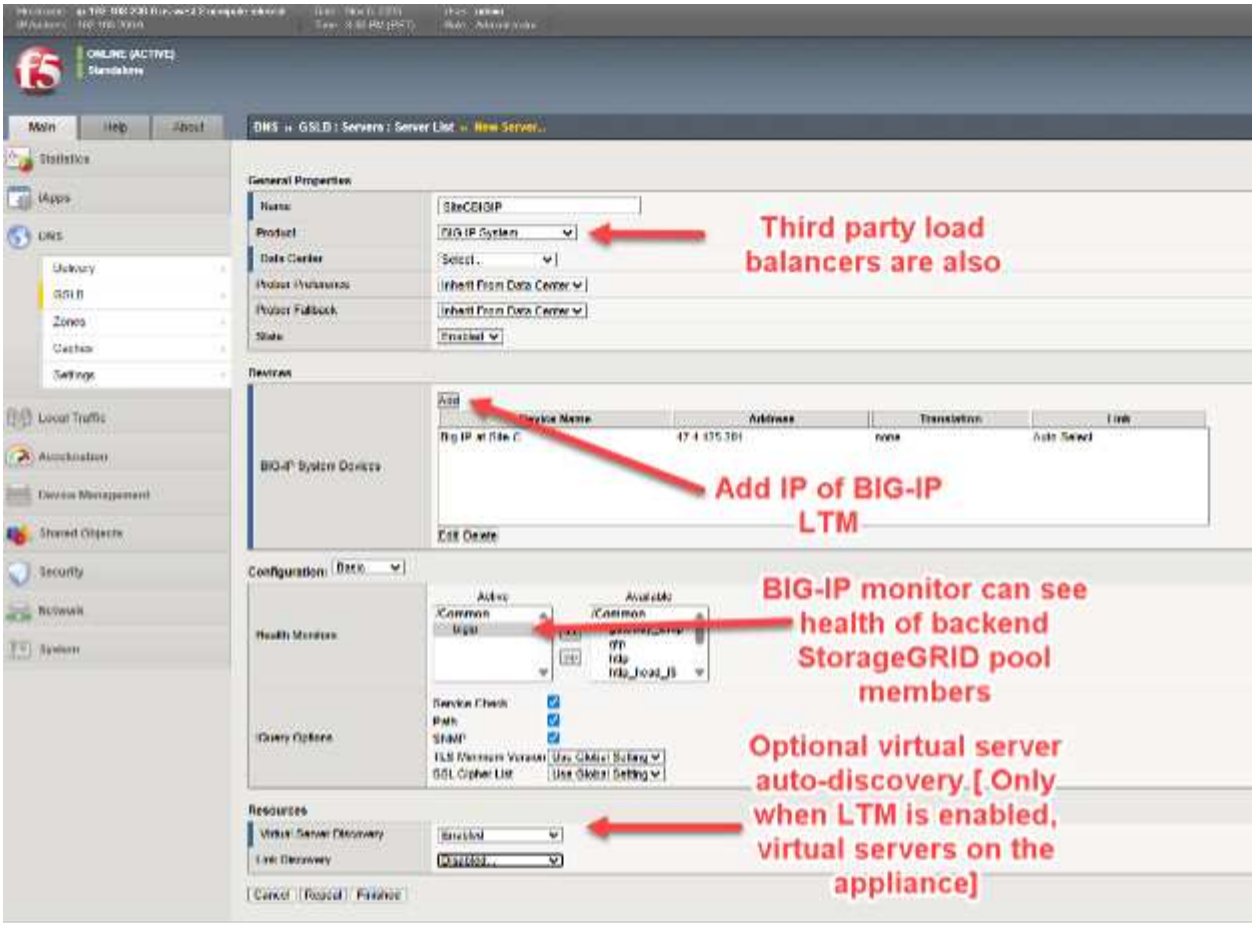
Per collegare i siti StorageGrid, in particolare i server virtuali locali, alla configurazione DNS BIG-IP su ciascun dispositivo, la configurazione deve essere eseguita una sola volta. In una fase successiva, le configurazioni dell'intero gruppo di dispositivi DNS BIG-IP saranno sincronizzate.

In parole povere, creeremo un elenco, denominato elenco server, di tutti i nostri dispositivi BIG-IP, indipendentemente dal fatto che siano dotati di licenza DNS, LTM o entrambi. Una volta completato l'elenco, questo elenco principale verrà sincronizzato con tutti gli appliance DNS BIG-IP.

Su un dispositivo con licenza DNS BIG-IP, seleziona DNS > GSLB > Server > Elenco server e fai clic sul pulsante Aggiungi (+).



I quattro elementi chiave da considerare quando si aggiunge ogni BIG-IP sono: * Selezionando BIG-IP dal menu a discesa del prodotto, sono possibili altri bilanciatori di carico, ma in genere non hanno la reattività di visibilità in tempo reale quando lo stato di salute del nodo back-end peggiora in ogni sito. * Aggiungere l'indirizzo IP dell'appliance DNS BIG-IP. Probabilmente, la prima volta che si aggiunge un'appliance DNS BIG-IP, l'indirizzo sarà quello dell'appliance corrente a cui si accede tramite GUI, mentre le appliance future saranno quelle delle altre appliance nella soluzione. * Scegliere un monitor di integrità, utilizzare sempre "BIG-IP" quando il bilanciatore del carico aggiunto è un'appliance BIG-IP, per la valutazione dello stato di integrità del nodo StorageGRID back-end. * Facoltativamente, richiedere il rilevamento automatico del server virtuale se l'appliance è un'appliance DNS/LTM doppia.

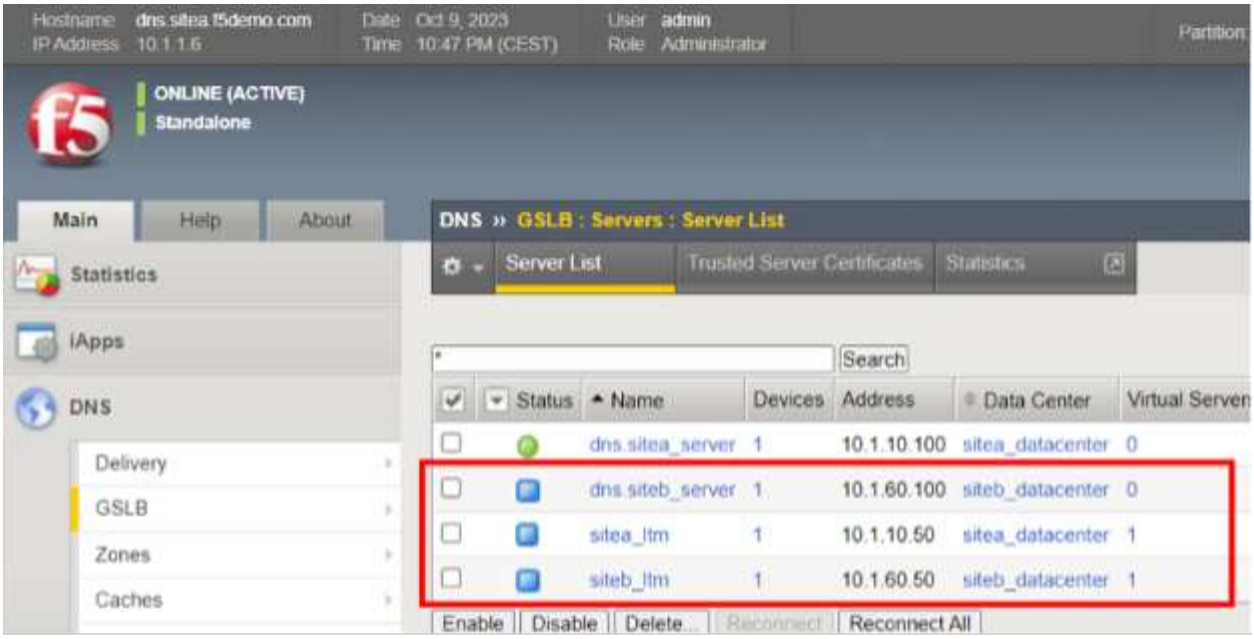


In alcune situazioni, ad esempio in caso di problemi di rete temporanei o regole ACL del firewall tra posizioni di rete, quando si aggiunge un dispositivo remoto in questa fase, la scoperta del server virtuale potrebbe non mostrare voci per i dispositivi remoti con LTM configurato. In questi casi, dopo aver aggiunto il nuovo dispositivo ("server"), è possibile aggiungere manualmente i server virtuali come indicato di seguito. Se si aggiunge un dispositivo BIG-IP DNS-only, non ci saranno server virtuali da scoprire o aggiungere a quel dispositivo.

Dobbiamo aggiungere queste voci server per ogni appliance nella nostra soluzione in tutti i siti, incluse le appliance BIG-IP DNS, le appliance BIG-IP LTM e tutte le appliance che svolgono il doppio ruolo di unità DNS e LTM.

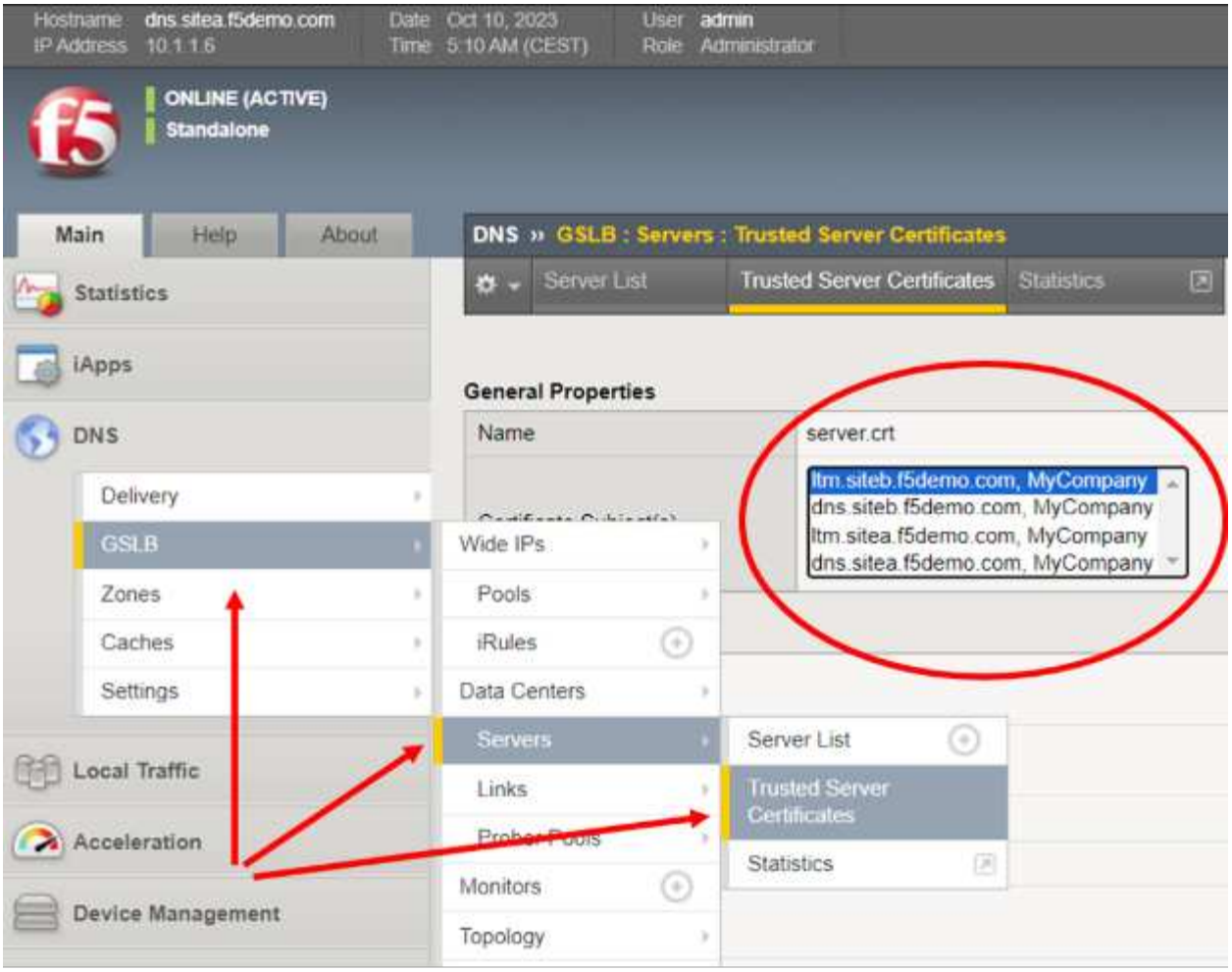
Fase tre: stabilire la fiducia tra tutti gli apparecchi BIG-IP

Nell'esempio seguente, sono stati aggiunti quattro dispositivi come server, distribuiti su due siti. Si noti che ogni sito ha un DNS BIG-IP e un LTM BIG-IP dedicati. Tuttavia, tutti gli elettrodomestici, ad eccezione di quello attualmente connesso, mostrano icone blu nella colonna "Stato". Ciò significa che non è ancora stata stabilita una relazione di fiducia con gli altri dispositivi BIG-IP.



Per aggiungere attendibilità, accedi tramite SSH al BIG-IP in cui sono stati appena inseriti i dettagli di configurazione tramite la GUI e usa l'account "root" per accedere all'interfaccia della riga di comando del BIG-IP. Emettere il seguente comando singolo al prompt: *bigip_add*

Il comando "bigip_add" estrae il certificato di gestione dai dispositivi BIGIP di destinazione per utilizzarlo durante la configurazione del canale crittografato "iQuery" tra i server GSLB nel cluster. iQuery, per impostazione predefinita, viene eseguito utilizzando la porta TCP 4353 ed è l'hearbeat che consente ai membri DNS BOG-IP di rimanere sincronizzati. Utilizza xml e gzip nel canale crittografato. Se si esegue "bigip_add" senza alcuna opzione, il comando verrà eseguito su tutti i dispositivi BIGIP nell'elenco del server GSLB utilizzando il nome utente corrente per connettersi agli endpoint. Per verificare rapidamente se l'operazione è andata a buon fine, basta tornare all'interfaccia utente grafica di BIG-IP e verificare che tutti i server abbiano ora i certificati elencati nel menu a discesa visualizzato.



Fase quattro: sincronizzare tutti gli appliance DNS BIG-IP con il gruppo DNS

Il passaggio finale consentirà di configurare completamente tutti gli apparecchi DNS BIG-IP semplicemente utilizzando l'interfaccia utente grafica TMUI di una singola unità. In un caso di esempio, in cui sono presenti due siti StorageGRID , ciò significa utilizzare SSH per raggiungere la riga di comando del DNS BIG-IP dell'**altro** sito. Dopo essersi connessi come root e aver verificato che i criteri/ACL del firewall consentano ai due dispositivi DNS BIG-IP di comunicare sulle porte TCP 22 (SSH), 443 (HTTPS) e 4354 (protocollo iQuery F5), immettere questo comando al prompt: *gtm_add <indirizzo IP del primo sito DNS BIG-IP, dove sono stati precedentemente eseguiti tutti i passaggi dell'interfaccia utente grafica>*

A questo punto, è possibile eseguire qualsiasi ulteriore lavoro di configurazione DNS su qualsiasi dispositivo DNS BIG-IP aggiunto al gruppo. Il comando precedente, gtm_add, non deve essere applicato ai membri dell'appliance che sono solo LTM. Solo gli apparecchi che supportano DNS richiedono questo comando per diventare parte del gruppo DNS sincronizzato.

Configurazione dei siti dei data center e creazione delle comunicazioni inter-BIG-IP

A questo punto, tutti i passaggi per creare il gruppo di appliance DNS BIG-IP sottostante e funzionante sono completati. Ora possiamo procedere con la creazione di nomi, FQDN, che puntano ai nostri servizi Web/S3 distribuiti esposti in ogni data center StorageGRID .

Questi nomi sono denominati "IP ampi" o WIP in breve e sono normali FQDN DNS con record di risorse DNS A. Tuttavia, anziché puntare a un server come un tradizionale record di risorse A, puntano internamente a pool di server virtuali BIG-IP. Ogni pool, singolarmente, può essere costituito da un insieme di uno o più server virtuali. Un client S3 che richiede un indirizzo IP per la risoluzione dei nomi riceverà l'indirizzo del server

virtuale S3 nel sito StorageGRID ottimale selezionato in base ai criteri.

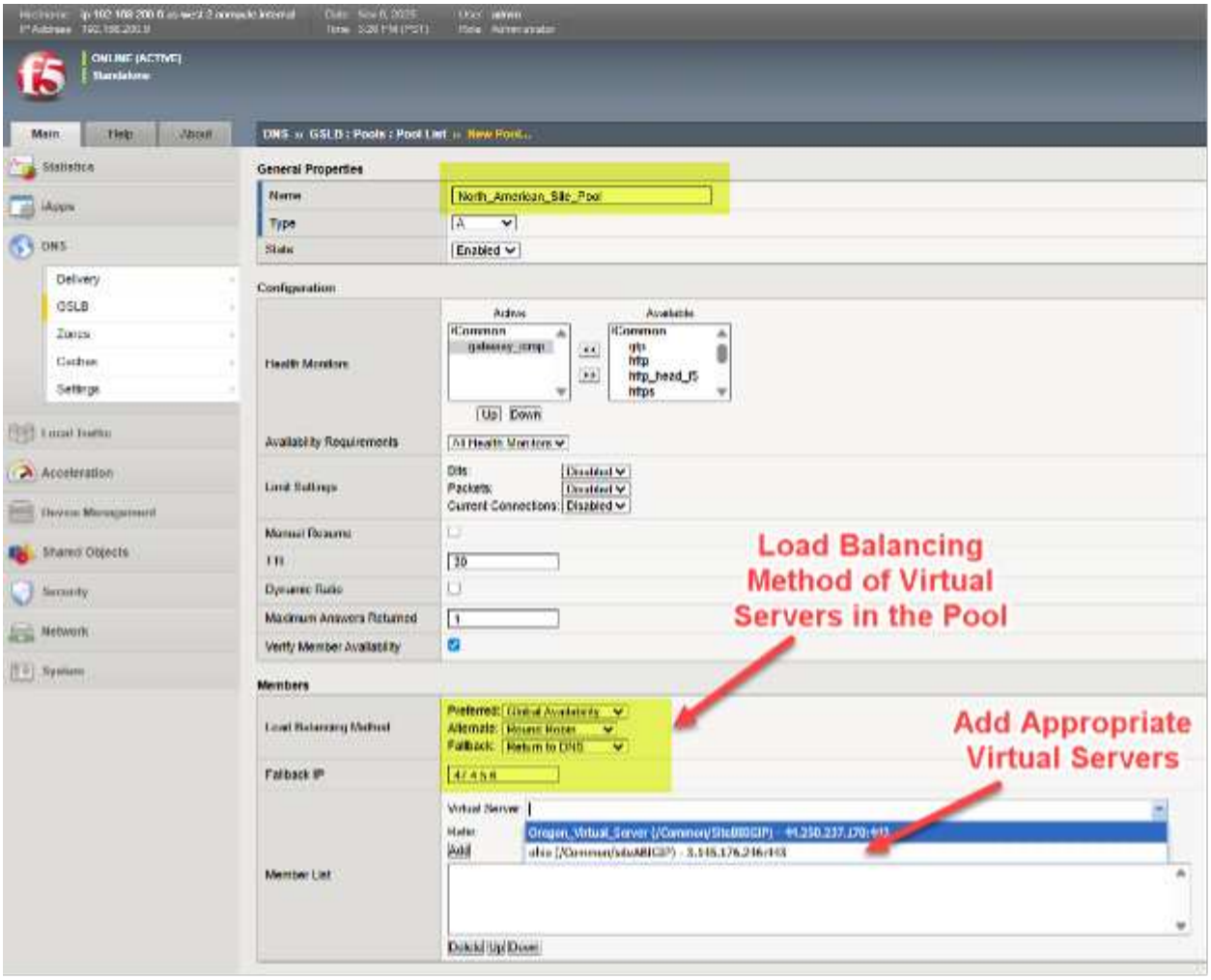
IP estesi, pool e server virtuali in breve

Per fare un esempio semplice e fittizio, un WIP per il nome **storage.quantumvault.com** potrebbe vedere la soluzione DNS BIG-IP collegata a due pool di potenziali server virtuali. Il primo pool potrebbe essere composto da 4 siti nel Nord America; il secondo pool potrebbe essere composto da 3 siti in Europa.

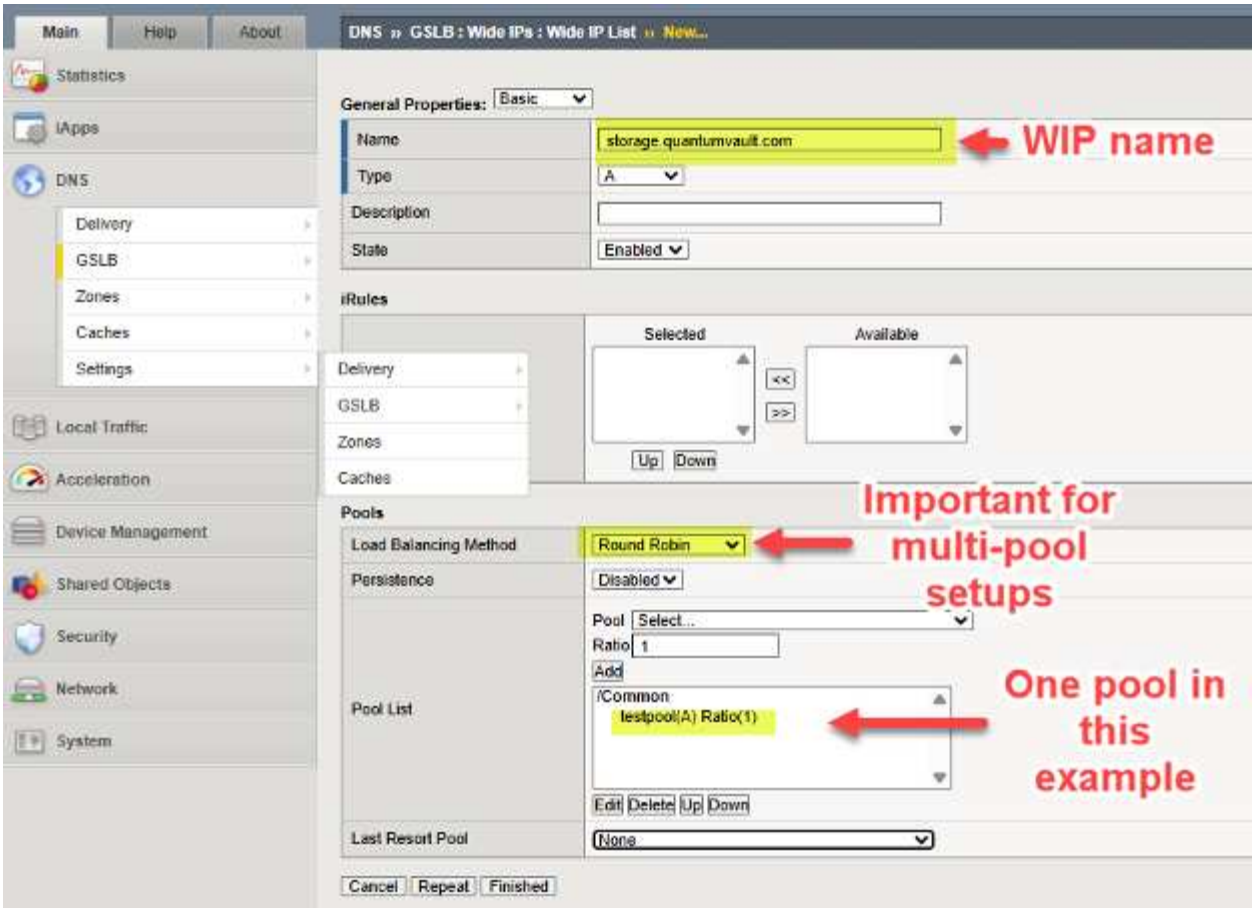
Il pool selezionato potrebbe essere ottenuto in base a una serie di decisioni politiche; forse si potrebbe utilizzare un semplice rapporto di 5:1 per indirizzare la maggior parte del traffico verso i siti StorageGRID del Nord America. Forse è più probabile una scelta basata sulla topologia, in cui il pool viene scelto in modo che, ad esempio, tutto il traffico S3 di origine europea venga indirizzato ai siti europei e il resto del traffico S3 mondiale venga indirizzato ai data center nordamericani.

Una volta che BIG-IP DNS ha individuato un pool, supponiamo che sia stato selezionato il pool nordamericano: il record di risorse DNS A effettivo restituito per risolvere storage.quantumvault.com può essere uno qualsiasi dei 4 server virtuali supportati da BIG-IP LTM in uno qualsiasi dei 4 siti nordamericani. Ancora una volta, la scelta è guidata dalle policy: esistono semplici approcci "statici" come Round-Robin, mentre selezioni "dinamiche" più avanzate come sonde di prestazioni per misurare la latenza di ciascun sito dai resolver DNS locali vengono mantenute e utilizzate come criterio per la selezione del sito.

Per impostare un pool di server virtuali su un DNS BIG-IP, seguire il percorso del menu **DNS > GSLB > Pool > Elenco pool > Aggiungi (+)**. In questo esempio, possiamo vedere che vari server virtuali nordamericani vengono aggiunti a un pool e l'approccio preferito al bilanciamento del carico, quando viene selezionato questo pool, viene scelto in modo graduale.



Aggiungiamo il WIP (Wide IP), il nome del nostro servizio che verrà risolto dal DNS, a una distribuzione seguendo DNS > GSLB > Wide IP > Wide IP List > Create (+). Nell'esempio seguente, forniamo un esempio di WIP per un servizio di archiviazione abilitato per S3.



Adattare il DNS per supportare la gestione del traffico globale

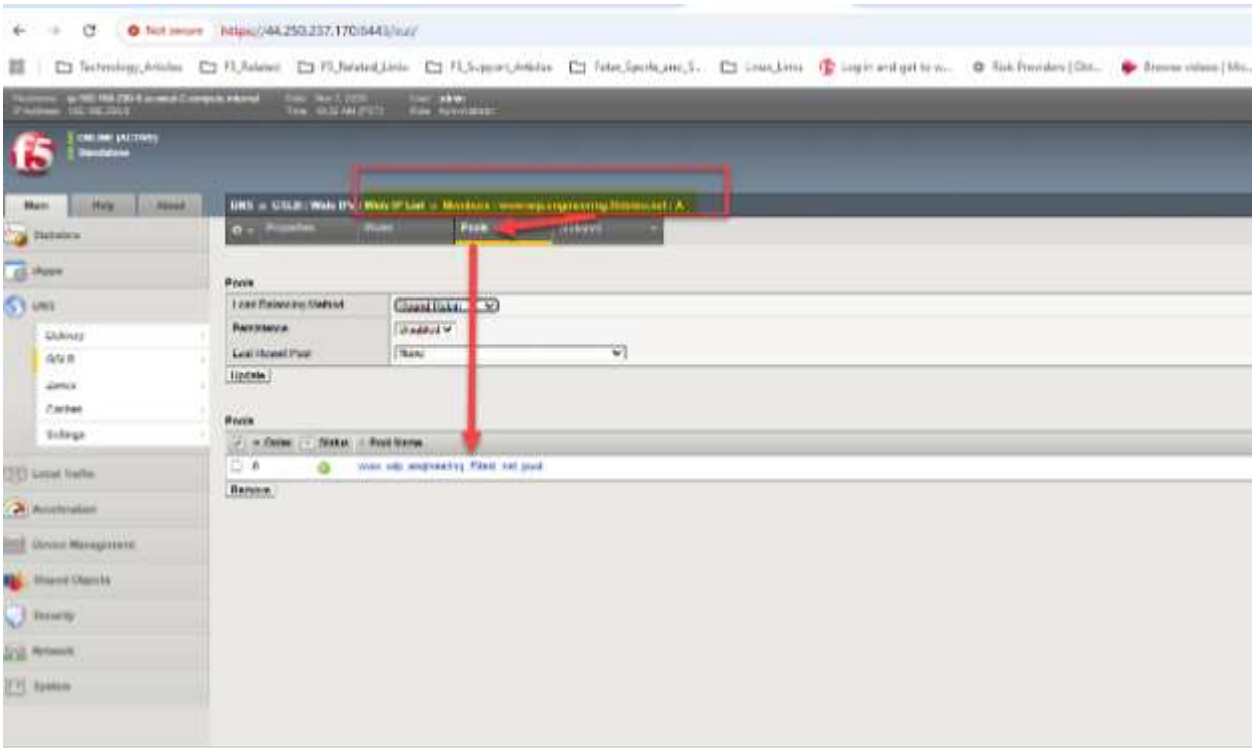
A questo punto tutti i nostri dispositivi BIG-IP sottostanti sono pronti per eseguire il GSLB (bilanciamento del carico globale del server). Per sfruttare la soluzione, dobbiamo semplicemente modificare e assegnare i nomi utilizzati per i flussi di traffico S3. L'approccio generale è quello di delegare parte di un dominio DNS esistente di un'azienda al controllo del DNS BIG-IP. Ciò significa "ritagliare" una sezione dello spazio dei nomi, un sottodominio, e delegare il controllo di questo sottodominio agli appliance DNS BIG-IP. Tecnicamente, ciò si ottiene assicurandosi che gli appliance DNS BIG-IP dispongano di record di risorse DNS (RR) nel DNS aziendale e quindi trasformando questi nomi/indirizzi in record di risorse DNS del server dei nomi (NS) per il dominio delegato.

Oggigiorno le aziende possono gestire il DNS in vari modi, uno dei quali è una soluzione completamente ospitata. Un esempio di ciò sarebbe la gestione e l'utilizzo del DNS tramite Windows Server 2025. Un approccio alternativo può essere quello di sfruttare provider cloud-DNS come AWS Route53 o Squarespace.

Ecco un esempio fittizio a scopo illustrativo. StorageGRID supporta la lettura e la scrittura di oggetti tramite il protocollo S3 con un dominio esistente gestito da AWS Route53; il dominio di esempio esistente è f5demo.net.

Vorremmo assegnare il sottodominio engineering.f5demo.net agli apparecchi DNS BIG-IP per la gestione del traffico globale. Per fare ciò, creiamo un nuovo record di risorse NS (name server) per engineering.f5demo.net e lo indirizziamo all'elenco dei nomi degli appliance DNS BIG-IP. Nel nostro esempio abbiamo due dispositivi DNS BIG-IP e pertanto creiamo per essi due record di risorse A.

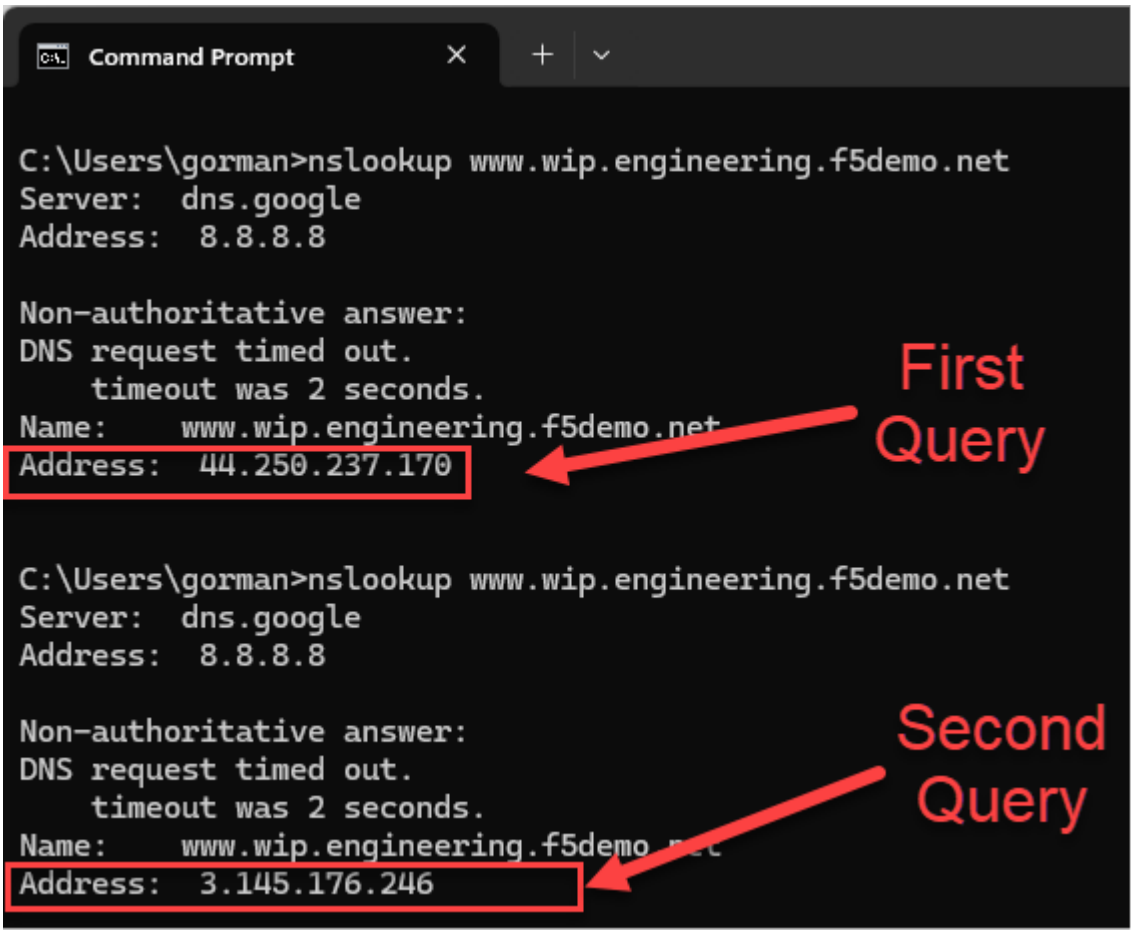
Ora, ad esempio, configureremo un IP ampio (WIP) nel nostro DNS BIG-IP, poiché il DNS utilizza la sincronizzazione di gruppo, dobbiamo solo effettuare la regolazione utilizzando l'interfaccia utente grafica di un dispositivo. All'interno dell'interfaccia utente grafica DNS BIG-IP, andare su **DNS > GSLB > IP estesi > Elenco IP estesi (+)**. Ricordiamo che in una configurazione FQDN DNS tradizionale si immetterebbero uno o più indirizzi IPv4; nel nostro caso puntiamo semplicemente a uno o più pool di server virtuali StorageGRID .



Nel nostro esempio, abbiamo server HTTPS web generici situati sia in Ohio che in Oregon. Con un semplice approccio "round robin", dovremmo essere in grado di vedere il DNS globale rispondere alle query per i mapping dei record di risorse A per *www.wip.engineering.f5demo.net* con entrambi gli IP del server virtuale.

Un test semplice può essere effettuato con i browser Web o, nel caso di S3 che utilizza StorageGRID, forse con strumenti grafici come S3Browser. Ogni query DNS vedrà il sito del data center successivo nel pool utilizzato come destinazione per il traffico successivo, grazie alla nostra scelta di Round Robin all'interno del pool.

Nella nostra configurazione di esempio, possiamo usare dig o nslookup per generare rapidamente una serie di due query DNS e assicurarci che il DNS BIG-IP stia effettivamente eseguendo un bilanciamento del carico round robin, con il risultato che entrambi i siti ricevono traffico nel tempo.

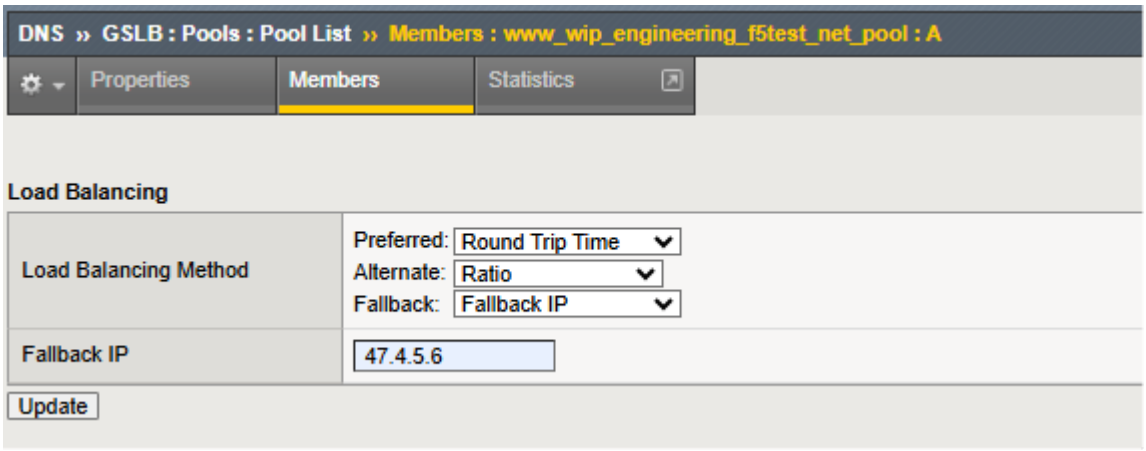


Esplorazione suggerita per tecniche più avanzate

Uno dei tanti approcci possibili prevede l'utilizzo della modalità "Disponibilità globale" anziché del semplice esempio "Round Robin" fornito sopra. Grazie alla disponibilità globale, è possibile indirizzare il traffico verso

l'ordine sequenziale dei pool o dei server virtuali all'interno di un singolo pool. In questo modo, tutto il traffico S3 potrebbe essere indirizzato per impostazione predefinita, ad esempio, verso un sito di New York City.

Se i controlli di integrità indicano un problema con la disponibilità del nodo StorageGRID in questo sito, il traffico potrebbe a quel punto essere indirizzato a St. Louis. Se St. Louis dovesse riscontrare problemi di salute, un sito a Francoforte potrebbe a sua volta iniziare a ricevere transazioni di lettura o scrittura S3. Pertanto, la disponibilità globale è un approccio alla resilienza complessiva della soluzione S3 StorageGRID . Un altro approccio consiste nel combinare e abbinare approcci di bilanciamento del carico, utilizzando un approccio a livelli.



In questo esempio, un'opzione "dinamica" è la prima scelta di bilanciamento del carico per i siti nel pool configurato. Nell'esempio mostrato, viene mantenuto un approccio di misurazione continua che utilizza il sondaggio attivo delle prestazioni del resolver DNS locale e funge da catalizzatore per la selezione del sito. Se questo approccio non fosse disponibile, i singoli siti possono essere selezionati in base al rapporto assegnato a ciascuno. Grazie al rapporto, i siti StorageGRID più grandi e con maggiore larghezza di banda possono ricevere più transazioni S3 rispetto ai siti più piccoli. Infine, come possibile scenario di disaster recovery, qualora tutti i siti nel pool diventassero non funzionanti, l'IP di fallback specificato viene utilizzato come sito di ultima istanza. Uno dei metodi di bilanciamento del carico più interessanti del DNS BIG-IP è la "Topologia", in cui viene osservata la sorgente in entrata delle query DNS, il resolver DNS locale dell'utente S3, e utilizzando le informazioni sulla topologia di Internet viene selezionato dal pool il sito apparentemente "più vicino".

Infine, se i siti si estendono in tutto il mondo, potrebbe valere la pena prendere in considerazione l'utilizzo della tecnologia "probe" dinamica discussa in dettaglio nel manuale DNS F5 BIG-IP. Con le sonde è possibile monitorare le fonti frequenti di query DNS, ad esempio un partner business-to-business il cui traffico utilizza generalmente lo stesso resolver DNS locale. Le sonde DNS BIG-IP possono essere avviate dal BIG-IP LTM in ogni sito nel mondo, per determinare in generale quale sito potenziale offrirebbe la latenza più bassa per le transazioni S3. Pertanto, il traffico proveniente dall'Asia potrebbe essere meglio servito dai siti StorageGRID asiatici rispetto ai siti situati in Nord America o in Europa.

Conclusione

L'integrazione di F5 BIG-IP con NetApp StorageGRID risolve le sfide tecniche legate alla disponibilità e alla coerenza dei dati su più siti e all'ottimizzazione del routing delle transazioni S3. L'implementazione di questa soluzione migliora la resilienza, le prestazioni e l'affidabilità dello storage, rendendola ideale per le aziende che cercano un'infrastruttura di storage solida, scalabile e flessibile.

Per saperne di più, la documentazione ufficiale F5 per BIG-IP DNS può essere trovata qui "collegamento". È inoltre possibile trovare una guida guidata allo stile di classe che fornisce istruzioni dettagliate su un esempio di configurazione "qui".

Configurazione SNMP Datadog

Di Aron Klein

Configurare Datadog per raccogliere le metriche e i trap snmp di StorageGRID.

Configurare Datadog

Datadog è una soluzione di monitoraggio che fornisce metriche, visualizzazioni e avvisi. La seguente configurazione è stata implementata con l'agente linux versione 7.43.1 su un host Ubuntu 22.04.1 distribuito localmente nel sistema StorageGRID.

Profilo Datadog e file trap generati dal file MIB StorageGRID

Datadog fornisce un metodo per convertire i file MIB del prodotto in file di riferimento datadog necessari per mappare i messaggi SNMP.

Questo file yaml di StorageGRID per la mappatura della risoluzione del trap Datadog generato in base alle istruzioni trovate "qui". + inserire questo file in /etc/datadog-Agent/conf.d/snmp.d/trap_db/ +

- "Scaricare il file yaml trap" +
  - **checksum md5** 42e27e4210719945a46172b98c379517 +
  - **sha256 checksum** d0fe5c8e6ca3c902d054f854b70a85f928cba8b7c76391d356f05d2cf73b6887 +

Questo file yaml del profilo StorageGRID per la mappatura delle metriche Datadog generato in base alle istruzioni trovate "qui". + inserire questo file in /etc/datadog-Agent/conf.d/snmp.d/profiles/ +

- "Scarica il file yaml del profilo" +
  - **checksum md5** 72b7784f4801adda4e0c3ea77df19aa +
  - **sha256 checksum** b6b7fadd33063422a8bb8e39b3ead8ab38349ee02229926eadc8585f0087b8cee +

Configurazione Datadog SNMP per metriche

La configurazione di SNMP per le metriche può essere gestita in due modi. È possibile configurare il rilevamento automatico fornendo un intervallo di indirizzi di rete contenente i sistemi StorageGRID o definendo gli IP dei singoli dispositivi. La posizione della configurazione è diversa in base alla decisione presa. Il rilevamento automatico viene definito nel file yaml dell'agente datadog. Le definizioni esplicite dei dispositivi vengono configurate nel file yaml di configurazione snmp. Di seguito sono riportati alcuni esempi di ciascuno per lo stesso sistema StorageGRID.

Rilevamento automatico

la configurazione si trova in /etc/datadog-agent/datadog.yaml

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

Singoli dispositivi

/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'
```

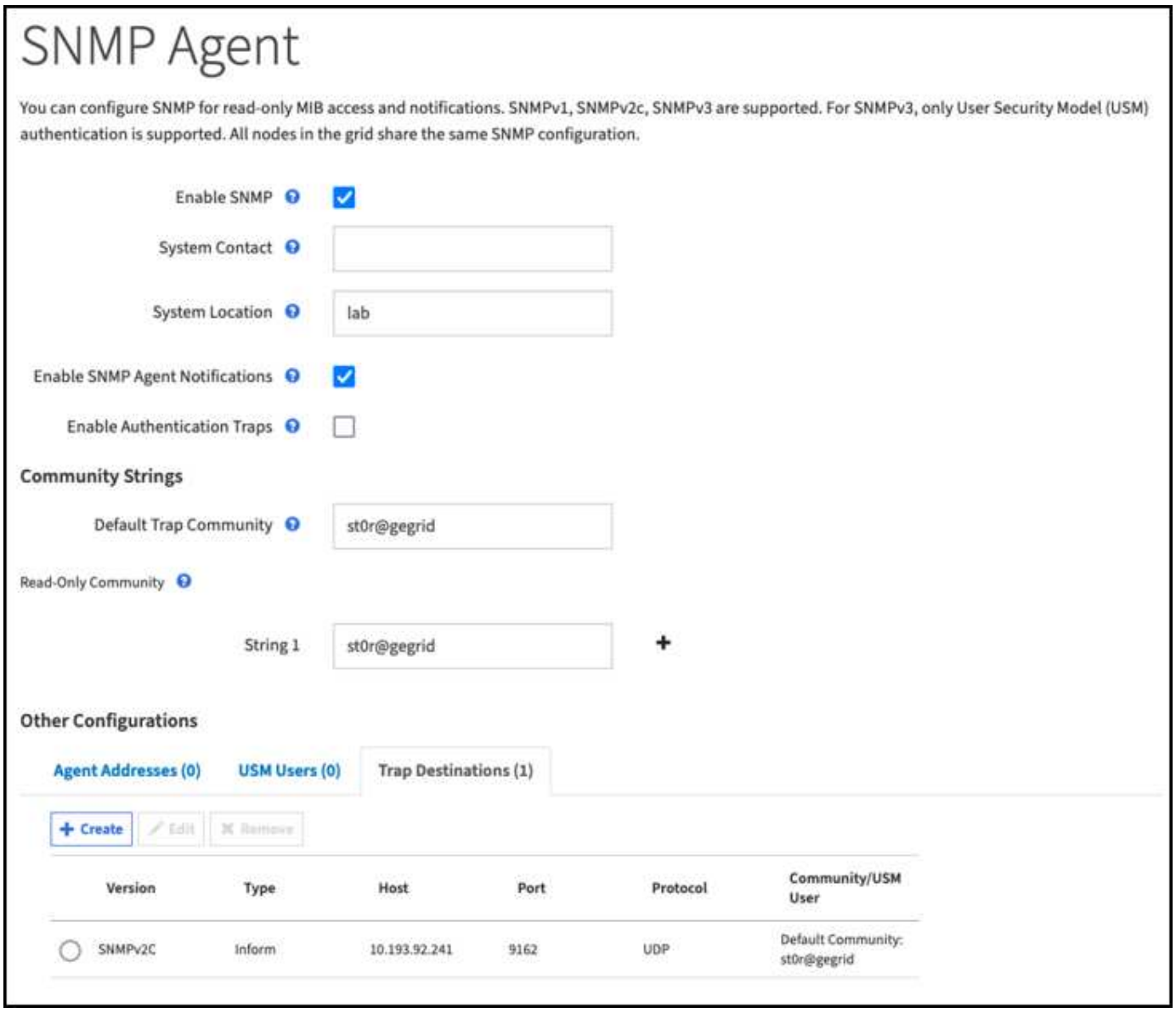
Configurazione SNMP per i trap

La configurazione per i trap SNMP è definita nel file yaml di configurazione del datadog /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
```

Esempio di configurazione SNMP StorageGRID

L'agente SNMP nel sistema StorageGRID si trova nella scheda di configurazione, colonna Monitoring (monitoraggio). Attivare SNMP e immettere le informazioni desiderate. Se si desidera configurare i trap, selezionare "Destinations trap" (Destinazioni trap) e creare una destinazione per l'host dell'agente Datadog contenente la configurazione trap.



Utilizzare rclone per migrare, INSERIRE ed ELIMINARE oggetti su StorageGRID

Di Siegfried Hepp e Aron Klein

Rclone è un tool e client a riga di comando gratuito per le operazioni S3. È possibile utilizzare rclone per migrare, copiare ed eliminare i dati degli oggetti su StorageGRID. rclone include la possibilità di eliminare i bucket anche quando non sono vuoti con una funzione di "purge", come illustrato nell'esempio riportato di seguito.

Installare e configurare rclone

Per installare rclone su una workstation o su un server, scaricarlo da "rclone.org".

Fasi iniziali della configurazione

1. Creare il file di configurazione rclone eseguendo lo script di configurazione o creando manualmente il file.
2. In questo esempio userò sgdemo come nome dell'endpoint remoto di StorageGRID S3 nella configurazione rclone.
   a. Creare il file di configurazione ~/.config/rclone/rclone.conf

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

   b. Eseguire rclone config

config. rclone

```
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
```

Esempi di comandi di base

- **Creare un bucket:**

  `rclone mkdir remote:bucket`

  ```
  mkdir sgdemo:test01
  ```

  > Utilizzare --no-check-certificate se si desidera ignorare i certificati SSL.

- **Elenca tutti i bucket:**

```
rclone lsd remote:
```

1. rclone lsd sgdemo:

- **Elenca oggetti in un bucket specifico:**

```
rclone ls remote:bucket
```

rclone ls sgdemo:test01

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

- **Eliminare un bucket:**

```
rclone rmdir remote:bucket
```

rclone rmdir sgdemo:test02

- **Mettere un oggetto:**

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

~/test/testfile.txt sgdemo:test01

- **Ottenere un oggetto:**

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

~/testfile.txt test/testfileS3.txt

- **Elimina un oggetto:**

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

rclone delete sgdemo:test01/testfile.txt

- **Migrare oggetti in un bucket**

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

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

rclone sync sgdemo:test01 sgdemo:clone01 --progress

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

Utilizzare --Progress o -P per visualizzare l'avanzamento dell'attività. In caso contrario, non viene visualizzato alcun output.

- **Elimina un bucket e tutti i contenuti degli oggetti**

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

rclone purge sgdemo:test01 --progress

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

rclone ls sgdemo:test01

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

Best practice di StorageGRID per l'implementazione con Veeam Backup and Replication

Di Oliver Haensel e Aron Klein

Questa guida si concentra sulla configurazione di NetApp StorageGRID e, in parte, su Veeam Backup and Replication. Questo documento è stato scritto per gli amministratori di storage e rete che hanno familiarità con i sistemi Linux e hanno il compito di mantenere o implementare un sistema NetApp StorageGRID in combinazione con Veeam Backup and Replication.

Panoramica

Gli amministratori dello storage cercano di gestire la crescita dei propri dati con soluzioni che soddisfino la disponibilità, gli obiettivi di recovery rapido, scalino per soddisfare le loro esigenze e automatizzino la policy per la conservazione dei dati a lungo termine. Queste soluzioni devono anche fornire protezione da perdite o attacchi dannosi. Insieme, Veeam e NetApp hanno avviato una partnership per creare una soluzione di data Protection che combina Veeam Backup & Recovery con NetApp StorageGRID per lo storage a oggetti on-premise.

Veeam e NetApp StorageGRID offrono una soluzione facile da utilizzare che lavorano insieme per contribuire a soddisfare le richieste di una rapida crescita dei dati e di maggiori normative in tutto il mondo. Lo storage a oggetti basato sul cloud è celebre per la sua resilienza, la sua capacità di scalare, le efficienze operative e i costi che lo rendono la scelta naturale come destinazione dei backup. Questo documento fornirà linee guida e consigli per la configurazione della soluzione Veeam Backup e del sistema StorageGRID.

Il workload a oggetti di Veeam crea un elevato numero di operazioni simultanee di PUT, DELETE ed LIST di piccoli oggetti. L'attivazione dell'immutabilità aumenta il numero di richieste all'archivio oggetti per l'impostazione della conservazione e dell'elenco delle versioni. Il processo di un processo di backup include la scrittura degli oggetti per la modifica giornaliera, quindi, una volta completate le nuove scritture, il processo eliminerà tutti gli oggetti in base al criterio di conservazione del backup. La pianificazione dei processi di backup si sovrappone quasi sempre. Questa sovrapposizione risulterà in un'ampia parte della finestra di backup che consiste in un carico di lavoro PUT/DELETE di 50/50 KB sull'archivio di oggetti. Rettificare in Veeam il numero di operazioni simultanee con l'impostazione dello slot di attività, aumentando la dimensione dell'oggetto aumentando la dimensione del blocco di lavoro di backup, riducendo il numero di oggetti nelle richieste di eliminazione multioggetto, inoltre, la scelta della finestra temporale massima per il completamento dei lavori ottimizzerà la soluzione in termini di prestazioni e costi.

Assicurati di leggere la documentazione del prodotto per "Backup e replica di Veeam" E "StorageGRID" prima di iniziare. Veeam fornisce calcolatori per comprendere il dimensionamento dell'infrastruttura Veeam e i requisiti di capacità da utilizzare prima di dimensionare la soluzione StorageGRID . Si prega di controllare sempre le configurazioni convalidate Veeam- NetApp sul sito Web del programma Veeam Ready per "Veeam Ready Object, immutabilità degli oggetti e Repository" .

Configurazione Veeam

Versione consigliata

Si consiglia sempre di restare aggiornati e di applicare gli aggiornamenti rapidi più recenti per il sistema Veeam Backup & Replication 12 o 12,1. Attualmente consigliamo almeno di installare Veeam 12 patch P20230718.

S3 Configurazione del repository

Un repository di backup scale-out (SOBR) è il Tier di capacità dello storage a oggetti S3. Il Tier di capacità è un'estensione del repository primario che offre periodi di conservazione dei dati più lunghi e una soluzione di storage a costi inferiori. Veeam offre la possibilità di fornire immutabilità tramite l'API S3 Object Lock. Veeam 12 può utilizzare bucket multipli in un repository scale-out. StorageGRID non ha un limite per il numero di oggetti o la capacità in un singolo bucket. L'utilizzo di bucket multipli può migliorare le performance durante il backup di set di dati molto grandi, dove i dati di backup possono raggiungere la scalabilità di petabyte in oggetti.

In base al dimensionamento della soluzione e ai requisiti specifici, potrebbe essere necessario limitare le attività simultanee. Le impostazioni predefinite specificano uno slot di attività del repository per ogni core della CPU e per ogni slot di attività un limite di 64 slot di attività simultanei. Ad esempio, se il server dispone di 2 core CPU, per l'archivio oggetti verrà utilizzato un totale di 128 thread simultanei. Questo include PUT, GET e Batch Delete. Si consiglia di selezionare un limite conservativo agli slot di attività da utilizzare per iniziare e regolare questo valore una volta che i backup Veeam hanno raggiunto lo stato stabile dei nuovi backup e dei dati di backup in scadenza. Collabora con l'account team di NetApp per dimensionare il sistema StorageGRID in modo appropriato e soddisfare le finestre temporali e le performance desiderate. Per fornire la soluzione ottimale, potrebbe essere necessario regolare il numero di slot di attività e il limite di attività per slot.

Configurazione del processo di backup

I job di backup Veeam possono essere configurati con diverse opzioni di dimensione del blocco che devono essere prese in considerazione con attenzione. Le dimensioni predefinite del blocco sono di 1MB KB e, grazie all'efficienza dello storage, Veeam offre funzionalità di compressione e deduplica crea dimensioni degli oggetti di circa 500KB KB per il backup completo iniziale e oggetti 100-200kB per i job incrementali. Possiamo aumentare enormemente le performance e ridurre i requisiti per l'archivio di oggetti scegliendo una dimensione maggiore dei blocchi di backup. Sebbene le dimensioni maggiori dei blocchi apportino notevoli miglioramenti nelle performance dell'archivio di oggetti, si presenta al costo di potenzialmente aumentare i requisiti di capacità dello storage primario grazie alle ridotte performance di efficienza dello storage. Si consiglia di configurare i processi di backup con una dimensione blocco di 4MB KB che crea circa 2MB oggetti per i backup completi e dimensioni oggetto di 700kB-1MB KB per i backup incrementali. I clienti possono prendere in considerazione anche la configurazione dei lavori di backup utilizzando blocchi di 8 MB, che possono essere abilitati con l'assistenza del supporto Veeam.

L'implementazione di backup immutabili utilizza S3 Object Lock nell'archivio oggetti. L'opzione immutabilità genera un numero maggiore di richieste all'archivio oggetti per l'elenco e la conservazione degli aggiornamenti sugli oggetti.

Alla scadenza delle trattenute di backup, i lavori di backup elaboreranno l'eliminazione degli oggetti. Veeam invia le richieste di eliminazione all'archivio oggetti nelle richieste di eliminazione di più oggetti di 1000 oggetti

per richiesta. Per le soluzioni di piccole dimensioni, potrebbe essere necessario regolarle per ridurre il numero di oggetti per richiesta. La riduzione di questo valore avrà il vantaggio di distribuire in modo più uniforme le richieste di eliminazione tra i nodi nel sistema StorageGRID. Si consiglia di utilizzare i valori nella tabella seguente come punto di partenza per configurare il limite di eliminazione di più oggetti. Moltiplicare il valore nella tabella per il numero di nodi per il tipo di appliance scelto per ottenere il valore per l'impostazione in Veeam. Se questo valore è uguale o superiore a 1000 non è necessario regolare il valore predefinito. Se questo valore deve essere modificato, si prega di collaborare con il supporto di Veeam per apportare la modifica.

Modello di appliance	S3MultiObjectDeleteLimit per nodo
SG5712	34
SG5760	75
SG6060	200

Contatta il tuo account team NetApp per ottenere la configurazione consigliata in base alle tue esigenze specifiche. Le raccomandazioni sulle impostazioni di configurazione di Veeam includono:

- Dimensione blocco processo di backup = 4MB
- Limite slot attività SOBR= 2-16
- Limite eliminazione oggetti multipli = 34-1000

Configurazione StorageGRID

Versione consigliata

Per le distribuzioni Veeam sono consigliate le versioni NetApp StorageGRID 11.9 o 12.0 con l'ultimo hotfix. Si consiglia sempre di rimanere aggiornati e di applicare gli ultimi hotfix per il sistema StorageGRID .

Bilanciamento del carico e configurazione dell'endpoint S3

Veeam richiede che l'endpoint sia connesso solo tramite HTTPS. Una connessione non crittografata non è supportata da Veeam. Il certificato SSL può essere un certificato autofirmato, un'autorità di certificazione privata attendibile o un'autorità di certificazione pubblica attendibile. Per garantire un accesso continuo al repository S3, si consiglia di utilizzare almeno due bilanciatori del carico in una configurazione ha. I bilanciatori del carico possono essere un servizio di bilanciamento del carico integrato fornito da StorageGRID situato in ogni nodo amministrativo e nodo gateway o soluzione di terze parti come F5, Kemp, HAproxy, Loadbalancer.org, ecc. L'utilizzo di un bilanciatore del carico StorageGRID fornirà la possibilità di impostare classificatori di traffico (regole QoS) che possono dare priorità al workload Veeam, o limitare Veeam a non influire sui carichi di lavoro a priorità più alta sul sistema StorageGRID.

Bucket S3

StorageGRID è un sistema di archiviazione multi-tenant sicuro. Si consiglia di creare un tenant dedicato per il carico di lavoro Veeam. Facoltativamente, è possibile assegnare una quota di archiviazione. Come buona pratica, abilitare "utilizza la propria fonte di identità". Proteggere l'utente di gestione root del tenant con una password appropriata. Veeam Backup 12 richiede una forte coerenza per i bucket S3. StorageGRID offre molteplici opzioni di coerenza configurate a livello di bucket. Per distribuzioni multi-sito con Veeam che accede ai dati da più posizioni, selezionare "strong-global". Se i backup e i ripristini Veeam vengono eseguiti solo in un singolo sito, il livello di coerenza deve essere impostato su "strong-site". Per ulteriori informazioni sui livelli di

coerenza del bucket, consultare "documentazione" . Per utilizzare StorageGRID per i backup di immutabilità di Veeam, S3 Object Lock deve essere abilitato globalmente e configurato sul bucket durante la creazione del bucket.

Gestione del ciclo di vita

StorageGRID supporta la replica e l'erasure coding per la protezione a livello di oggetto in siti e nodi StorageGRID. L'erasure coding richiede almeno una dimensione dell'oggetto di 200kB KB. Le dimensioni predefinite dei blocchi per Veeam di 1MB producono dimensioni degli oggetti che possono spesso essere inferiori a questa dimensione minima consigliata di 200kB KB dopo le efficienze di storage di Veeam. Per le performance della soluzione, non è consigliabile utilizzare un profilo di erasure coding su più siti, a meno che la connettività tra i siti non sia sufficiente per non aggiungere latenza o limitare la larghezza di banda del sistema StorageGRID. In un sistema StorageGRID multisito, la regola ILM può essere configurata per memorizzare una singola copia in ciascun sito. Per garantire la massima durata, è possibile configurare una regola per memorizzare una copia con erasure coding in ogni sito. L'utilizzo di due copie locali nei server Veeam Backup è l'implementazione più consigliata per questo workload.

Elimina le prestazioni

Veeam fornisce la regolazione della frequenza delle richieste di eliminazione e la pianificazione del processo di eliminazione del backup. Per ottimizzare ulteriormente le prestazioni di eliminazione, è possibile disabilitare le eliminazioni sincrone e lasciare che lo scanner ILM gestisca l'eventuale eliminazione degli oggetti.

Passaggi per disabilitare le eliminazioni sincrone

1. Aprire StorageGRID Grid Manager.
2. Nell'angolo in alto a destra, seleziona il punto interrogativo e poi Documentazione API.
3. Nell'angolo in alto a destra, fare clic sul collegamento alla pagina Documentazione API privata.
4. Espandi ilm-advanced.
5. Selezionare OTTIENI ilm-advanced.
6. Selezionare Provalo, quindi Esegui.
7. Controllare il risultato della risposta.
    a. Se i valori sono nulli, significa che sono in uso i valori ilm-advanced predefiniti.
    b. Se i valori non sono nulli, significa che sono in uso valori avanzati ILM personalizzati. Copia tutto l'output dopo "data":, iniziando con { fino al penultimo }.
        i. Salvalo in un editor di testo.

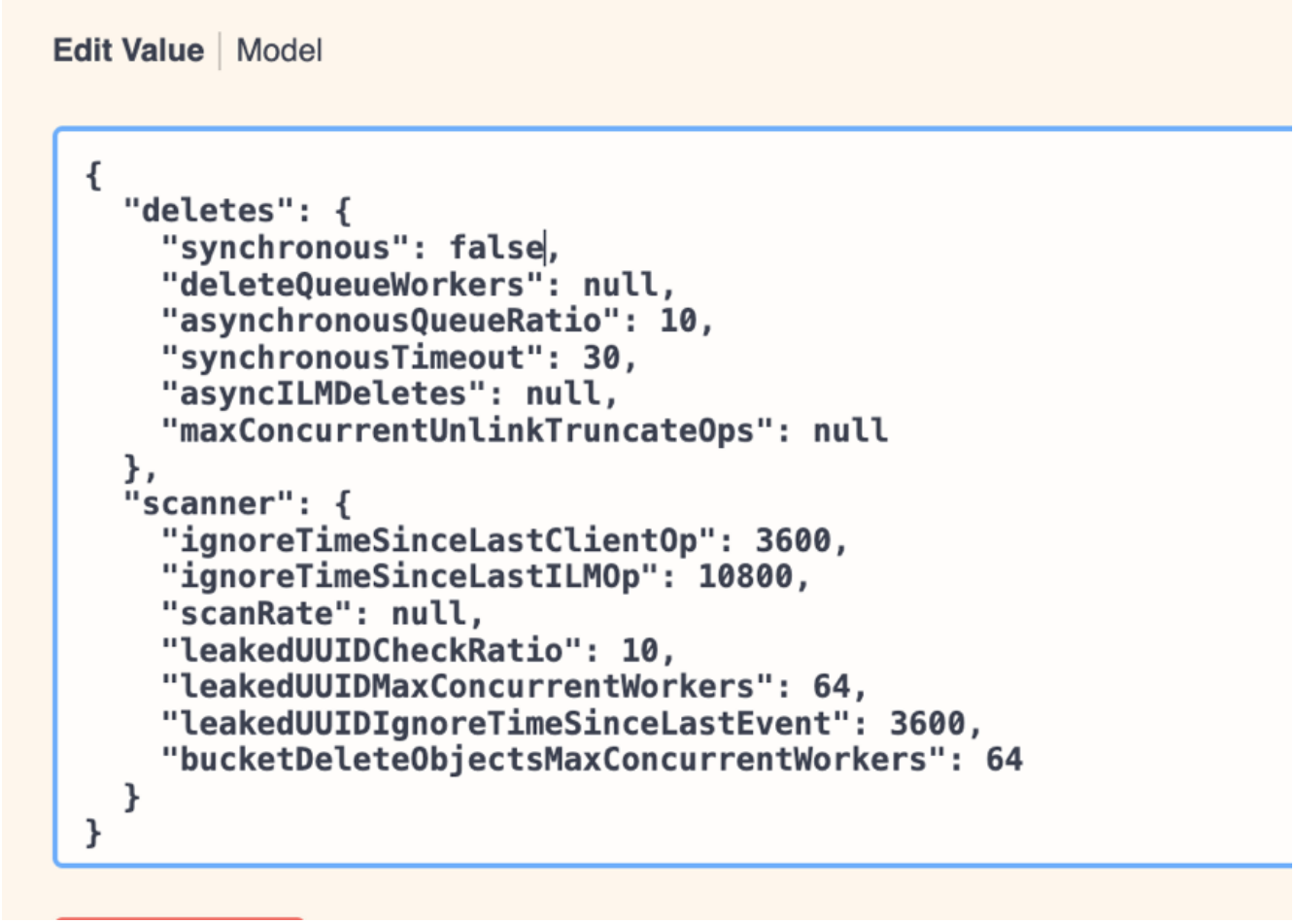
        Esempio di risposta:

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. Selezionare PUT ilm-advanced.
9. Seleziona Provalo per iniziare a modificare il corpo dell'API.
    a. Per impostazione predefinita, il corpo dell'API conterrà valori predefiniti e non valori personalizzati configurati in precedenza. Ecco perché è MOLTO importante eseguire i passaggi da 5 a 7.
10. Se nei passaggi da 5 a 7 vengono rilevati valori non predefiniti, sostituire il corpo dell'API con l'output salvato nel passaggio 7. . Altrimenti, se i valori erano nulli nei passaggi 5-7, lasciare il corpo dell'API così com'è.
11. Regola i seguenti parametri nella casella del corpo dell'API:
    a. Impostare il valore sincrono su falso.

       Esempio di testo del corpo dell'API:

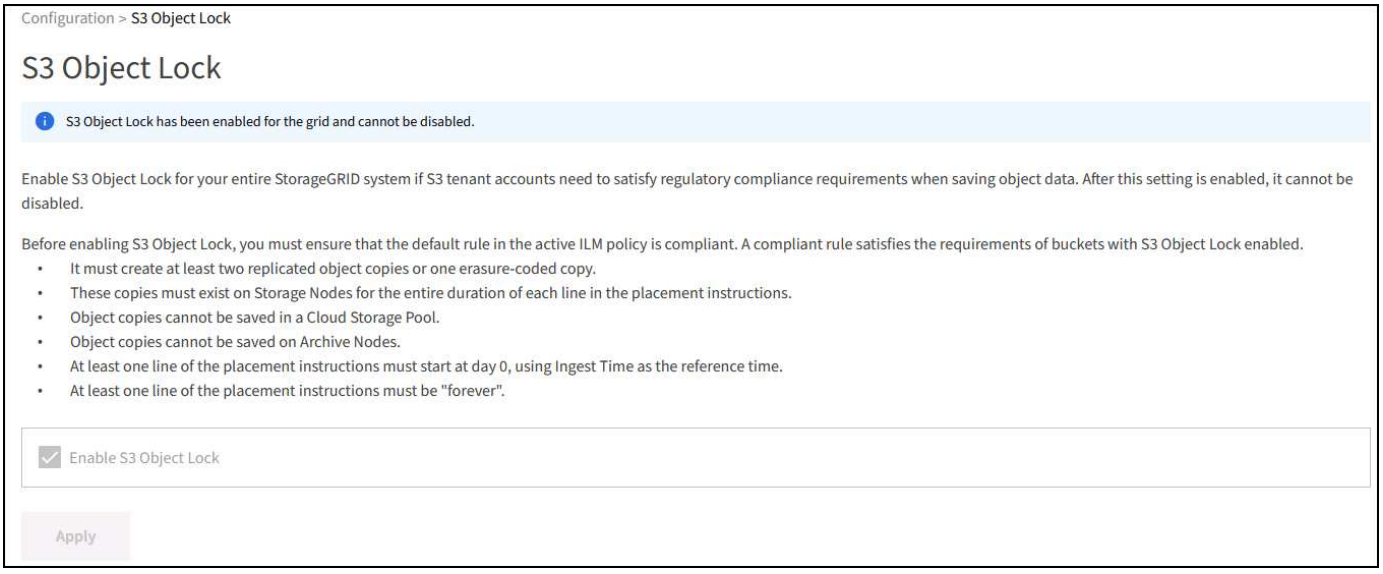


12. Una volta completato, seleziona Esegui

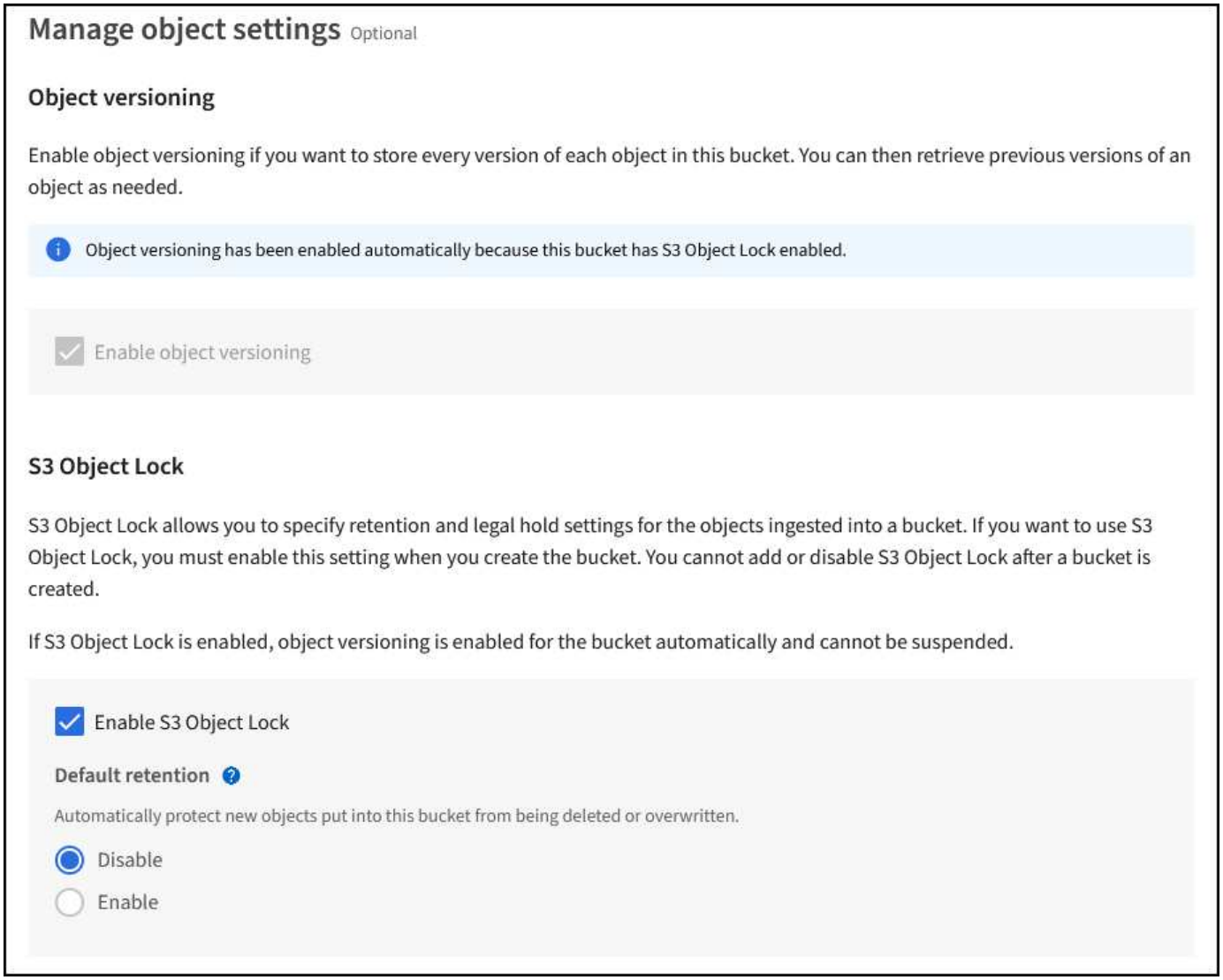
Punti chiave di implementazione

StorageGRID

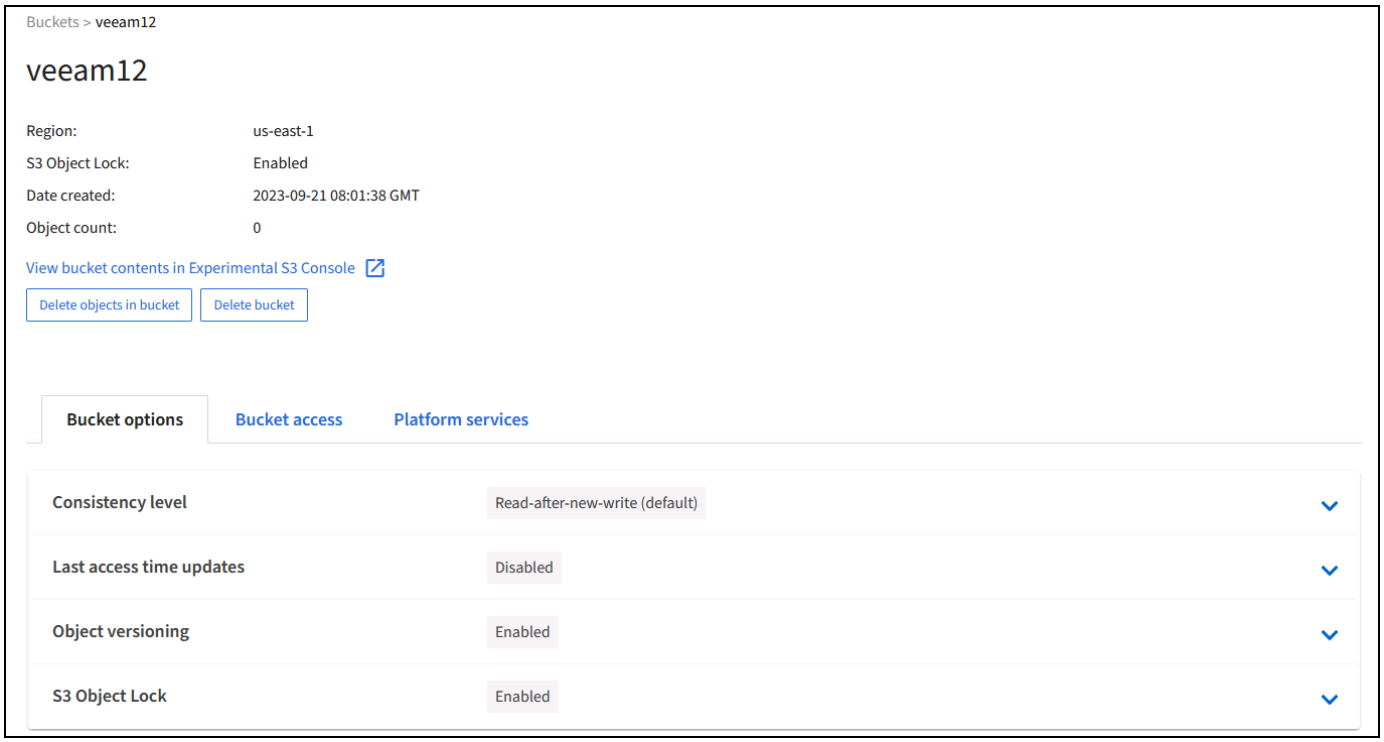
Assicurarsi che blocco oggetti sia attivato sul sistema StorageGRID se è necessaria l'immutabilità. Individuare l'opzione nell'interfaccia utente di gestione in Configurazione/blocco oggetti S3.



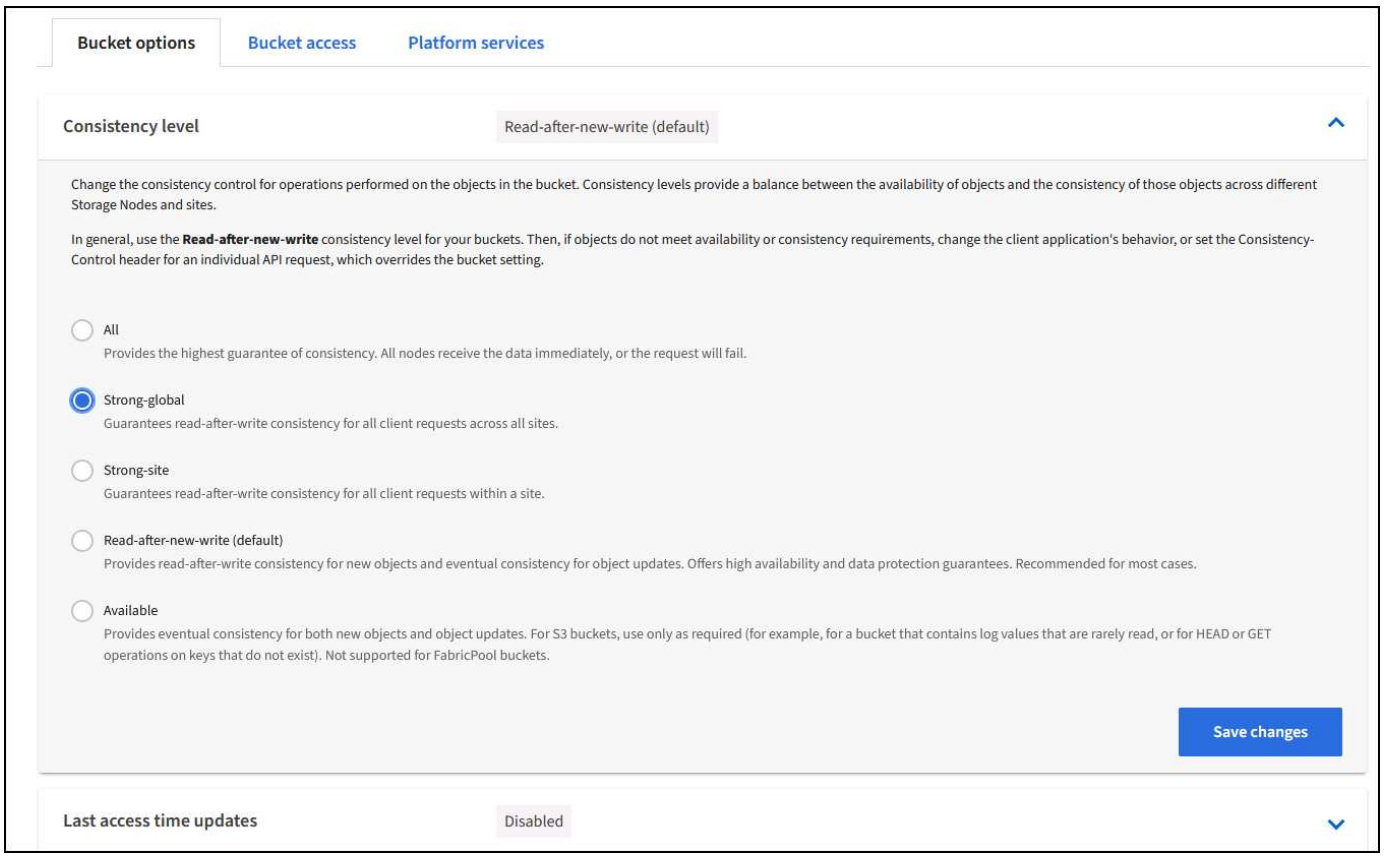
Quando si crea il bucket, selezionare "Enable S3 Object Lock" (attiva blocco oggetti 3D) se questo bucket deve essere utilizzato per i backup di immutabilità. In questo modo si attiva automaticamente la versione bucket. Lasciare disattivata la conservazione predefinita poiché Veeam imposterà esplicitamente la conservazione degli oggetti. Versioning e blocco oggetto S3 non devono essere selezionati se Veeam non sta creando backup immutabili.



Una volta creato il bucket, andare alla pagina dei dettagli del bucket creato. Selezionare il livello di coerenza.



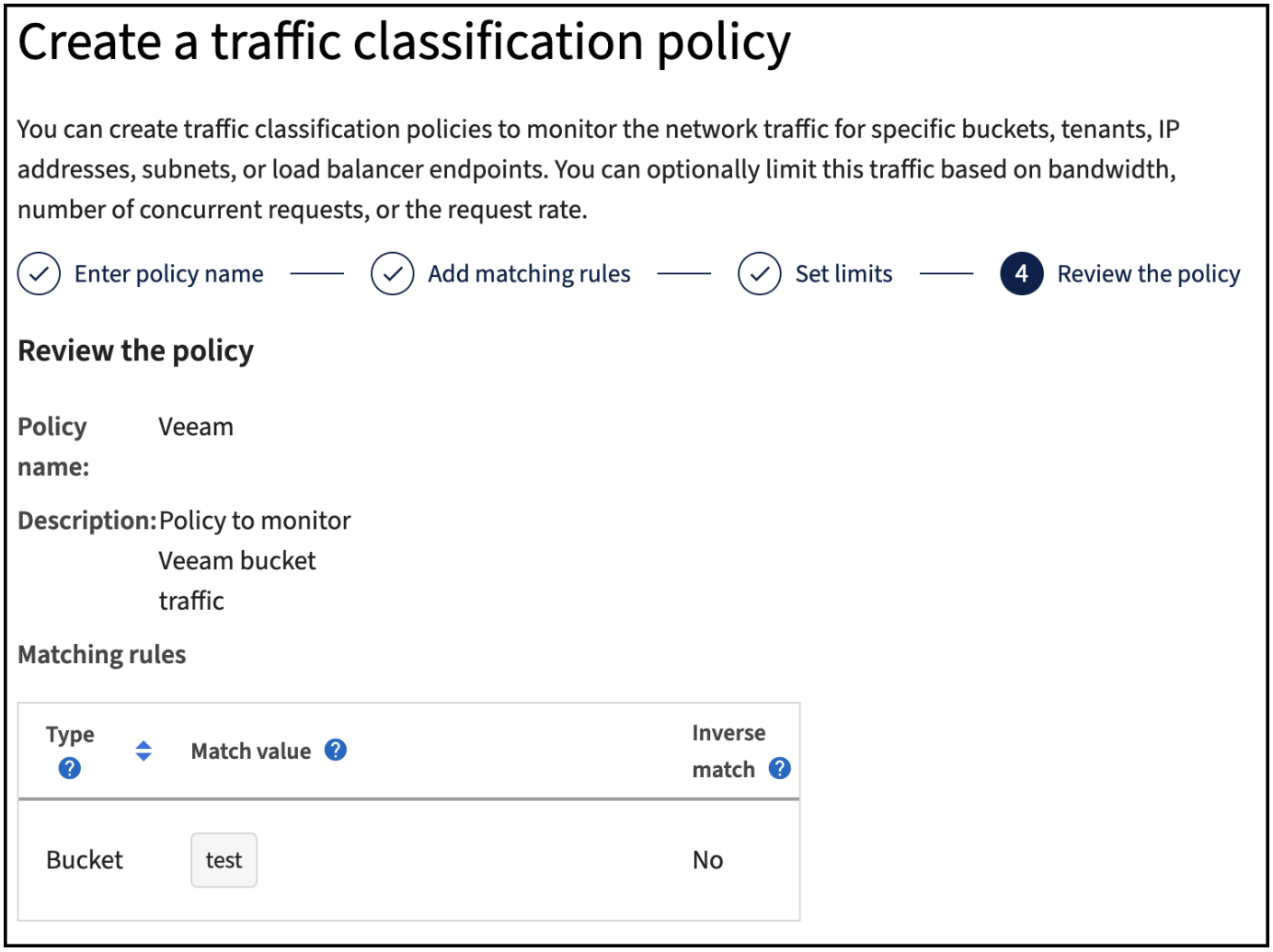
Veeam richiede una forte coerenza per i bucket S3. Quindi, per implementazioni multi-sito con Veeam che accede ai dati da posizioni multiple, seleziona "strong-Global". Se Veeam effettua backup e ripristini solo su un singolo sito, dovrebbe essere impostato su "strong-site". Salvare le modifiche.



StorageGRID fornisce un servizio di bilanciamento del carico integrato in ogni nodo amministrativo e nodo di gateway dedicato. Uno dei numerosi vantaggi dell'utilizzo di questo bilanciamento del carico è la possibilità di

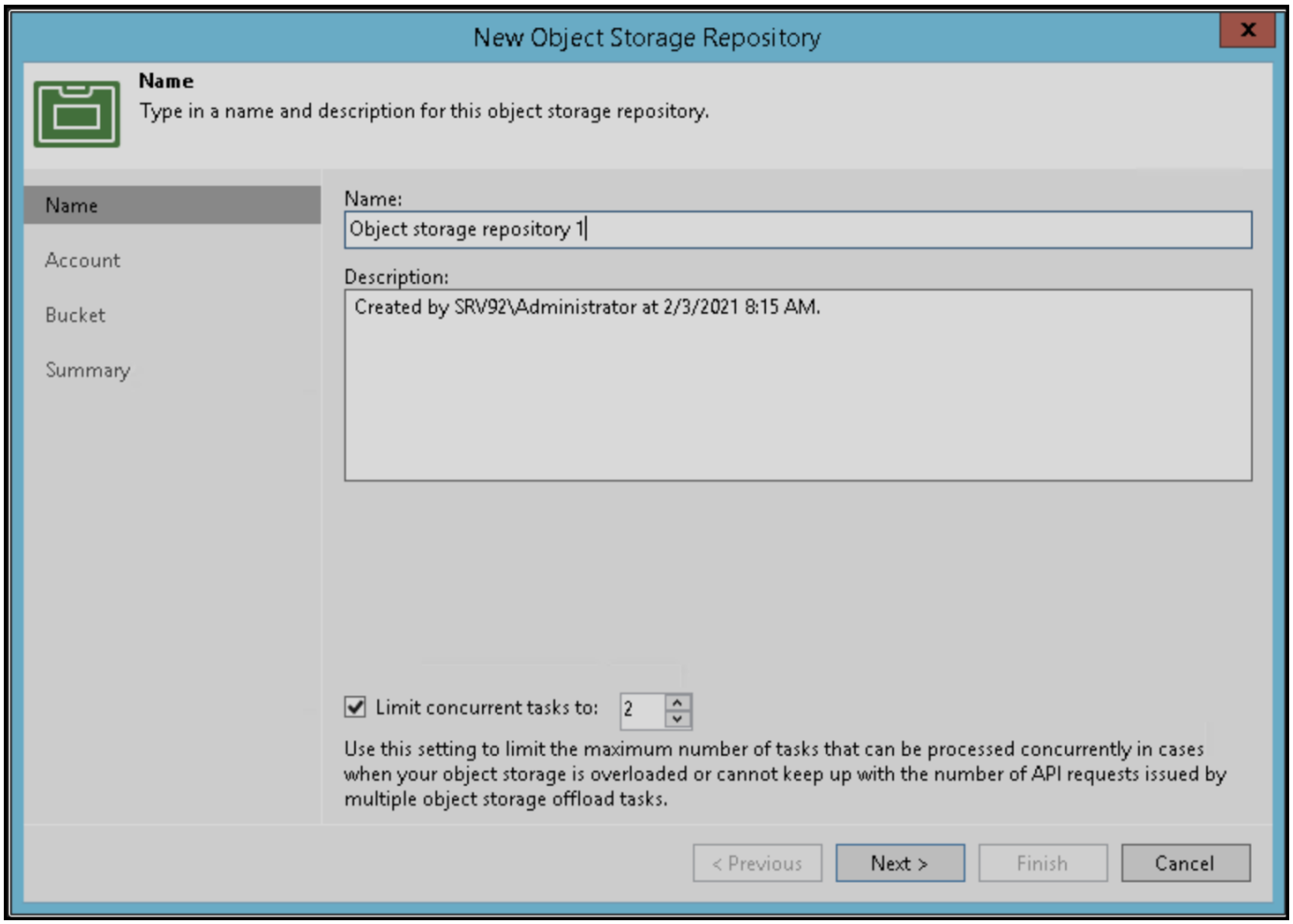
configurare i criteri di classificazione del traffico (QoS). Sebbene vengano utilizzati principalmente per limitare l'impatto di un'applicazione su altri carichi di lavoro dei client o per assegnare priorità a un carico di lavoro rispetto ad altri, forniscono anche un bonus di raccolta di metriche aggiuntive per agevolare il monitoraggio.

Nella scheda di configurazione, selezionare "Traffic Classification" (classificazione traffico) e creare una nuova policy. Assegnare un nome alla regola e selezionare il bucket o il tenant come tipo. Immettere i nomi dei bucket o locatario. Se la QoS è necessaria, impostare un limite, ma per la maggior parte delle implementazioni, è sufficiente aggiungere i vantaggi di monitoraggio che questo fornisce, quindi non impostare un limite.

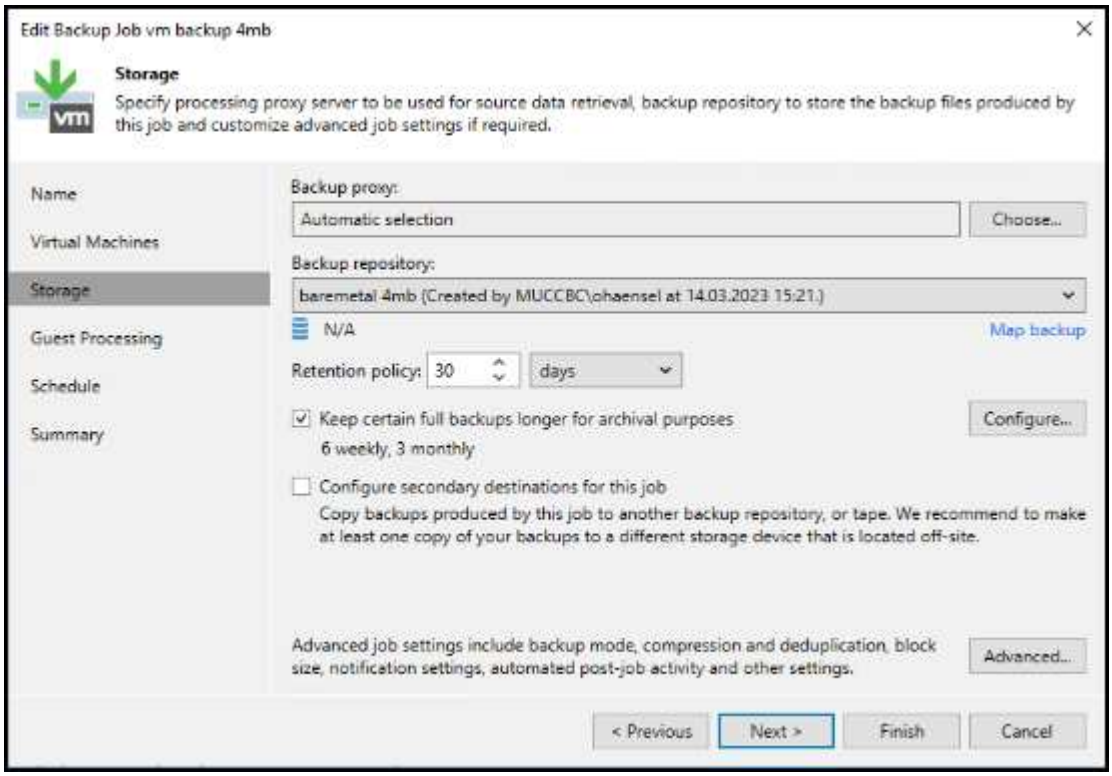


Veeam

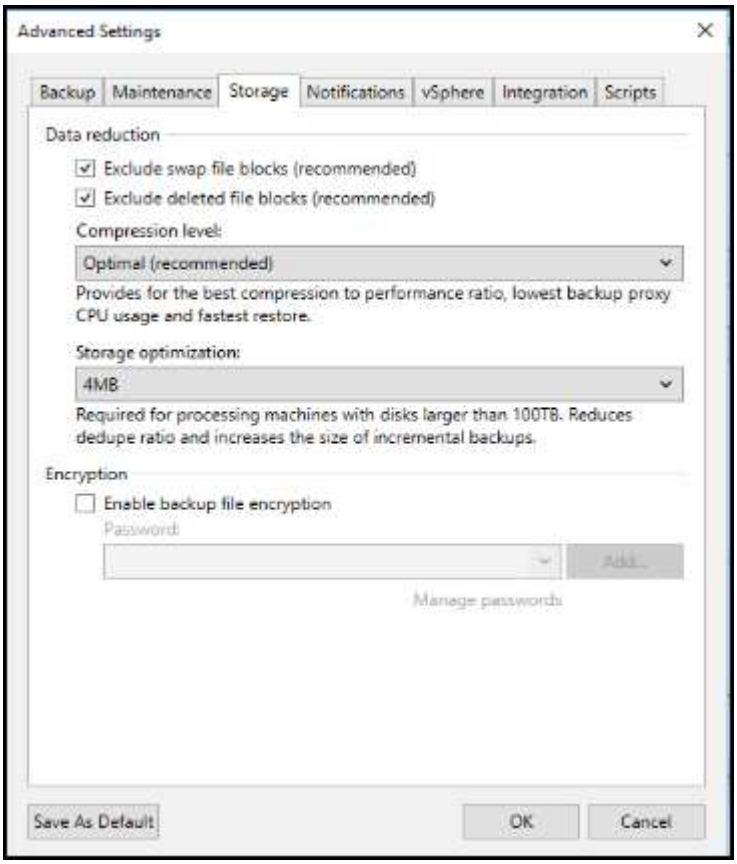
A seconda del modello e della quantità di appliance StorageGRID, potrebbe essere necessario selezionare e configurare un limite al numero di operazioni simultanee nel bucket.



Seguite la documentazione Veeam sulla configurazione del lavoro di backup nella console Veeam per avviare la procedura guidata. Dopo aver aggiunto le VM, selezionare il repository SOBR.



Fare clic su Impostazioni avanzate e modificare le impostazioni di ottimizzazione dell'archiviazione a 4 MB o più. Compressione e deduplica devono essere abilitate. Modificare le impostazioni guest in base ai requisiti e configurare la pianificazione del processo di backup.



Monitoraggio di StorageGRID

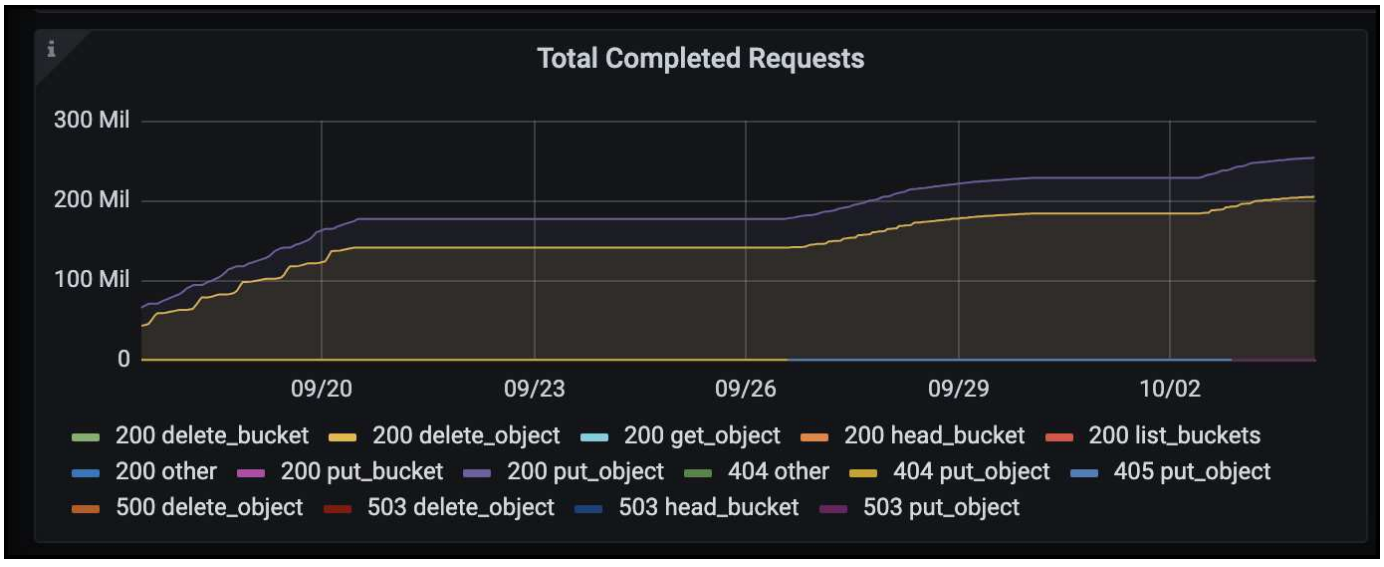
Per avere un quadro completo delle prestazioni congiunte di Veeam e StorageGRID, devi attendere la scadenza del tempo di conservazione dei primi backup. Fino a questo punto il workload Veeam è costituito principalmente da operazioni PUT e non si sono verificati eliminazioni. Una volta che i dati di backup stanno per scadere e le operazioni di pulizia sono in corso, è ora possibile vedere l'utilizzo completo e coerente nell'archivio oggetti e regolare le impostazioni in Veeam, se necessario.

StorageGRID fornisce utili grafici per monitorare il funzionamento del sistema nella pagina metriche della scheda supporto. I dashboard principali da esaminare saranno S3 Overview, ILM e Traffic Classification Policy, se è stato creato un criterio. Nel dashboard Panoramica di S3 sono disponibili informazioni su velocità operative, latenze e risposte delle richieste di S3.

Osservando le velocità S3 e le richieste attive è possibile visualizzare la quantità di carico gestita da ciascun nodo e il numero complessivo di richieste in base al tipo.

Il grafico durata media mostra il tempo medio impiegato da ciascun nodo per ciascun tipo di richiesta. Questa è la latenza media della richiesta e potrebbe essere un buon indicatore che potrebbe essere necessaria una regolazione aggiuntiva o che il sistema StorageGRID può assumere più carico.

Nel grafico Total Completed Requests (Richieste totali completate), è possibile visualizzare le richieste per tipo e codici di risposta. Se si visualizzano risposte diverse da 200 (OK) per le risposte, questo potrebbe indicare un problema come il sistema StorageGRID sta caricando pesantemente inviando 503 risposte (rallentando) e potrebbe essere necessario un ulteriore tuning, o è arrivato il momento di espandere il sistema per il carico aumentato.



Nel dashboard ILM è possibile monitorare le prestazioni di eliminazione del sistema StorageGRID. StorageGRID utilizza una combinazione di eliminazioni sincrone e asincrone su ciascun nodo per provare e

ottimizzare le performance complessive per tutte le richieste.

Con una Traffic Classification Policy, possiamo visualizzare le metriche sul carico bilanciatore richiesta throughput, tassi, durata, così come le dimensioni oggetto che Veeam sta inviando e ricevendo.

Dove trovare ulteriori informazioni

Per ulteriori informazioni sulle informazioni descritte in questo documento, consultare i seguenti documenti e/o siti Web:

- "Documentazione del prodotto NetApp StorageGRID"
- "Backup e replica di Veeam"

Configurare l'origine dati Dremio con StorageGRID

Di Angela Cheng

Dremio supporta una varietà di origini dati, incluso lo storage a oggetti on-premise o basato su cloud. È possibile configurare Dremio in modo che utilizzi StorageGRID come origine dati dello storage a oggetti.

Configurare l'origine dati Dremio

Prerequisiti

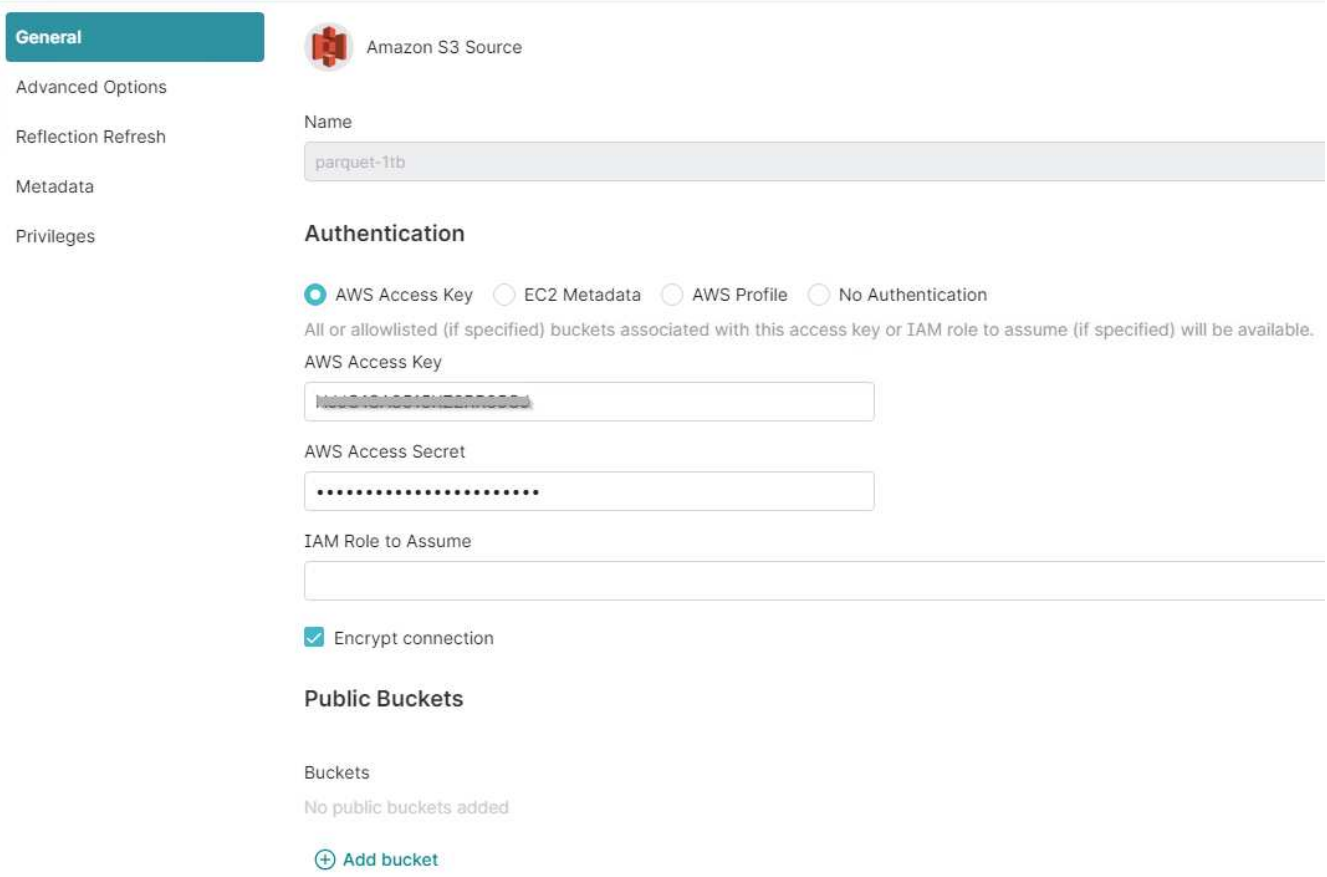
- Un URL dell'endpoint StorageGRID S3, un ID della chiave di accesso tenant S3 e una chiave di accesso segreta.
- Raccomandazione per la configurazione di StorageGRID: Disattivare la compressione (disattivata per impostazione predefinita).
  Dremio utilizza l'intervallo di byte GET per recuperare contemporaneamente diversi intervalli di byte dall'interno dello stesso oggetto durante la query. Le dimensioni tipiche per le richieste di intervalli di byte sono 1MB. L'oggetto compresso riduce le prestazioni di LETTURA DELL'intervallo di byte.

Guida di Dremio

"Connessione ad Amazon S3 - Configurazione dell'archiviazione compatibile con S3".

Istruzioni

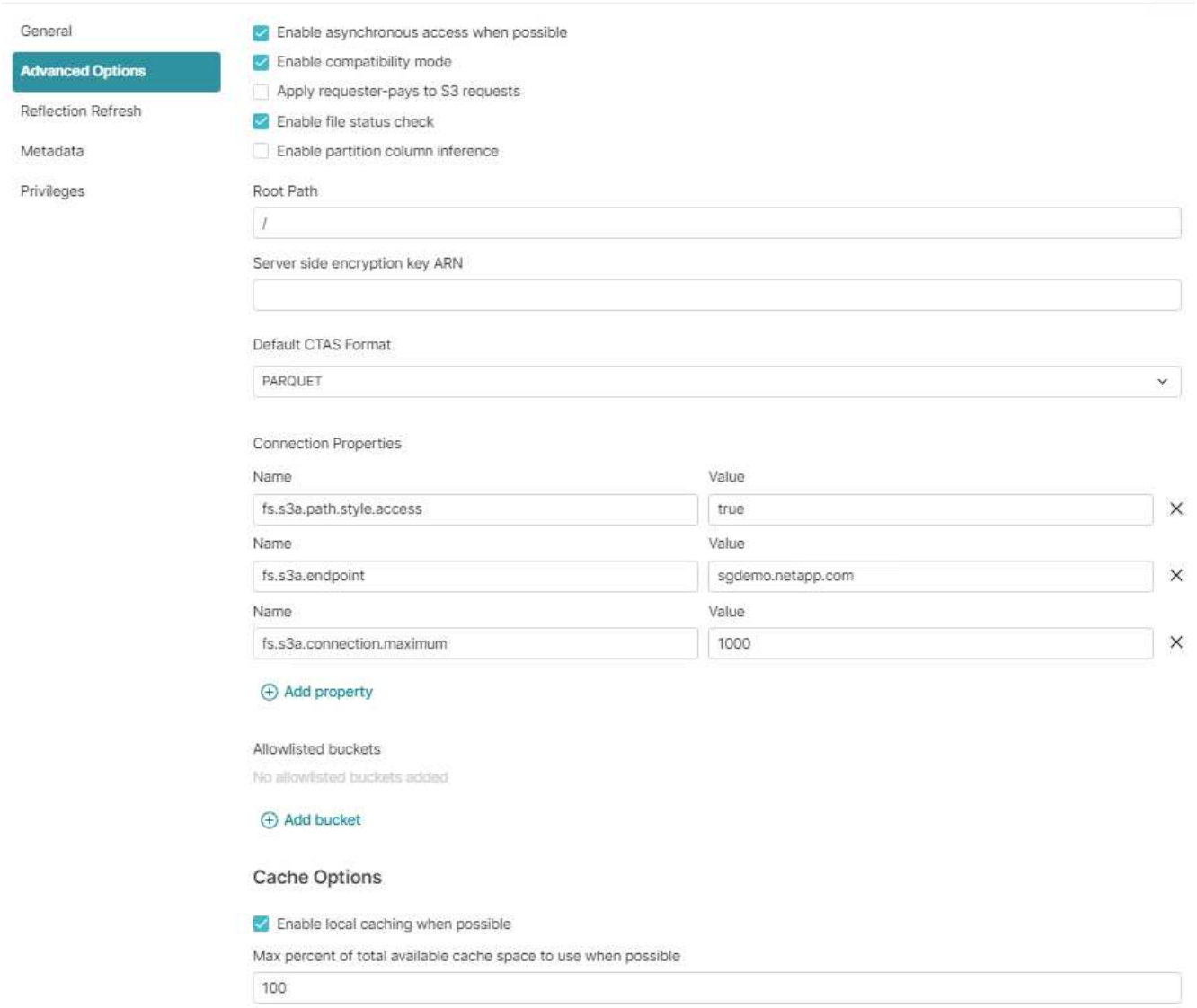
1. Nella pagina Datasets di Dremio, fare clic sul segno + per aggiungere un'origine, selezionare "Amazon S3".
2. Immettere un nome per la nuova origine dati, l'ID della chiave di accesso tenant StorageGRID S3 e la chiave di accesso segreta.
3. Selezionare la casella 'Crittografa connessione' se si utilizza https per la connessione all'endpoint StorageGRID S3.
   Se si utilizza un certificato CA autofirmato per questo endpoint S3, seguire la procedura della guida Dremio per aggiungere questo certificato CA a <JAVA_HOME>/jre/lib/Security + del server Dremio
   Esempio di screenshot



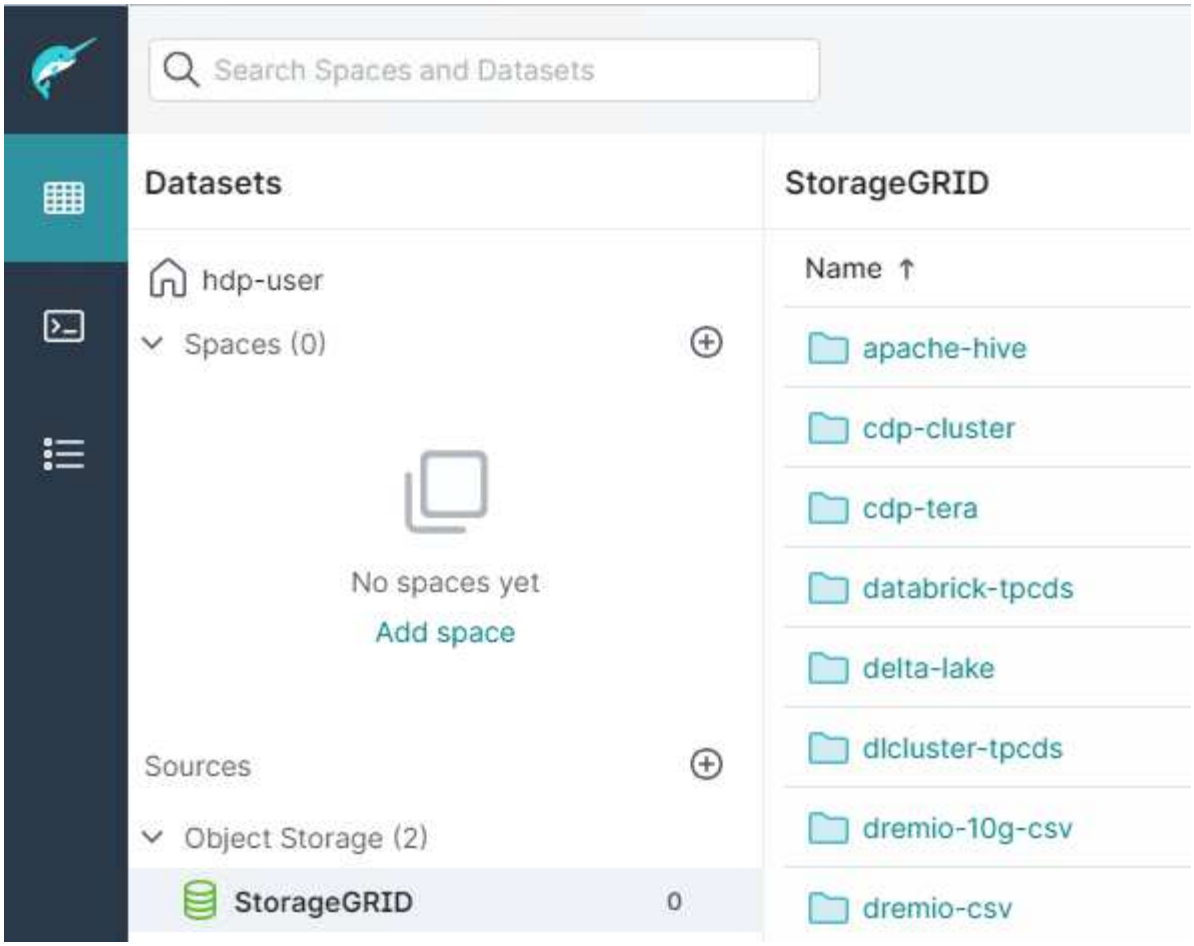
4. Fare clic su "Opzioni avanzate" e selezionare "attiva modalità di compatibilità"
5. In Proprietà di connessione, fare clic su + Aggiungi proprietà e aggiungere queste S3A proprietà.
6. fs.s3a.connection.il valore massimo predefinito è 100. Se i set di dati S3 includono file Parquet di grandi dimensioni con 100 o più colonne, è necessario immettere un valore maggiore di 100. Per questa impostazione, fare riferimento alla guida Dremio.

Nome	Valore
fs.s3a.endpoint	*<StorageGRID S3 endpoint:porta>*
fs.s3a.path.style.access	vero
fs.s3a.connection.maximum	*<un valore maggiore di 100>*

Esempio di screenshot



7. Configurare altre opzioni Dremio in base ai requisiti dell'organizzazione o delle applicazioni.
8. Fare clic sul pulsante Salva per creare questa nuova origine dati.
9. Una volta aggiunta correttamente l'origine dati StorageGRID, viene visualizzato un elenco di bucket sul pannello di sinistra.
   Esempio di screenshot



NetApp StorageGRID con GitLab

Di Angela Cheng

NetApp ha testato StorageGRID con GitLab. Vedere l'esempio di configurazione GitLab riportato di seguito. Fare riferimento a. "Guida alla configurazione dello storage a oggetti GitLab" per ulteriori informazioni.

Esempio di connessione allo storage a oggetti

Per le installazioni dei pacchetti Linux, questo è un esempio di `connection` impostazione nel modulo consolidato. Modifica `/etc/gitlab/gitlab.rb` e aggiungere le seguenti righe, sostituendo i valori desiderati:

```
# 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'
```

Informazioni sul copyright

Informazioni sul marchio commerciale