AWS S3

Amazon S3

S3 Performance Optimization

Optimizing Performance for High-Throughput Workloads

- **Parallelizing Data Transfers**: Use **multipart uploads** for large files.
- **Prefixing Object Keys**: Distribute objects across partitions using prefixes.
- **S3 Transfer Acceleration**: Uses AWS edge locations to speed up global transfers.
- **Amazon CloudFront Integration**: Reduces latency by caching objects closer to users.
- **Compression**: Reduce storage size and improve transfer speeds using gzip or other compression methods.

Amazon S3 Lifecycle Policies

What Are Lifecycle Policies?

- Automates transitions of objects between different S3 storage classes.
- Helps optimize cost by moving less frequently accessed data to cheaper storage tiers.
- Can be used for automatic deletion of objects after a specified period.

Lifecycle Policy Use Cases

Action	Storage Class	Example Use Case
Transition to S3 Standard-IA	After 30 days	Archive infrequently accessed logs
Move to S3 Glacier	After 90 days	Retain data for compliance
Delete objects	After 365 days	Remove old backups automatically

Example: Create a Lifecycle Policy to Move Objects to Glacier

```
{
  "Rules": [
    {
      "ID": "MoveToGlacier",
      "Status": "Enabled",
      "Filter": {},
      "Transitions": [
        {
          "Days": 90,
          "StorageClass": "GLACIER"
        }
      ],
      "Expiration": {
        "Days": 365
      }
    }
  ]
}
```

Cross-Region Replication (CRR) & Same-Region Replication (SRR)

What is S3 Replication?

- **CRR (Cross-Region Replication)**: Replicates objects across AWS regions for disaster recovery.
- **SRR (Same-Region Replication)**: Duplicates objects within the same AWS region for compliance.

Steps to Enable Cross-Region Replication

1. Enable **versioning** on both source and destination buckets.
2. Create an **IAM role** with necessary permissions.
3. Configure replication rules in the S3 console.

aws s3api put-bucket-replication --bucket source-bucket --replication-configuration file://replication.json

Example replication.json:

```json
{
  "Role": "arn:aws:iam::123456789012:role/s3-replication-role",
  "Rules": [
    {
      "Status": "Enabled",
      "Priority": 1,
      "Filter": {},
      "Destination": {
        "Bucket": "arn:aws:s3:::destination-bucket",
        "StorageClass": "STANDARD"
      }
    }
  ]
}
```

S3 Transfer Acceleration

What is S3 Transfer Acceleration?

- Uses AWS CloudFront edge locations to speed up uploads.
- Reduces latency for global users.

How to Enable S3 Transfer Acceleration

1. Open the S3 Console.
2. Navigate to the bucket settings.
3. Enable **Transfer Acceleration** under **Properties**.
4. Use the accelerated endpoint for uploads:

aws s3 cp largefile.zip s3://my-bucket/ --endpoint-url https://my-bucket.s3-accelerate.amazonaws.com

Summary

- Optimize S3 performance using multipart uploads, prefixing, and Transfer Acceleration.
- Automate data lifecycle management with **Lifecycle Policies**.
- Ensure disaster recovery with **Cross-Region Replication**.
- Improve upload speeds with **S3 Transfer Acceleration**.