

# AMAZON SIMPLE STORAGE SERVICE REPLICATION

P SRI VAISHNAVI<sup>1</sup>

\*\*\*

**Abstract** - Customers often have operational requirements or company policies that require additional copies of their existing or new Amazon S3 items for Disaster Recovery or to reduce latency. Amazon S3 Replication is used extensively to replicate newly uploaded objects between S3 buckets. With the help of Amazon S3 Batch Replication, we can replicate existing objects which were created or uploaded before enabling S3 replication. Whereas, Amazon S3 Same Region Replication and Amazon S3 Cross Region Replication will automatically replicate newly created objects in the bucket.

**Key Words:** S3 Replication, S3 Same Region Replication (SRR), S3 Cross Region Replication(CRR),S3 Batch Replication, S3 Replication Time Control(RTC).

## 1. INTRODUCTION

Replication is a process of automatically copying objects between buckets in the same or different AWS Regions. This copying happens when you create new objects or update existing ones.

Amazon Simple Storage Service (S3) replication allows you to have an exact copy of your objects stored in other buckets. This is a very flexible and cost-effective way to store your data in the cloud. It can be used to help you share your data across different accounts, or to store large amounts of data where you need it to be available anywhere.

### 1.1 PURPOSE OF S3 REPLICATION

- Reduce Latency
- Enhance availability
- Disaster Recovery
- Object Replication within 15 minutes
- Store objects in Multiple Region
- Copy Objects to cost effective storage class
- Data redundancy
- Meet compliance requirements
- Copying Objects while preserving the metadata
- Sustain object copies under various Owners

## 2. REPLICATION OPTIONS:

### 2.1 AMAZON S3 SAME REGION REPLICATION

S3 Same Region Replication(SRR) will automatically replicate objects from a source bucket to destination bucket within same AZ or different AZ in same region. We can easily aggregate the logs from different S3 buckets for intra-region processing or for configuration of live replication between developer and test environment. It uses asynchronous replication, that means objects are not copied to destination bucket as soon as it is created or modified.



**Fig -1:** S3 Same Region Replication(SRR)

We can configure SRR using AWS Console, AWS CLI and AWS SDKs.

Replication can be done at bucket-level, for certain prefix, for specific tags or object level within same account or different account in same region.

It allow us to change account ownership to protect replicated objects from accidental deletion.

### 2.2 AMAZON S3 CROSS REGION REPLICATION

S3 Cross Region Replication(CRR) will automatically replicate objects or data from source bucket to destination bucket in different region. It minimizes latency for data access in different geographic regions. Similar to SRR, we can configure CRR using AWS Console, AWS CLI and AWS SDKs.

Replication can be done at bucket-level, for certain prefix, for specific tags or object level within same account or different account in different region.

It allow us to change account ownership to protect replicated objects from accidental deletion.

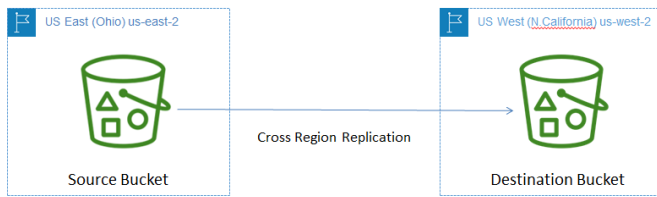


Fig -2: S3 Cross Region Replication (CRR)

We can configure S3 CRR to copy objects into one or more buckets in a different region to improve resilience. It helps us in maintaining compliance, disaster recovery and security enforcement.

### 2.3 S3 BATCH REPLICATION:

With CRR and SRR, Amazon S3 protects your data by automatically replicating new objects that you upload. S3 Batch Replication, on the other hand, lets you replicate existing objects using S3 Batch Operations, which are managed jobs.

We need to pay S3 charges for storing data in the selected storage class in the primary copy, for PUT requests to replicate data across AWS Regions, and for retrieving data from storage in the selected storage class. If something already exists in the destination bucket, we will check to see if the data is in sync. If it's not in sync, we have to replicate the data using a PUT request, but you won't pay the inter-Region Data Transfer OUT charge. If the data is in sync, Batch Replication won't do anything and we will incur no charge.

If we want to do a batch replication job, we can choose to have it only replicate objects that were created recently or that are in a certain status.

- Pending: The object is pending replication.
- In Progress: The object is being replicated.
- Available: The object is available for replication.
- Error: An error occurred while replicating the object.

### 3. STEPS TO SETUP S3 REPLICATION:

1. Navigate to S3 console.
2. Create two buckets by enabling bucket versioning -Source bucket in Region A and Destination bucket in Region A(S3 SRR) or in Region B (S3 CRR).

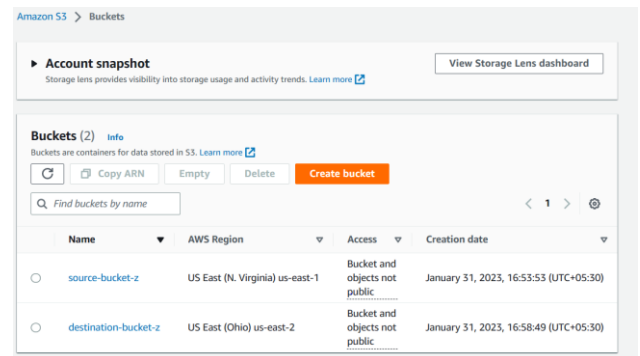


Fig -3: S3 Cross Region Replication(CRR)

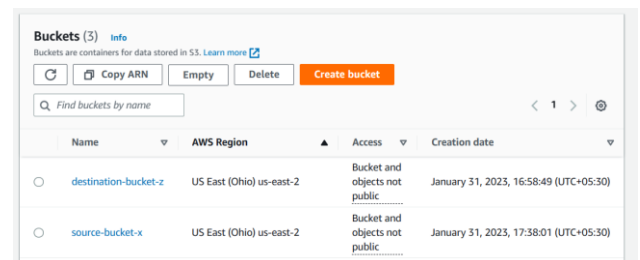


Fig -4: S3 Same Region Replication(SRR)

3. In source bucket, in Management tab, Click on create replication rule.

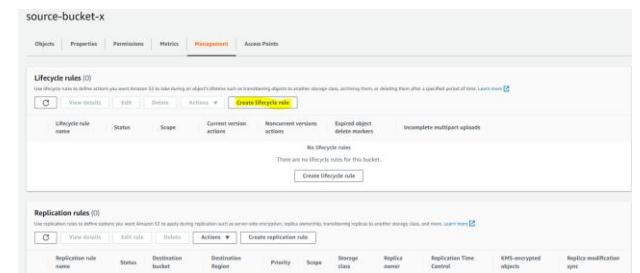
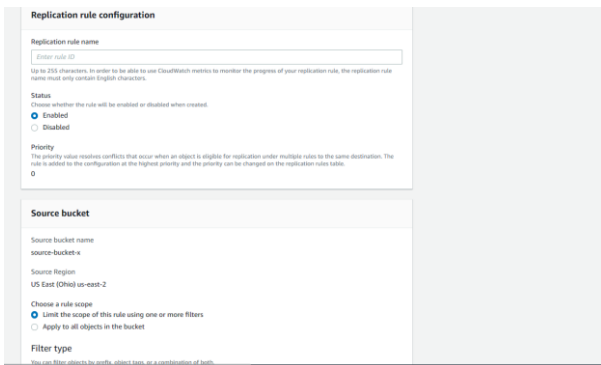


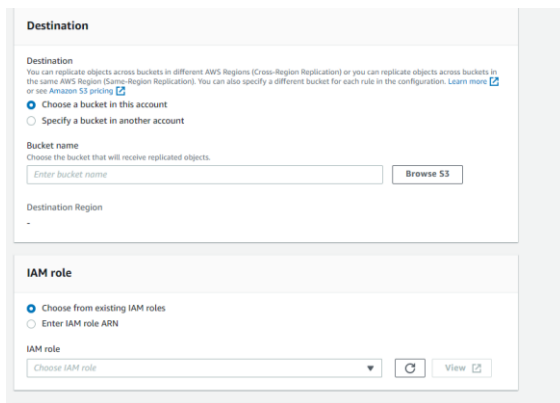
Fig -5: Replication rule creation

4. In replication rule configuration, enter Replication rule name, choose Status of replication rule as Enable/Disable after creating bucket, set Priority value. In Source bucket, choose scope rule as Limit the scope of this rule using one or more filters (if we want to limit to specific tag or prefix or object) with Filter type as specific prefix or **Apply to all objects in the bucket** (if we want to all objects to be replicated).



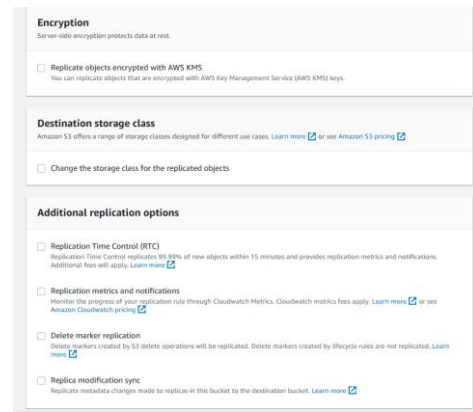
**Fig -6:** Replication rule configuration

5. In Destination, Select Choose a bucket in this account if destination bucket exist in same account or Specify a bucket in another account if destination bucket is present in other account and also enable check box if you want to change ownership to destination bucket owner.



**Fig -7:** Destination and IAM role configuration

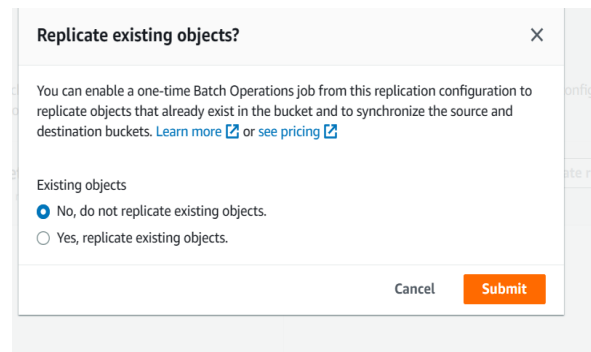
6. For IAM role, if you already created any role, then select Choose from existing IAM roles or select Enter IAM role and choose Create new role.
7. To replicate encrypted, we can enable Replicate objects encrypted with AWS KMS and provide KMS Key details. To change storage class at destination, enable Change the storage class for the replicated objects and choose the storage class.



**Fig -8:** Additional replication option

8. We can also enable Additional replication options like Replication Time Control(RTC) to replicate new objects within 15 minutes, Replication metrics and notifications to get metrics to monitor replication rule, Delete marker replication (delete markers are replicated on destination bucket object) and Replica modification sync to replicate changes made in source bucket to destination bucket.

9. After we click on Save. We will get two options as shown as below figure.



**Fig -9:** Replicate existing object

- a. We can NO, if we don't want to replicate existing objects in source bucket and Click on Submit.
- b. Select Yes, if you want to replicate existing objects and Submit. It will redirect to Batch Operation Job Creation page.

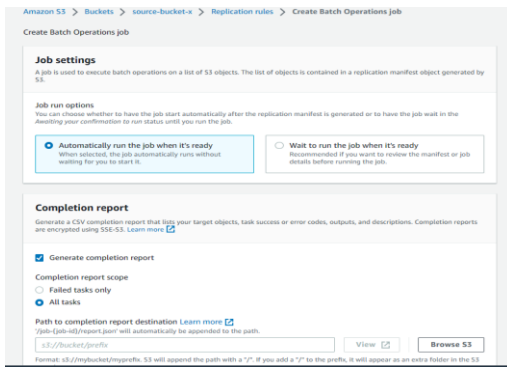


Fig -10: Batch Job Configuration

- c. Choose Job run options.
- d. If we want to completion report for all tasks or for failed tasks, Enable Generate completion report and provide path to import report.

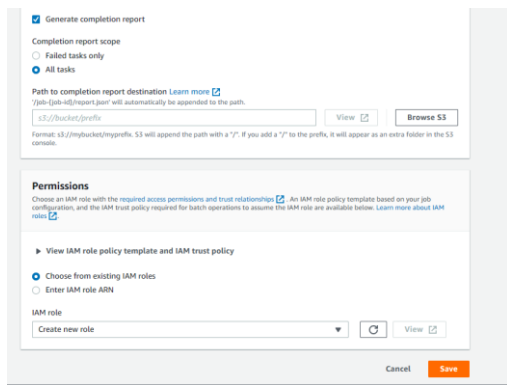


Fig -11: Report Generation and iam Role Configuration

- e. We can select existing permission for batch operations or create new role and click on Save.

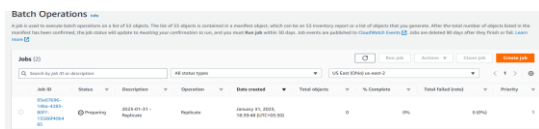


Fig -12: Batch Operation

- f. Batch operation will be created. After batch operation is completed, a folder will be created for completion report in provided destination bucket.

## 4. RESULTS

### a. Replication

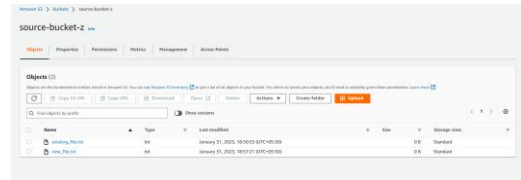


Fig -13: Source Bucket

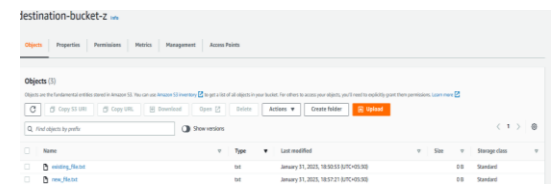


Fig -14: Destination Bucket after Replication

### b. Batch Replication

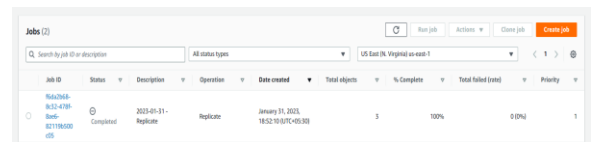


Fig -15: Batch Job

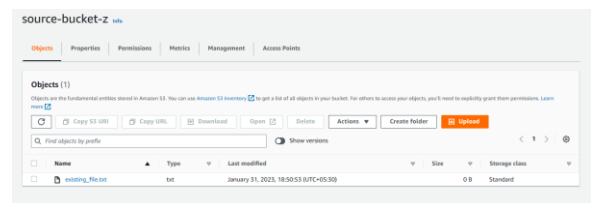


Fig -16: Source Bucket before Replication

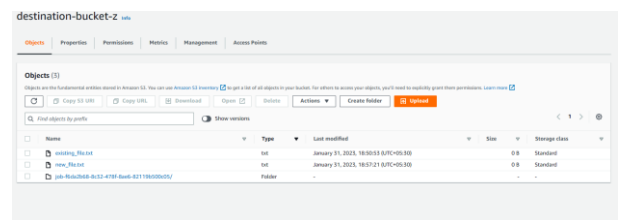


Fig -17: Destination Bucket after Batch Replication

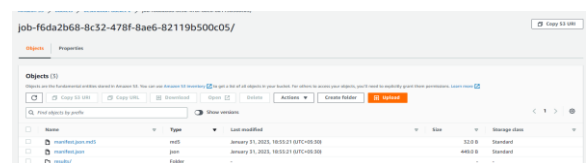


Fig -18: Completion Report

### 3. CONCLUSION

The main objective of this S3 Replication is to copy objects or data to same region (same account or different region) for intra-testing, log aggregation or to different region (same account or different region) for disaster recovery, compliance requirement. We have implemented S3 batch replication to replicate for existing object which was created before enabling replication as S3 SRR and S3 CRR will only replicate newly uploaded or created object after creating replication rule.

### REFERENCES

- [1] <https://docs.aws.amazon.com/AmazonS3/latest/userguide/replication.html>
- [2] <https://docs.aws.amazon.com/AmazonS3/latest/userguide/replication.html>
- [3] <https://www.geeksforgeeks.org/amazon-s3-cross-region-replication/>
- [4] [https://medium.com/@maxy\\_ermayank/aws-s3-bucket-replication-56dcfec2f424](https://medium.com/@maxy_ermayank/aws-s3-bucket-replication-56dcfec2f424)
- [5] <https://aws.plainenglish.io/what-is-amazon-s3-replication-fbb9c45507b3>
- [6] <https://medium.com/cloud-techies/s3-same-region-replication-srr-and-cross-region-replication-crr-34d446806bab>
- [7] <https://aws.amazon.com/s3/pricing/>
- [8] <https://aws.plainenglish.io/aws-s3-replication-60c4663efc2f>
- [9] <https://www.workfall.com/learning/blog/replicate-existing-s3-objects-using-s3-batch-replication/>
- [10] <https://docs.aws.amazon.com/AmazonS3/latest/userguide/s3-batch-replication-batch.html>
- [11] <https://aws.amazon.com/blogs/aws/new-cross-region-replication-for-amazon-s3/>
- [12] <https://aws.amazon.com/blogs/storage/replicating-existing-objects-between-s3-buckets/>
- [13] <https://www.workfall.com/learning/blog/replicate-existing-s3-objects-using-s3-batch-replication/>