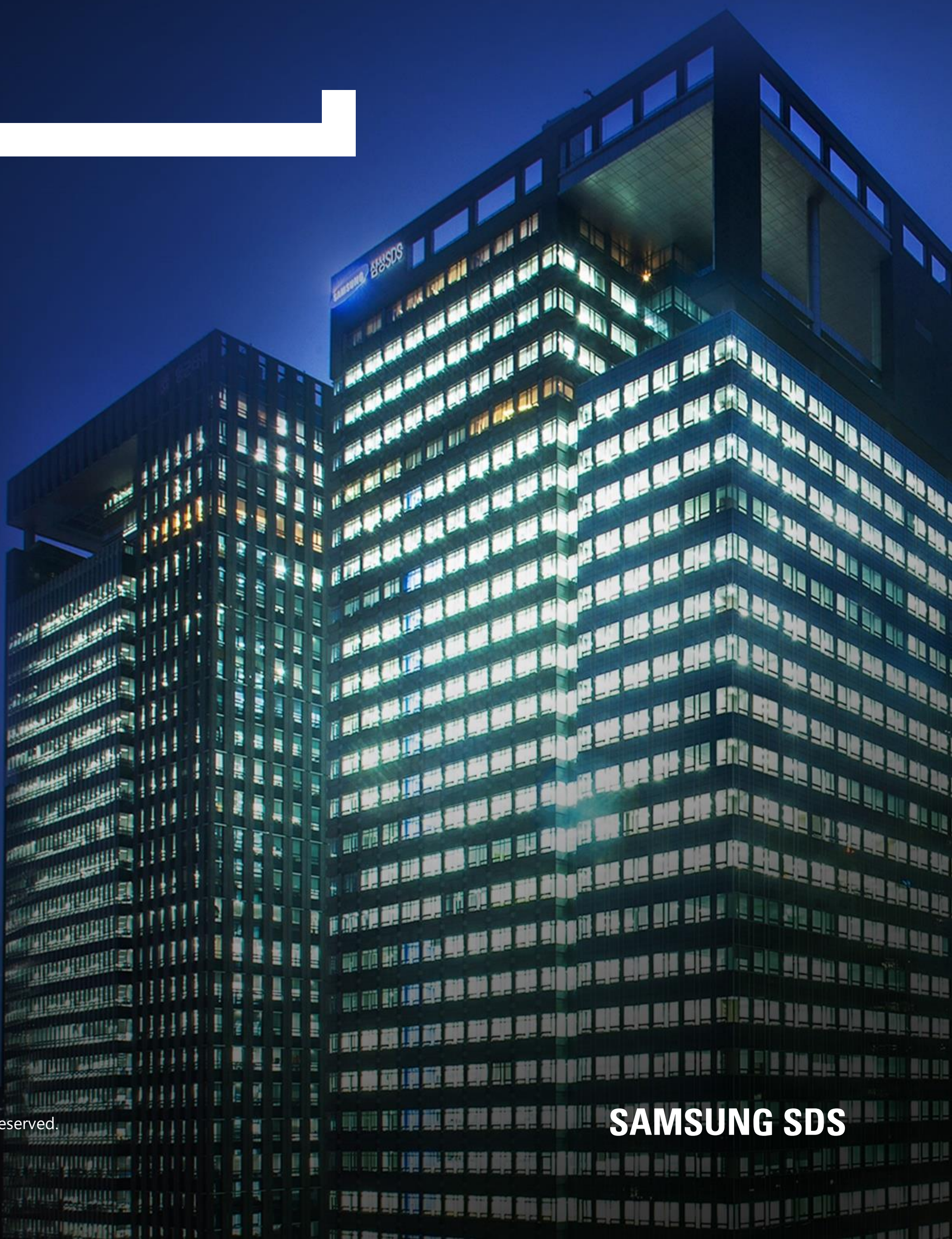
Technical Guide

Backup configuration in a cloud environment

October 2021



SAMSUNG SDS

Contents

1. Overview

This document will cover how to establish a backup plan for data protection in the SDS Cloud environment.

Data loss can occur at any time due to human errors, H/W problems (power failure or damage caused by aging or shock), S/W issues, theft, disasters, or viruses. In the event of an accident, the importance of a backup plan for recovery based on thorough backups is particularly highlighted.

A well-designed and tested backup scheme can minimize the impact on the business in case of data loss. SDS Cloud provides materials for constructing a robust, flexible and cost-effective backup recovery measure that can be used to build or extend the continuity of customer-built systems.

The definition of backup, or data backup, is temporary storage. It means that data is temporarily replicated in advance in order to ensure recovery in case of a problem.

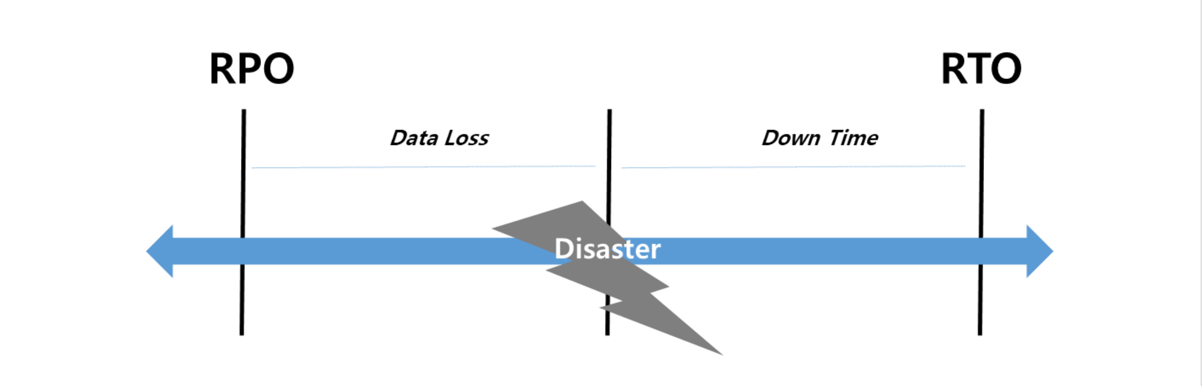


Figure 1. RPO and RTO as of time of data loss

Recovery resumes the IT services that have been interrupted by data loss. For recovery, finding a strategy that meets both Recovery Time Objective (RTO) and Recovery Point Objective (RPO) is the key. RTO refers to the maximum allowed time it takes to restore the service when the service is interrupted due to data loss. RPO refers to the point at which data loss is tolerated when the interrupted service is restored.

In general, the smaller the RTO and RPO values and the faster the application must recover from an interruption, the higher the cost of running the application. Smaller RTO and RPO values also increase the administrative overhead with greater complexity. For more secure data protection, replication between two physically separated data centers also needs to be managed.

2. Backup plan overview

The elements composing the backup S/W are as shown in the diagram below.

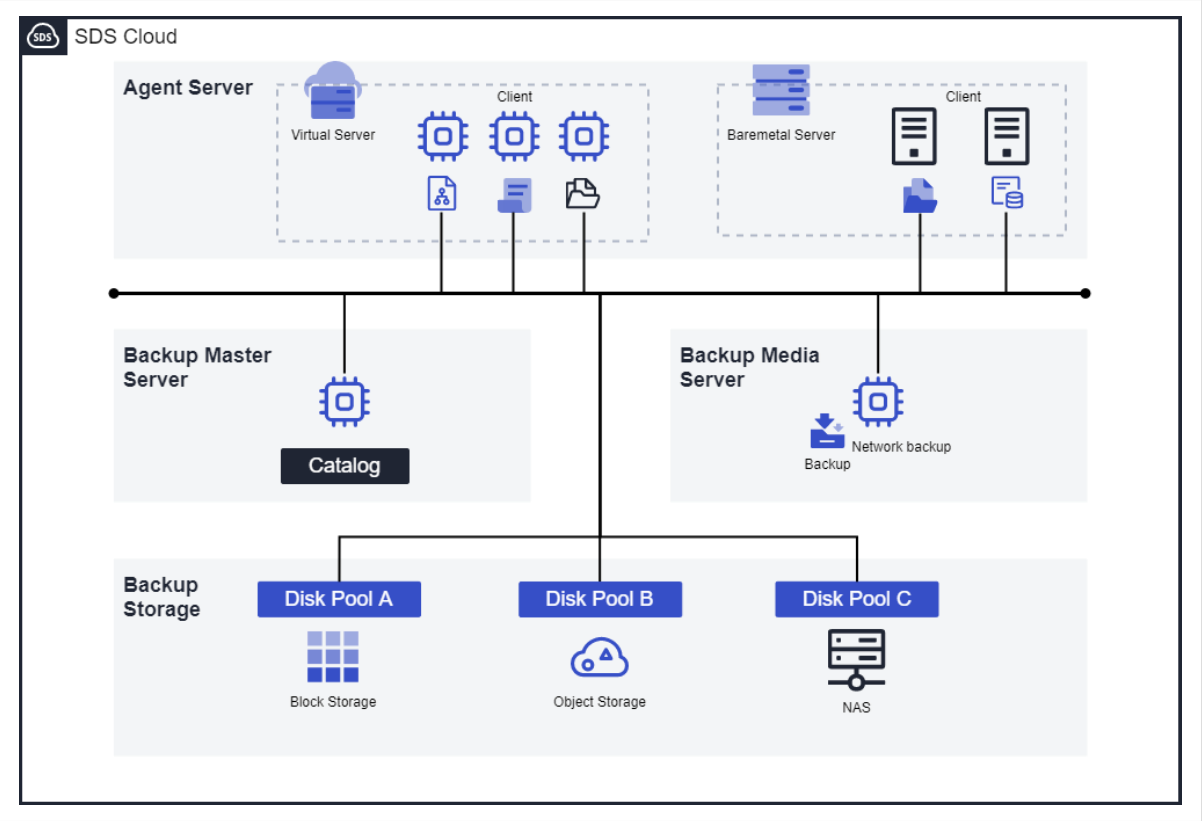


Figure 2. Backup S/W components

2.1 Backup S/W Configuration

Backup S/W consists of a backup master server, media server, and agent client system.

- **Backup master server**
  - Centralized operation and management, and overall backup environment operation
  - Archive operation
  - Recovery operation
  - Automated backup scheduling
  - Record of backup configuration information
  - Record of job status information
  - Record of backup image information

- **Backup media server**
  - A system that provides storage resources for backup data
  - A system that receives commands from the master server and stores the data stream sent from the client in the storage unit attached to the server
  - Deduplication and replication for vaulting to a remote location

- **Backup client**
  - The system where the backup agent is installed as the backup target system
  - Data delivery to backup media server via network

2.2 Backup target

The backup target can be defined in units of files, folders, disks, or partitions.

- **Backing up files and folders**
  - Log files such as OS system, security, and application logs
  - User and application data
  - Database data
  - NAS files
  - Recovery often requires a sequence.

- **System backup**
  - Backup of the entire system including the operating system or application S/W in the hard disk
  - Recovery requires a restoration of the entire system, resulting in a simple recovery sequence.

2.3 Backup storage devices

- **Block Storage**
  - Internal and external storage devices
  - High performance and high cost

- **File Storage**
  - NFS and CIFS remote mount
  - Management convenience and low cost (higher cost compared to Object Storage)

- **Object Storage**
  - AWS cloud or S3-compatible storage
  - Large capacity and low cost
  - Compatibility verification with backup S/W required

2.4 Estimating backup server capacity

The backup master server's capacity is calculated according to the compatibility of each backup S/W and the frequency of using catalog DB, while the backup media server's capacity is estimated according to the backup target's capacity, backup cycle, and storage cycle.

Backup target capacity	Backup master (+media) server specifications			Object Storage
	CPU	Memory	SSD	
10TB	over 1	over 48GB	over 500GB	over 15TB
20TB	over 2	over 128GB	over 1TB	over 25TB
40TB	over 2	over 192GB	over 2TB	over 45TB
100TB	over 2	over 384GB	over 3TB	over 110TB

Table 1. Backup master + media server capacity (based on 1-day backup and 30-day storage)

Table 1 includes the recommended values for specifications when configuring the backup master server and media server in a single server.

	Partition	Required Capacity	Remarks
Linux	/tmp	3GB	1 GB of additional index storage space is required for every million backup files. *Additional capacity is required if performing NAS backup through a client server.
	/opt	3GB	
	/usr	3-5GB	
Windows	C:\	3-5GB	

Table 2. Capacity required for client installation

Table 2 is the recommended capacity for each partition to install the agent on the backup target server of the client. To estimate detailed backup server capacity, please refer to the installation guide for each solution.

2.5 Backup agent server deployment and port settings

Agent servers can be installed by manual or automatic push method. Before installing the agent server, communications and firewalls with the backup master/media server should be checked, and the server information should be registered in the hosts file.

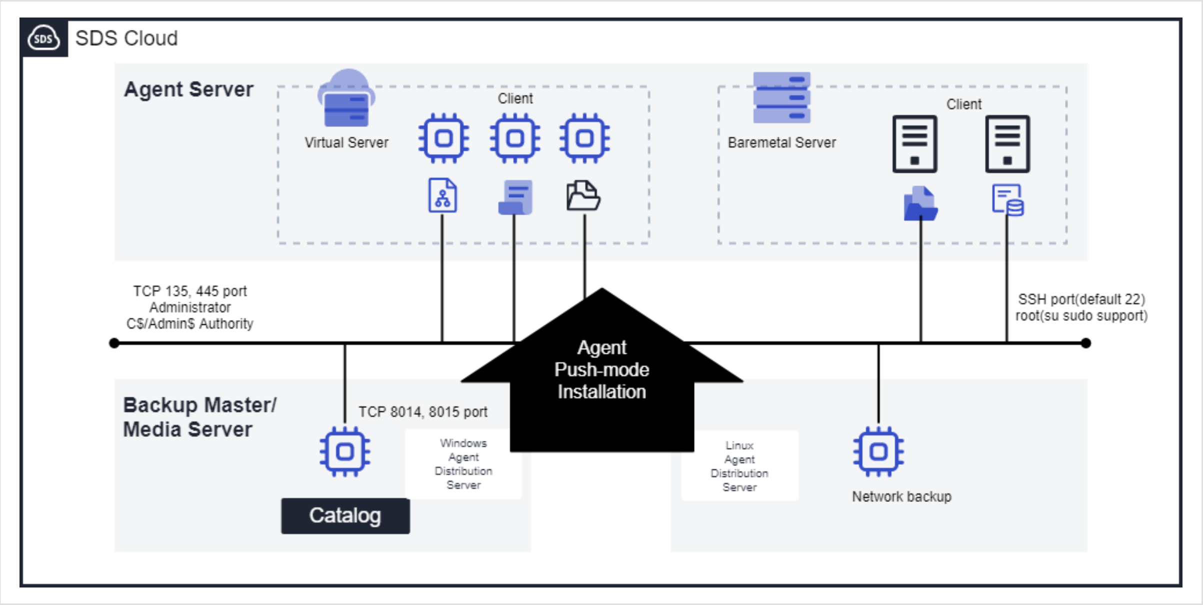


Figure 3. Arcserve UDP agent deployment

Figure 3 shows the port configuration required for installing an Arcserve UDP agent server and automatic deployment method of push method as an example.

This product requires a distribution server for Windows and Linux servers, and the backup master server acts as a distribution server for Windows. The distribution server offers communications using TCP ports 135, 445 and SSH port 22 while the master server uses TCP ports 8014 and 8015. The firewall settings need to be checked before installation.

For detailed information on communications ports and agent installation methods for each backup S/W, please refer to the installation guide for each vendor.

3. SDS Cloud backup configuration options

SDS Cloud recommends using a low-cost, large-capacity Object Storage product as a backup storage device. Two ways to recognize Object Storage as a backup storage device are provided. Users can configure the backup environment using commercial backup S/W or self-developed tools.

3.1 Object Storage access using backup S/W

3.1.1 Backup S/W compatible with SDS Cloud Object Storage

SDS Cloud provides S3-compatible Object Storage and can be recognized as a disk pool by directly interworking with backup S/W. For SDS Cloud, it is recommended to use Veritas Netbackup or Arcserve UDP, which have been pre-validated for compatibility.

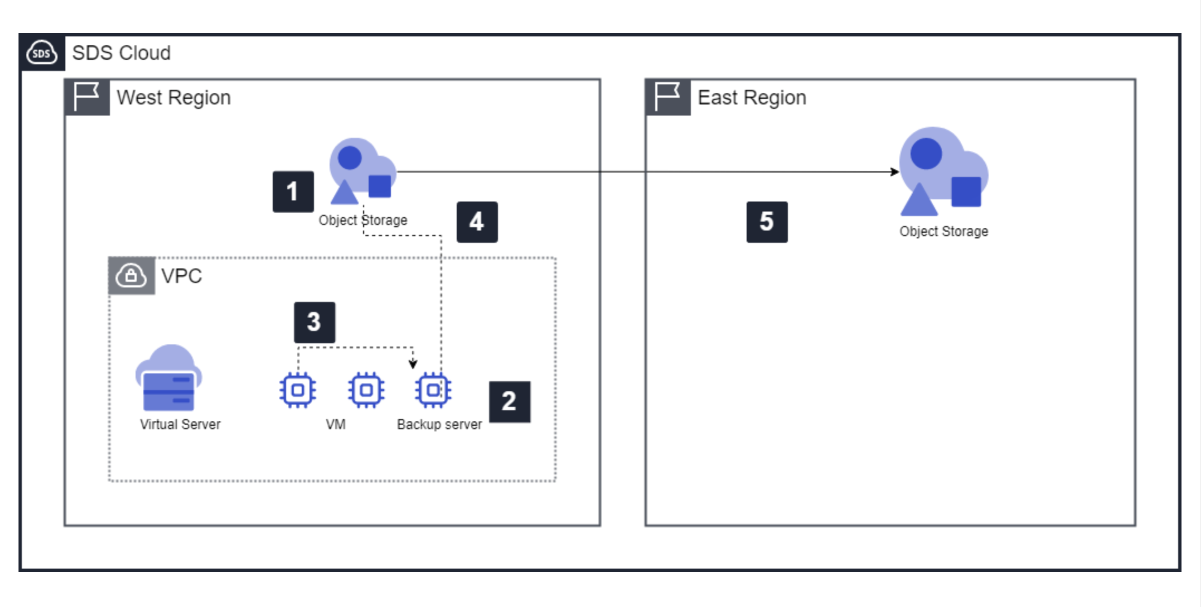


Figure 4. Backup S/W compatible Object Storage

1. Create a bucket in Object Storage and issue an access URL, access key, and secret key.
2. Install the backup server on the SDS Cloud Virtual Server and register the bucket as a backup disk (disk pool).
3. Install the agent on the Virtual Server targeted for backup and register the client on the backup master server.
4. Set backup plan and policy, and perform agent backup.
5. Duplicate to an Object Storage in another region.

3.2 Object Storage access using mount solution

3.2.1 Backup S/W incompatible with SDS Cloud Object Storage

If the installed backup S/W that is not compatible with SDS Cloud Object Storage, backup/recovery can be configured by locally mounting the bucket to the backup server.

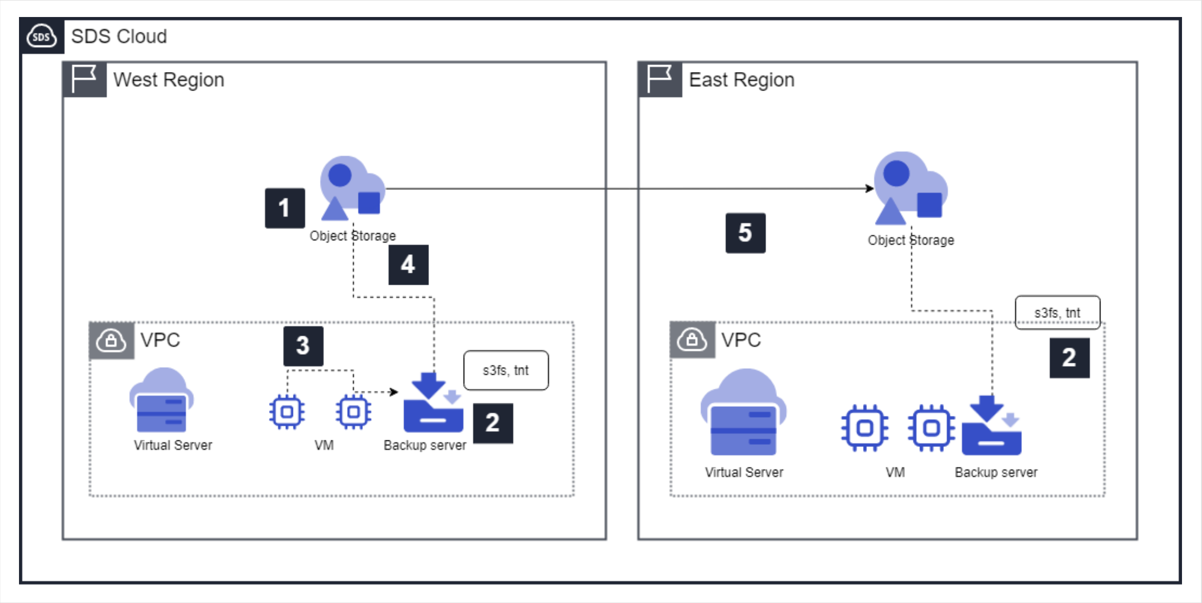


Figure 5. Backup S/W backup/recovery through locally mounted bucket

1. Create a bucket in Object Storage and issue an access URL, access key, and secret key.
2. After installing the backup server on SDS Cloud Virtual Server, mount the bucket locally on the backup server using the s3fs solution and register it as a disk pool.
3. Install the agent on the Virtual Server targeted for backup and register the client on the backup master server.
4. Set backup plan and policy, and perform agent backup.
5. If necessary, duplicate to an Object Storage in another region.

3.2.2 SDS Cloud Object Storage using self-developed tools

If a self-developed tool/script is used, backup/recovery can be configured by using a solution that mounts the bucket locally on the server.

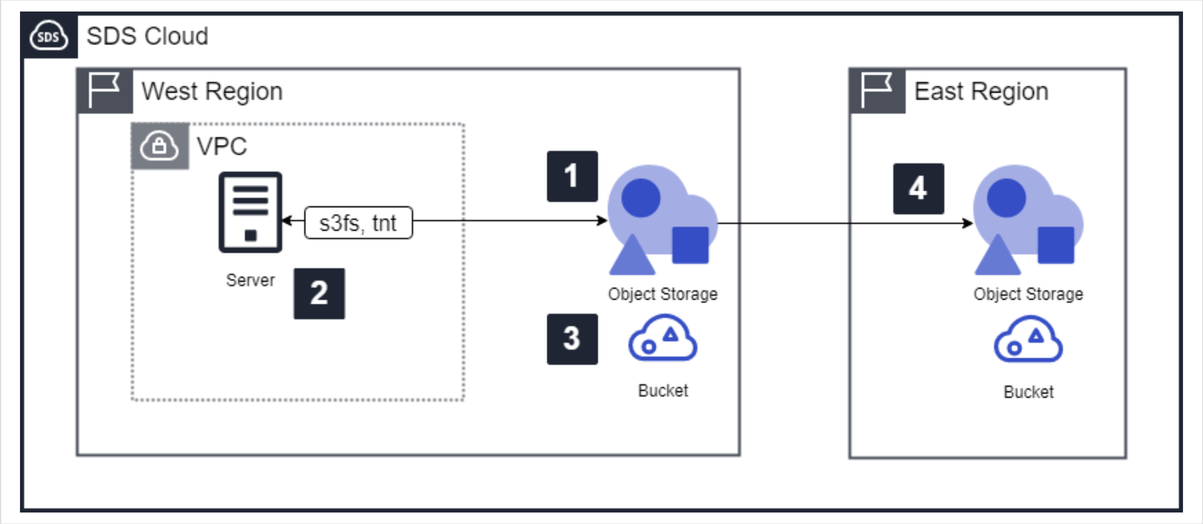


Figure 6. Self-developed tool backup/recovery through locally mounted bucket

1. Create a bucket iObject Storage and issue an access URL, access key, and secret key.
2. Mount the bucket locally on the server using the s3fs solution.
3. Back up to bucket with a self-developed tool.
4. If necessary, duplicate to an Object Storage in another region.

3.3 Detailed configuration of locally mounted bucket

3.3.1 Installation and configuration

- **Step 1: Install solutions**

Linux: s3fs rpm installation

```
# yum install -y s3fs-fuse.x86_64
```

Windows: Rclone, Rclone-browser, WinFsp installation

- https://rclone.org/downloads/
- https//github.com/billziss-gh/winfsp/releases

Step 2: Set up access environments for the bucket

Linux

```
# echo ACCESSKEY:SECRETKEY >   /etc/passwd-s3fs
# chmod 600 /etc/passwd-s3fs
```

Windows: rclone configurations

```
1) cd c:\rclone
2) rclone config
   - n/s/q> n // New remote
   - name> MYBUCKET // Set the name
   - Type of storage
     4: Amazon S3 Compliant Storage Providers
     14: Any other S3 compatible provider
   - Authentication environment
     false: Enter AWS credentials in the next step
     access_key_id: Access key id
     secret_access_key: Secret access key
   - Region and endpoint URL
     1: v4 signature and an empty region
     endpoint: Endpoint URL
   - Location constraint
     “”(default): Owner gets FULL_CONTROL
   - Edit advanced config and save
     n: Edit advanced config
     y: Yes this is OK
```

Step 3: Mount the bucket

Linux

```
s3fs mybucket /MYBUCKET -o passwd_file=~/.passwd-s3fs –o url=BUCKET ENDPOINT –o allow_other -o nomultipart –o use_path_request_style
```

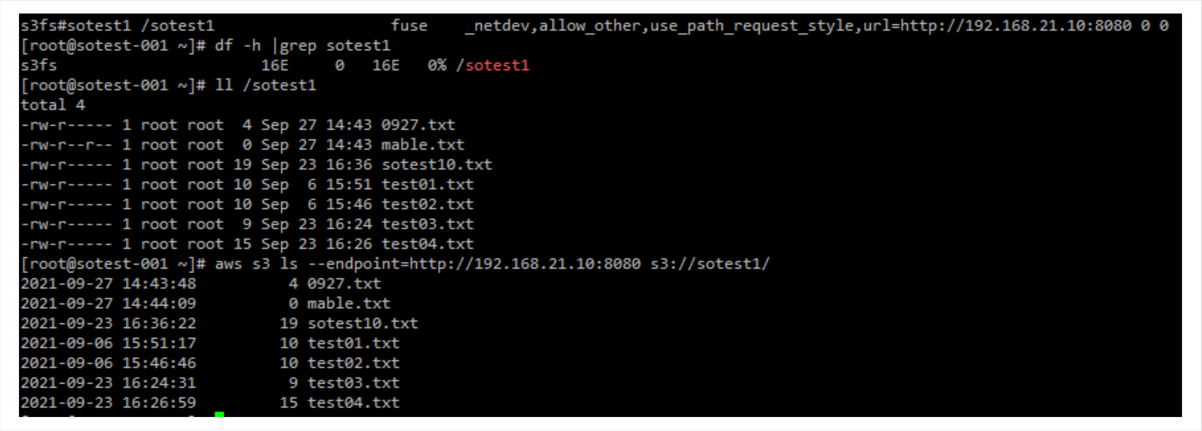
Windows: Register a bucket as a drive

```
1) Rclone.exe mount NAME:/ DRIVE:
2) Check a drive mount
```

3.3.2 Verification method

A. Linux OS log backup environment

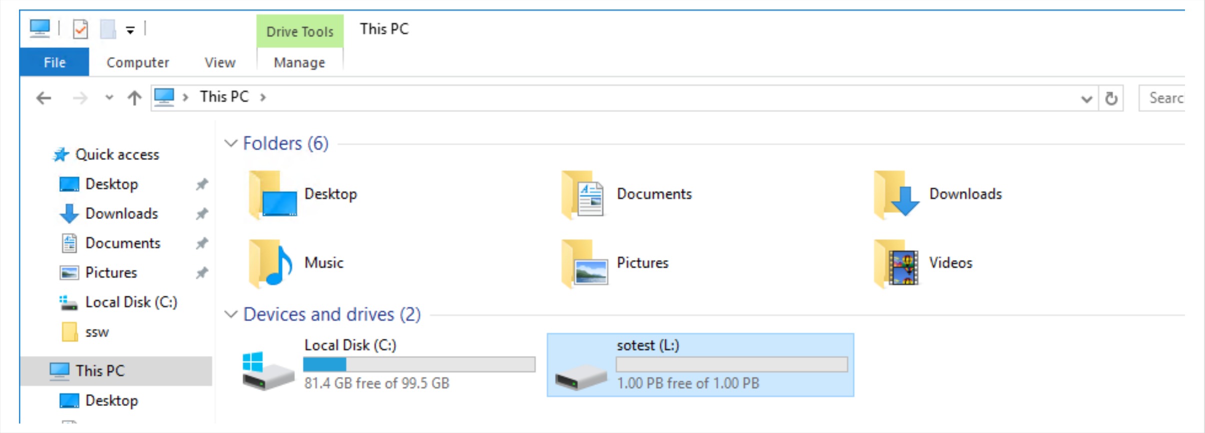
- Compare the bucket’s mount information and s3api information. (Refer to S3 API documents).



```
s3fs#sotest1 /sotest1                  fuse    _netdev,allow_other,use_path_request_style,url=http://192.168.21.10:8080 0 0
[root@sotest-001 ~]# df -h |grep sotest1
s3fs                       16E     0   16E    0% /sotest1
[root@sotest-001 ~]# ll /sotest1
total 4
-rw-r----- 1 root root  4 Sep 27 14:43 0927.txt
-rw-r--r-- 1 root root  0 Sep 27 14:43 mable.txt
-rw-r----- 1 root root 19 Sep 23 16:36 sotest10.txt
-rw-r----- 1 root root 10 Sep  6 15:51 test01.txt
-rw-r----- 1 root root 10 Sep  6 15:46 test02.txt
-rw-r----- 1 root root  9 Sep 23 16:24 test03.txt
-rw-r----- 1 root root 15 Sep 23 16:26 test04.txt
[root@sotest-001 ~]# aws s3 ls --endpoint=http://192.168.21.10:8080 s3://sotest1/
2021-09-27 14:43:48          4 0927.txt
2021-09-27 14:44:09          0 mable.txt
2021-09-23 16:36:22         19 sotest10.txt
2021-09-06 15:51:17         10 test01.txt
2021-09-06 15:46:46         10 test02.txt
2021-09-23 16:24:31          9 test03.txt
2021-09-23 16:26:59         15 test04.txt
```

B. Windows OS log backup environment

- Mount the bucket to L drive.



3.3.3 Prerequisites

The private Access Control List (ACL) information of the bucket to be mounted must be checked in advance.

- Object Storage Endpoint, access key ID, and secret access key

4. Considerations

If DR is configured with an inter-region replication, the overall DR cost should be predicted by estimating the traffic usage on the replication line. The backup policy (backup frequency setting) and replication frequency also need to be configured to align with the RPO goals.

In particular, for public cloud services, DR configuration using SDS Cloud resources should analyze the outbound replication traffic volume of the public cloud service provider and the resulting network cost and the decision should be made accordingly.

In addition, users must either use their own backup software licenses, or purchase, install and sign a supporting contract on their own.

If access control is required, it can be set in two ways. In order to transfer the backup data to Object Storage for storage on the Virtual Server located inside the VPC, it is necessary to set up communications on both the Virtual Server and the VPC.

For this, a Security Group (for Virtual Server or VPC) needs to be set as follows:

- Source: Virtual Server IP Address
- Destination: Object Storage End-Point IP address
- Port: 80 and 443 TCP
- Direction: One-way from Virtual Server to Object Storage

In Object Storage, the IP address can be set to access in units of buckets.

- Target: Bucket to mount
- Settings: Virtual Server Public IP address