

Ceph Object Gateway - S3 Bucket Notifications

Versioning

Version	Description	Date
0.1	First draft	2015-08-18
0.2	RPUSH instead of LPUSH, some minor changes	2015-08-18
0.3	LPUSH instead of RPUSH	2015-08-24
0.4	Support S3 PUT object copy, specify behaviour for logging on error	2015-08-28
1.0	Final version	2015-11-09

Goal

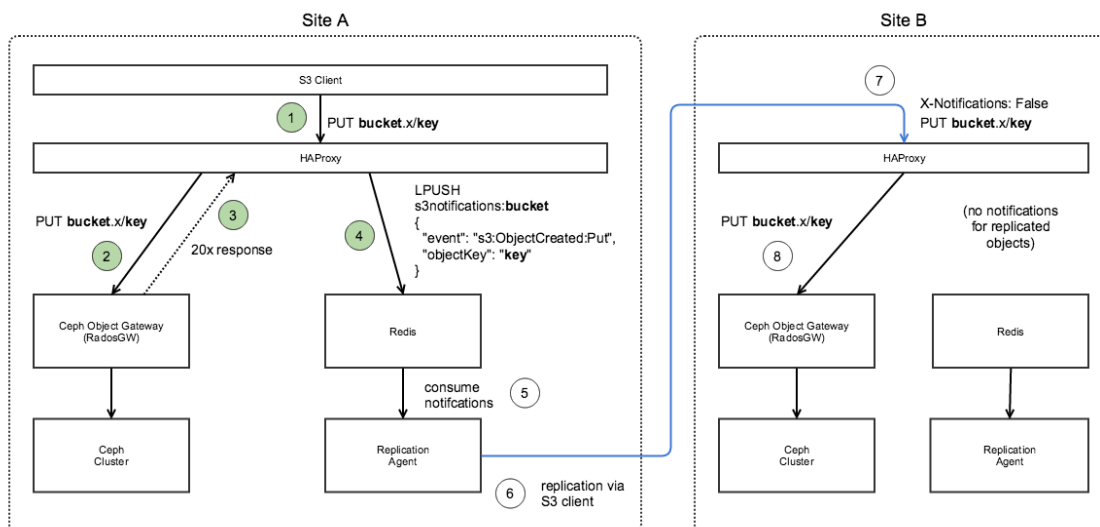
Provide a simple way of being informed about newly created or deleted objects for a given set of S3 buckets, e.g. to trigger replication of buckets with millions of objects.

Realization

Description:

- add a transparent proxy in front of Ceph Object Gateway
 - extend HAProxy
 - intercept every PUT, POST and DELETE request
 - determine the event, bucket name and object key
 - publish a message (notification)
 - use [hiredis](#) C client for Redis with HAProxy
 - LPUSH the notification (specified below) if the RadosGW call was successful
 - only if notifications are enabled via configuration
 - only if value of X-Notifications header is not False
- message queue to publish the event (object created or deleted)
 - use Redis (with hiredis C client)

Architecture



1. Client makes a PUT to create object with key "key" in bucket "bucket"
2. HAProxy routes the call to this local RadosGW
3. On success (20x response), prepare sending a notification
4. Publish the notification to Redis asynchronously (try once, log the whole notification on error)
(...)
5. Replication agent consumes the notification
6. Replication agent creates the object with key "key" in bucket "bucket" on remote site
7. Set "X-Notifications: False" HTTP header to signal the HAProxy to not send notifications for a replicated object
8. HAProxy routes the call to his local RadosGW

HAProxy Configuration

The extended version of HAProxy must provide a way to configure:

- the name of the buckets to enable notifications for (default: nothing set = no notifications)
- a switch to globally enable or disable notifications (default: enabled)
- a prefix of the Redis queue (default: "s3notifications")
- Redis endpoint for notifications
- a switch to enable logging of the whole message if writing to Redis queue fails

The extended HAProxy must not send notifications if a HTTP header "X-Notifications: False" is present.

Message Specification

Queue name

s3notifications:<bucketName>

Example:

s3notifications:images-eu

Message format

The message is in JSON format.

Field name	Description	Data type	Possible values
event	type	String (fixed set)	s3:ObjectCreated:Put s3:ObjectCreated:Post s3:ObjectCreated:Copy s3:ObjectRemoved:Delete
objectKey	key of created or deleted object	String	(see S3 documentation for possible values)
source <i>(optional)</i>	only for PUT operations; value of x-amz-copy-source header (if set) (see http://docs.aws.amazon.com/AmazonS3/latest/API/RESTObjectCOPY.html)	String	/<bucketName>/<objectKey>

Example message:

(object with key "foobar" was created via PUT operation)

```
{
```

```
"event": "s3:ObjectCreated:Put",  
"objectKey": "foobar"  
}
```

Example message (with source):

(object with key "foobar" copied from source /sourceBucketName/baz)

```
{  
  "event": "s3:ObjectCreated:Put",  
  "objectKey": "foobar",  
  "source": "/sourceBucketName/baz"  
}
```

Additional notes

- The source code of the extended HAProxy is publicly available at <https://github.com/spreadshirt/s3gw-haproxy>