What are platform services?

StorageGRID 11.7

NetApp
January 09, 2024

Table of Contents

What are platform services?

StorageGRID platform services can help you implement a hybrid cloud strategy by allowing you to send event notifications and copies of S3 objects and object metadata to external destinations.

If the use of platform services is allowed for your tenant account, you can configure the following services for any S3 bucket:

- **CloudMirror replication**: Use StorageGRID CloudMirror replication service to mirror specific objects from a StorageGRID bucket to a specified external destination.

  For example, you might use CloudMirror replication to mirror specific customer records into Amazon S3 and then leverage AWS services to perform analytics on your data.

  > CloudMirror replication is not supported if the source bucket has S3 Object Lock enabled.

- **Notifications**: Use per-bucket event notifications to send notifications about specific actions performed on objects to a specified external Amazon Simple Notification Service™ (SNS).

  For example, you could configure alerts to be sent to administrators about each object added to a bucket, where the objects represent log files associated with a critical system event.

  > Although event notification can be configured on a bucket with S3 Object Lock enabled, the S3 Object Lock metadata (including Retain Until Date and Legal Hold status) of the objects will not be included in the notification messages.

- **Search integration service**: Use the search integration service to send S3 object metadata to a specified Elasticsearch index where the metadata can be searched or analyzed using the external service.

  For example, you could configure your buckets to send S3 object metadata to a remote Elasticsearch service. You could then use Elasticsearch to perform searches across buckets, and perform sophisticated analyses of patterns present in your object metadata.

  > Although Elasticsearch integration can be configured on a bucket with S3 Object Lock enabled, the S3 Object Lock metadata (including Retain Until Date and Legal Hold status) of the objects will not be included in the notification messages.

Because the target location for platform services is typically external to your StorageGRID deployment, platform services give you the power and flexibility that comes from using external storage resources, notification services, and search or analysis services for your data.

Any combination of platform services can be configured for a single S3 bucket. For example, you could configure both the CloudMirror service and notifications on a StorageGRID S3 bucket so that you can mirror specific objects to the Amazon Simple Storage Service, while sending a notification about each such object to a third party monitoring application to help you track your AWS expenses.

> The use of platform services must be enabled for each tenant account by a StorageGRID administrator using the Grid Manager or the Grid Management API.

How platform services are configured

Platform services communicate with external endpoints that you configure using the Tenant Manager or the Tenant Management API. Each endpoint represents an external destination, such as a StorageGRID S3 bucket, an Amazon Web Services bucket, a Simple Notification Service (SNS) topic, or an Elasticsearch cluster hosted locally, on AWS, or elsewhere.

After you create an external endpoint, you can enable a platform service for a bucket by adding XML configuration to the bucket. The XML configuration identifies the objects that the bucket should act on, the action that the bucket should take, and the endpoint that the bucket should use for the service.

You must add separate XML configurations for each platform service that you want to configure. For example:

- If you want all objects whose keys start with `/images` to be replicated to an Amazon S3 bucket, you must add a replication configuration to the source bucket.
- If you also want to send notifications when these objects are stored to the bucket, you must add a notifications configuration.
- Finally, if you want to index the metadata for these objects, you must add the metadata notification configuration that is used to implement search integration.

The format for the configuration XML is governed by the S3 REST APIs used to implement StorageGRID platform services:

Platform service	S3 REST API
CloudMirror replication	• GET Bucket replication • PUT Bucket replication
Notifications	• GET Bucket notification • PUT Bucket notification
Search integration	• GET Bucket metadata notification configuration • PUT Bucket metadata notification configuration These operations are custom to StorageGRID.

Related information

Considerations for platform services

Use S3 REST API

CloudMirror replication service

You can enable CloudMirror replication for an S3 bucket if you want StorageGRID to replicate specified objects added to the bucket to one or more destination buckets.

CloudMirror replication operates independently of the grid's active ILM policy. The CloudMirror service replicates objects as they are stored to the source bucket and delivers them to the destination bucket as soon as possible. Delivery of replicated objects is triggered when object ingest succeeds.

CloudMirror replication has important similarities and differences with the cross-grid replication feature. To learn more, see Compare cross-grid replication and CloudMirror replication.

If you enable CloudMirror replication for an existing bucket, only the new objects added to that bucket are replicated. Any existing objects in the bucket aren't replicated. To force the replication of existing objects, you can update the existing object's metadata by performing an object copy.

If you are using CloudMirror replication to copy objects to an Amazon S3 destination, be aware that Amazon S3 limits the size of user-defined metadata within each PUT request header to 2 KB. If an object has user-defined metadata greater than 2 KB, that object will not be replicated.

In StorageGRID, you can replicate the objects in a single bucket to multiple destination buckets. To do so, specify the destination for each rule in the replication configuration XML. You can't replicate an object to more than one bucket at the same time.

Additionally, you can configure CloudMirror replication on versioned or unversioned buckets, and you can specify a versioned or unversioned bucket as the destination. You can use any combination of versioned and unversioned buckets. For example, you could specify a versioned bucket as the destination for an unversioned source bucket, or vice versa. You can also replicate between unversioned buckets.

Deletion behavior for the CloudMirror replication service is the same as the deletion behavior of the Cross Region Replication (CRR) service provided by Amazon S3 — deleting an object in a source bucket never deletes a replicated object in the destination. If both source and destination buckets are versioned, the delete marker is replicated. If the destination bucket is not versioned, deleting an object in the source bucket does not replicate the delete marker to the destination bucket or delete the destination object.

As objects are replicated to the destination bucket, StorageGRID marks them as "replicas." A destination StorageGRID bucket will not replicate objects marked as replicas again, protecting you from accidental replication loops. This replica marking is internal to StorageGRID and does not prevent you from leveraging AWS CRR when using an Amazon S3 bucket as the destination.

The custom header used to mark a replica is `x-ntap-sg-replica`. This marking prevents a cascading mirror. StorageGRID does support a bidirectional CloudMirror between two grids.

The uniqueness and ordering of events in the destination bucket aren't guaranteed. More than one identical copy of a source object might be delivered to the destination as a result of operations taken to guarantee delivery success. In rare cases, when the same object is updated simultaneously from two or more different StorageGRID sites, the ordering of operations on the destination bucket might not match the ordering of events on the source bucket.

CloudMirror replication is typically configured to use an external S3 bucket as a destination. However, you can also configure replication to use another StorageGRID deployment or any S3-compatible service.

Understand notifications for buckets

You can enable event notification for an S3 bucket if you want StorageGRID to send notifications about specified events to a destination Amazon Simple Notification Service (SNS).

You can configure event notifications by associating notification configuration XML with a source bucket. The notification configuration XML follows S3 conventions for configuring bucket notifications, with the destination SNS topic specified as the URN of an endpoint.

Event notifications are created at the source bucket as specified in the notification configuration and are delivered to the destination. If an event associated with an object succeeds, a notification about that event is created and queued for delivery.

The uniqueness and ordering of notifications aren't guaranteed. More than one notification of an event might be delivered to the destination as a result of operations taken to guarantee delivery success. And because delivery is asynchronous, the time ordering of notifications at the destination is not guaranteed to match the ordering of events on the source bucket, particularly for operations that originate from different StorageGRID sites. You can use the `sequencer` key in the event message to determine the order of events for a particular object, as described in Amazon S3 documentation.

Supported notifications and messages

StorageGRID event notifications follow the Amazon S3 API with some limitations:

- The following event types are supported:
  - s3:ObjectCreated:*
  - s3:ObjectCreated:Put
  - s3:ObjectCreated:Post
  - s3:ObjectCreated:Copy
  - s3:ObjectCreated:CompleteMultipartUpload
  - s3:ObjectRemoved:*
  - s3:ObjectRemoved:Delete
  - s3:ObjectRemoved:DeleteMarkerCreated
  - s3:ObjectRestore:Post
- Event notifications sent from StorageGRID use the standard JSON format but don't include some keys and use specific values for others, as shown in the table:

Key name	StorageGRID value
eventSource	`sgws:s3`
awsRegion	not included
x-amz-id-2	not included
arn	`urn:sgws:s3:::bucket_name`

Understand search integration service

You can enable search integration for an S3 bucket if you want to use an external search and data analysis service for your object metadata.

The search integration service is a custom StorageGRID service that automatically and asynchronously sends S3 object metadata to a destination endpoint whenever an object or its metadata is updated. You can then use sophisticated search, data analysis, visualization, or machine learning tools provided by the destination service

to search, analyze, and gain insights from your object data.

You can enable the search integration service for any versioned or unversioned bucket. Search integration is configured by associating metadata notification configuration XML with the bucket that specifies which objects to act on and the destination for the object metadata.

Notifications are generated in the form of a JSON document named with the bucket name, object name, and version ID, if any. Each metadata notification contains a standard set of system metadata for the object in addition to all of the object's tags and user metadata.

> For tags and user metadata, StorageGRID passes dates and numbers to Elasticsearch as strings or as S3 event notifications. To configure Elasticsearch to interpret these strings as dates or numbers, follow the Elasticsearch instructions for dynamic field mapping and for mapping date formats. You must enable the dynamic field mappings on the index before you configure the search integration service. After a document is indexed, you can't edit the document's field types in the index.

Notifications are generated and queued for delivery whenever:

- An object is created.
- An object is deleted, including when objects are deleted as a result of the operation of the grid's ILM policy.
- Object metadata or tags are added, updated, or deleted. The complete set of metadata and tags is always sent on update — not just the changed values.

After you add metadata notification configuration XML to a bucket, notifications are sent for any new objects that you create and for any objects that you modify by updating its data, user metadata, or tags. However, notifications aren't sent for any objects that were already in the bucket. To ensure that object metadata for all objects in the bucket is sent to the destination, you should do either of the following:

- Configure the search integration service immediately after creating the bucket and before adding any objects.
- Perform an action on all objects already in the bucket that will trigger a metadata notification message to be sent to the destination.

The StorageGRID search integration service supports an Elasticsearch cluster as a destination. As with the other platform services, the destination is specified in the endpoint whose URN is used in the configuration XML for the service. Use the NetApp Interoperability Matrix Tool to determine the supported versions of Elasticsearch.

Related information

Configuration XML for search integration

Object metadata included in metadata notifications

JSON generated by search integration service

Configure search integration service

Copyright information

Trademark information