

# S3 Batch Operation

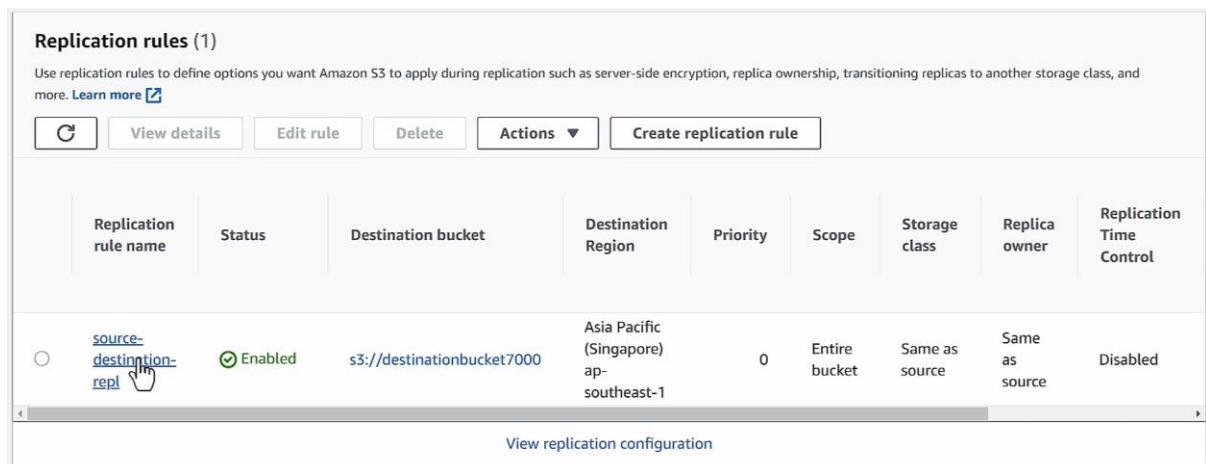
S3 Batch Operations is an Amazon S3 data management feature that lets you manage billions of objects at scale with just a few clicks in the Amazon S3 Management Console or a single API request. With this feature, you can make changes to object metadata and properties, or perform other storage management tasks, such as copying or replicating objects between buckets, replacing object tag sets, modifying access controls, and restoring archived objects from S3 Glacier — instead of taking months to develop custom applications to perform these tasks.

## To begin with the lab

1. Login to AWS Console. Then navigate to S3.
2. Now in the previous lab you performed **Cross Region Replication**.
3. But the problem with Cross Region Replication is that it does not support replication of existing objects.
4. Instead of it you are going to use S3 batch operation.
5. For that you will need an IAM role.
6. So, while doing cross region replication an IAM role was created.
7. You are going to make some changes in that role itself.
8. Now open your Source bucket from where you replicated the data and go to management.



9. There you will see your replication job. And if you open it you can see different option and job details.



Replication rule name	Status	Destination bucket	Destination Region	Priority	Scope	Storage class	Replica owner	Replication Time Control
source-destination-repl	Enabled	s3://destinationbucket7000	Asia Pacific (Singapore) ap-southeast-1	0	Entire bucket	Same as source	Same as source	Disabled

10. Now if you go back to replication rules.
11. You can see that you have your IAM role in place.

12. Now you have to click on this IAM and role and you will be on a new page in a new tab for IAM.

The screenshot shows the 'Replication rules' section of the AWS S3 console. At the top, it displays the path: Amazon S3 > Buckets > sourcebucket7000 > Replication rules. Below this, the title 'Replication rules' has an 'Info' link. A descriptive text explains that replication enables automatic and asynchronous copying of objects across buckets in the same or different AWS Regions. Configuration settings affect all replication rules in the bucket. The configuration settings panel shows a 'Source bucket' set to 'sourcebucket7000', a 'Source Region' set to 'Asia Pacific (Mumbai) ap-south-1', and an 'IAM role' set to 's3crr\_role\_for\_sourcebucket7000'. There are 'Create replication job' and 'Edit' buttons at the top right of this panel.

13. So, this role has certain permission. Now you are going to add some permission to this role.

14. **The JSON code is used from AWS Documentation for S3 batch operation.**

15. Click on add permission, then choose Create inline policy.

The screenshot shows the 'Add permissions' section of the IAM role configuration. It includes buttons for 'Simulate', 'Remove', 'Add permissions', 'Attach policies', and 'Create inline policy'. The 'Create inline policy' button is highlighted with a red box. Navigation controls like '<', '1', and ' >' are also visible.

16. Inside the policy, you have to choose JSON.

The screenshot shows the 'JSON' tab selected in the IAM policy editor. Other tabs like 'Visual' and 'Actions' are also visible. A small icon is present on the right side of the toolbar.

17. Then you have to paste this JSON code there. Remember to change the destination bucket and source bucket name with yours.

{

```
"Version": "2012-10-17",
"Statement": [
    {
        "Action": [
            "s3:PutObject",

```

```
        "s3:PutObjectAcl",
        "s3:PutObjectTagging"
    ],
    "Effect": "Allow",
    "Resource": "arn:aws:s3:::destinationbucket7122/*"
},
{
    "Action": [
        "s3:GetObject",
        "s3:GetObjectAcl",
        "s3:GetObjectTagging",
        "s3>ListBucket"
    ],
    "Effect": "Allow",
    "Resource": [
        "arn:aws:s3:::sourcebucket7123",
        "arn:aws:s3:::sourcebucket7123/*"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "s3:GetObject",
        "s3:GetObjectVersion"
    ],
    "Resource": [
        "arn:aws:s3:::sourcebucket7123/*"
    ]
},
```

```

    {
        "Effect": "Allow",
        "Action": [
            "s3:PutObject"
        ],
        "Resource": [
            "arn:aws:s3:::sourcebucket7123/*"
        ]
    }
}

```

The screenshot shows the AWS IAM Policy editor interface. The left pane displays the JSON code of the policy:

```

1 "Version": "2012-10-17",
2 "Statement": [
3     {
4         "Action": [
5             "s3:PutObject",
6             "s3:PutObjectAcl",
7             "s3:PutObjectTagging"
8         ],
9         "Effect": "Allow",
10        "Resource": "arn:aws:s3:::destinationbucket7122/*"
11    },
12    {
13        "Action": [
14            "s3:GetObject",
15            "s3:GetObjectAcl",
16            "s3:GetObjectTagging",
17            "s3>ListBucket"
18        ],
19        "Effect": "Allow",
20        "Resource": [
21            "arn:aws:s3:::sourcebucket7123",
22            "arn:aws:s3:::sourcebucket7123/*"
23        ]
24    },
25    {
26        "Effect": "Allow",
27        "Action": [
28            ...
29        ]
30    }
]

```

The right pane shows the 'Edit statement' interface with a 'Select a statement' dropdown and a '+ Add new statement' button.

18. Then move to next page, give this policy a name then save it.
19. Now again you have to create an Inline policy. This time this policy is all about initiating the replication.

```

{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:InitiateReplication"
            ],

```

```
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3:::sourcebucket7123/*"
        ],
    },
    {
        "Action": [
            "s3:GetReplicationConfiguration",
            "s3:PutInventoryConfiguration"
        ],
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3:::sourcebucket7123"
        ],
    },
    {
        "Action": [
            "s3:GetObject",
            "s3:GetObjectVersion"
        ],
        "Effect": "Allow",
        "Resource": [
            "arn:aws:s3:::sourcebucket7123/*"
        ],
    },
    {
        "Effect": "Allow",
        "Action": [
            "s3:PutObject"
        ]
    }
}
```

```

        ],
        "Resource": [
            "arn:aws:s3:::sourcebucket7123/*",
            "arn:aws:s3:::sourcebucket7123/*"
        ]
    }
}

```

Policy editor

```

1: {
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "s3:InitiateReplication"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::sourcebucket7123/*"
      ]
    },
    {
      "Action": [
        "s3:GetReplicationConfiguration",
        "s3:PutInventoryConfiguration"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::sourcebucket7123"
      ]
    },
    {
      "Action": [
        "s3.GetObject",
        "s3:GetObjectVersion"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::sourcebucket7123/*"
      ]
    }
  ]
}

```

Visual    **JSON**    Actions ▾   

Select a statement  
Select an existing statement in the policy or add a new statement.

+ Add new statement

JSON Ln 1, Col 0    9162 of 9681 characters remaining

Security: 0 Errors: 0 Warnings: 1 Suggestions: 1

20. Then move to next page, give this policy a name then save it.
21. Now two inline policies have been created so far.
22. Now click on Trust relationships. Click on Edit trust policy.
23. And paste the policy code there.

Permissions    **Trust relationships**    Tags    Access Advisor    Revoke sessions

Permissions    **Trust relationships**    Tags    Access Advisor    Revoke sessions

**Trusted entities**

Entities that can assume this role under specified conditions.

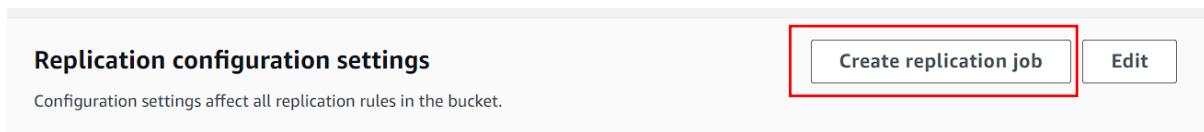
```

1: {
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "batchoperations.s3.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}

```

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": [
                    "batchoperations.s3.amazonaws.com", "s3.amazonaws.com"
                ],
                "Action": "sts:AssumeRole"
            }
        }
    ]
}
```

24. Now this code is giving batch operation the ability to assume this role along with S3.
25. Once this is done go back to your source bucket and navigate to replication rules.
26. There in replication rules. You will see this option to create replication job. Click on it.



27. In replication leave the settings as they are just click on next to move to other page.

### Choose Region and manifest Info

The screenshot shows the 'Choose Region and manifest' step of the replication wizard. It features a 'AWS Region' section with a dropdown menu set to 'Asia Pacific (Mumbai) ap-south-1'. A cursor is hovering over the dropdown menu.

## Manifest

Manifests must only reference objects in a single S3 bucket. To generate a new manifest, [configure an S3 inventory list for a bucket or prefix](#).

### Manifest format

S3 inventory report (manifest.json)

CSV

CSV format must be either 2 or 3 columns in the following order: bucket name, object key, and optionally version ID.

Create manifest using S3 Replication configuration

A list of objects will be generated using the replication configuration and optionally saved to the destination you choose. When using a replication configuration to generate the manifest, the only operation that will be available is replicate. [Learn more](#) or [see pricing](#).

## Source

### Replication configuration source bucket

The source bucket must be located in the same region as your job.

Bucket in this AWS account

Bucket in another AWS account

### Source bucket

Choose or enter the source bucket to generate a manifest using its replication configuration.

s3://sourcebucket7000

[Browse S3](#)

Format: s3://bucket

28. On the next page choose to replicate the objects. Then click on next.

Replicate

Replicates every object to the destinations specified in the replication configuration.



[Cancel](#)

[Previous](#)

[Next](#)

29. Now you will see this option for completion report, on the third page or say step 3 of the replication job.

30. Here you have to choose All tasks and give a S3 URI to it.

31. For that go back to your source bucket and create a folder with name of Reports or whatever you want to name then copy its URI and paste it here.

Amazon S3 > Buckets > sourcebucket7000 > reports/

[Copy S3 URI](#)

reports/

[Objects](#) [Properties](#)

### Folder overview

AWS Region  
Asia Pacific (Mumbai) ap-south-1

S3 URI  
<s3://sourcebucket7000/reports/>

Amazon Resource Name (ARN)  
<arn:aws:s3:::sourcebucket7000/reports/>

**Completion report**

Generate a CSV completion report that lists your target objects, task success or error codes, outputs, and descriptions. Completion reports are encrypted using SSE-S3. [Learn more](#)

Generate completion report

Completion report scope

Failed tasks only

All tasks

Completion report destination

Specify a general purpose bucket location to store the completion report. '/job-{job-id}/report.json' will automatically be appended to the specified destination. [Learn more](#)

s3://sourcebucket7000/reports/ View Browse S3

Format: s3://<bucket>/<optional-prefix-with-path>. S3 will append the path with a "/". If you add a "/" to the prefix, it will appear as an extra folder in the S3 console.

### 32. Now choose the IAM role.

▶ [View IAM role policy template and IAM trust policy](#)

Choose from existing IAM roles

Enter IAM role ARN

IAM role

s3crr\_role\_for\_sourcebucket7000 View

### 33. Now click on next then on the review page create your replication job.

34. Once your replication job is created, you can see this in the batch operations.

Jobs (1)									<span style="border: 1px solid #ccc; padding: 2px;">C</span>	<span style="border: 1px solid #ccc; padding: 2px;">Run job</span>	<span style="border: 1px solid #ccc; padding: 2px;">Actions</span>	<span style="border: 1px solid #ccc; padding: 2px;">Clone job</span>	<span style="border: 1px solid #ccc; padding: 2px;">Create job</span>
<span style="border: 1px solid #ccc; padding: 2px;">Q</span> Search by job ID or description			All status types						<	1	>	<span style="border: 1px solid #ccc; padding: 2px;">⟳</span>	
Job ID	Status	Description	Operation	Date created	Total objects	% Complete	Total failed (rate)						
1e450f11-feb2-40a3-9d70-0ca6eb5944a2	<span style="color: #0070C0;">🕒</span> Preparing	2023-08-14 - Replicate	Replicate	August 14, 2023, 12:43:15 (UTC+04:00)	0	0%	0 (0%)						

### 35. Now to run this job first you need to wait until it asked you to run this job.

36. Then open this job.

Jobs (1)								
Search by job ID or description			All status types		Actions			
Asia Pacific (Mumbai) ap-south-1							< 1 >	
Job ID	Status	Description	Operation	Date created	Total objects	% Complete	Total failed (rate)	
<a href="#">1e450f11-feb2-40a3-9d70-0ca6eb5944a2</a>	Awaiting your confirmation to run	2023-08-14 - Replicate	Replicate	August 14, 2023, 12:43:15 (UTC+04:00)	4	0%	0 (0%)	P

37. Now click on run job.

Amazon S3 > Batch Operations > View details: 1e450f11-feb2-40a3-9d70-0ca6eb5944a2

Job ID 1e450f11-feb2-40a3-9d70-0ca6eb5944a2

Run job Actions Clone job C

38. Once you have run the job, it will show that the job is in active state.

39. Now you have to wait for it to get completed. It will take some time.

40. Once it gets completed you can see all your objects replicated to destination bucket.

Jobs (1)								
Search by job ID or description			All status types		Actions			
Asia Pacific (Mumbai) ap-south-1							< 1 >	
Job ID	Status	Description	Operation	Date created	Total objects	% Complete	Total failed (rate)	
<a href="#">1e450f11-feb2-40a3-9d70-0ca6eb5944a2</a>	Active	2023-08-14 - Replicate	Replicate	August 14, 2023, 12:43:15 (UTC+04:00)	4	0%	0 (0%)	P

41. After couple of minutes, you can see that your job is completed.

Jobs (1)								
Search by job ID or description			All status types		Actions			
Asia Pacific (Mumbai) ap-south-1							< 1 >	
Job ID	Status	Description	Operation	Date created	Total objects	% Complete	Total failed (rate)	
<a href="#">1e450f11-feb2-40a3-9d70-0ca6eb5944a2</a>	Completed	2023-08-14 - Replicate	Replicate	August 14, 2023, 12:43:15 (UTC+04:00)	4	100%	0 (0%)	P

42. Now you can go to your destination bucket to check for the objects.

43. So, about deletion of the job, it can't be deleted. It will automatically will delete itself after 90 days.

44. You just need to delete your source and destination bucket.

## Batch Operations [Info](#)

A job is used to execute batch operations on a list of S3 objects. The list of S3 objects is contained in a manifest object, which can be an S3 inventory report or a list of objects that you generate. After the total number of objects listed in the manifest has been confirmed, the job status will update to *Awaiting your confirmation to run*, and you must **Run job** within 30 days. Job events are published to [CloudWatch Events](#). Jobs are deleted 90 days after they finish or fail. [Learn more](#)