**Module 14 – Guided Lab: Hybrid Storage and Data Migration with AWS Storage Gateway File Gateway**

**Lab overview and objectives**

In this lab, you will use the AWS Storage Gateway File Gateway service to attach a Network File System (NFS) mount to an on-premises data store. You will then replicate that data to an S3 bucket in AWS. Additionally, you will configure advanced Amazon S3 features, like Amazon S3 lifecycle policies and cross-Region replication.

After completing this lab, you should be able to:

* Configure a File Gateway with an NFS file share and attach it to a Linux instance
* Migrate a set of data from the Linux instance to an S3 bucket
* Create and configure a primary S3 bucket to migrate on-premises server data to AWS
* Create and configure a secondary S3 bucket to use for cross-Region replication
* Create an S3 lifecycle policy to automatically manage data in a bucket

**Duration**

This lab will require approximately **90 minutes** to complete.

**AWS service restrictions**

In this lab environment, access to AWS services and service actions might be restricted to the ones that are needed to complete the lab instructions. You might encounter errors if you attempt to access other services or perform actions beyond the ones that are described in this lab.

**Accessing the AWS Management Console**

1. At the top of these instructions, choose Start Lab to launch your lab.

A **Start Lab** panel opens, and it displays the lab status.

**Tip**: If you need more time to complete the lab, restart the timer for the environment by choosing the Start Lab button again.

1. Wait until the **Start Lab** panel displays the message *Lab status: ready*, then close the panel by choosing the **X**.

1. At the top of these instructions, choose AWS.

This action opens the AWS Management Console in a new browser tab. The system automatically logs you in.

**Tip**: If a new browser tab does not open, a banner or icon is usually at the top of your browser with the message that your browser is preventing the site from opening pop-up windows. Choose the banner or icon, and then choose **Allow pop-ups**.

1. Arrange the **AWS Management Console** tab so that it displays alongside these instructions. Ideally, you will have both browser tabs open at the same time so that you can follow the lab steps more easily.

**Do not change the Region unless specifically instructed to do so**.

**Task 1: Reviewing the lab architecture**

This lab environment uses a total of three AWS Regions. A Linux EC2 instance that emulates an on-premises server is deployed to the us-east-1 (N. Virginia) Region. The Storage Gateway virtual appliance is deployed to the same Region as the Linux server. In a real-world scenario, the appliance would be deployed in a VMware vSphere or Microsoft Hyper-V environment, or as a physical Storage Gateway appliance.

The primary S3 bucket is created in the us-east-2 (Ohio) Region. Data from the Linux host is copied to the primary S3 bucket. This bucket can also be called the *source*.

The secondary S3 bucket is created in the us-west-2 (Oregon) Region. This secondary bucket is the *target* for the cross-Region replication policy. It can also be called the *destination*.

**Task 2: Creating the primary and secondary S3 buckets**

Before you configure the File Gateway, you must create the primary S3 bucket (or the source) where you will replicate the data. You will also create the secondary bucket (or the destination) that will be used for cross-Region replication.

1. In the search box to the right of **Services**, search for and choose **S3** to open the S3 console.

1. Choose **Create bucket** then configure these settings:
   * **Bucket name:** Create a name that you can remember easily. It must be globally unique.
   * **Region:** *US East (Ohio) us-east-2*
   * **Bucket Versioning:** *Enable*

For cross-Region replication, you must enable versioning for both the source and destination buckets.

1. Choose **Create bucket**

1. Repeat the previous steps in this task to create a second bucket with the following configuration:
   * **Bucket name:** Create a name you can easily remember. It must be globally unique.
   * **Region:** *US West (Oregon) us-west-2*
   * **Versioning:** *Enable*

**Task 3: Enabling cross-Region replication**

Now that you created your two S3 buckets and enabled versioning on them, you can create a replication policy.

1. Select the name of the *source* bucket that you created in the US East (Ohio) Region.

1. Select the **Management** tab and under *Replication rules* select **Create replication rule**

1. Configure the Replication rule:
   * **Replication rule name:** crr-full-bucket
   * **Status** *Enabled*
   * **Source bucket**:
     + For **Choose a rule scope**, select *Apply to all objects in the bucket*
   * **Destination:**
     + *Choose a bucket in this account*
     + Choose **Browse S3** and select the bucket you created in the US West (Oregon) Region.
     + Select **Choose path**
     + **IAM role:** *S3-CRR-Role*
       - **Note:** To find the AWS Identity and Access Management (IAM) role, in the search box, enter: S3-CRR (This role was pre-created with the required permissions for this lab)

1. Choose **Save**. When prompted, if you want to replicate existing objects, choose **No**, and then choose **Submit**

**Note**: there are no objects currently in the bucket, so the answer will have no effect in this case.

1. Return to and select the link to the bucket you created in the *US East (Ohio)* Region.

1. Choose **Upload** to upload a file from your local computer to the bucket.

*For this lab, use a small file that does not contain sensitive information, such as a blank text file.*

1. Choose **Add files**, locate and open the file, then choose **Upload**

1. Wait for the file to upload, then choose **Close**. Return to the bucket you created in the *US West (Oregon)* Region.

The file that you uploaded should also now have been copied to this bucket.

**Note:** You may need to refresh the console for the object to appear.

**Task 4: Configuring the File Gateway and creating an NFS file share**

In this task, you will deploy the File Gateway appliance as an Amazon Elastic Compute Cloud (Amazon EC2) instance. You will then configure a cache disk, select an S3 bucket to synchronize your on-premises files to, and select an IAM policy to use. Finally, you will create an NFS file share on the File Gateway.

1. In the search box to the right of **Services**, search for and choose **Storage Gateway** to open the Storage Gateway console.

1. At the top-right of the console, verify that the current Region is *N. Virginia*.

1. Choose **Create gateway** then begin configuring the **Step 1: Set up gateway** settings:
   * Gateway name: File Gateway
   * Gateway time zone: Choose **GMT -5:00 Eastern Time (US & Canada), Bogota, Lima**
   * Gateway type: **Amazon S3 File Gateway**
   * Host platform: choose **Amazon EC2**. Choose **Customize your settings**. Then choose the **Launch instance** button.

A new tab opens to the EC2 instance launch wizard. This link automatically selects the correct Amazon Machine Image (AMI) that must be used for the File Gateway appliance.

2. In the Launch an instance screen, begin configuring the gateway as described:
   * **Name**: File Gateway Appliance
   * **AMI from catalog**: Accept the default aws-storage-gateway AMI.
   * **Instance type:** Select the **t2.xlarge** instance type

**Note:** t2.xlarge is the only instance type that you can select in this lab environment. If you select any other instance type, it will result in an error message when you attempt to launch the instance.

The t2.xlarge instance type is used only as an example in this lab. For correct appliance sizing when you deploy a Storage Gateway appliance, refer to the [Storage Gateway documentation](https://docs.aws.amazon.com/storagegateway/latest/userguide/Requirements.html#requirements-hardware-storage).

* + **Key pair name - *required***: choose the existing **vockey** key pair.

**Note:** This SSH key pair is provided on the Details > Show page for this lab.

1. Configure the network and security group settings for the gateway.
   * Next to *Network settings*, choose **Edit**, then configure:
     + **VPC**: On-Prem-VPC
     + **Subnet:** *On-Prem-Subnet*
     + **Auto-assign public IP:** *Enable*
     + Under Firewall (security groups), choose **Select an existing security group**.
   * For **Common security groups**:
     + Select the security group with **FileGatewayAccess** in the name

**Note**: This security group is configured to allow traffic through ports 80 (HTTP), 443 (HTTPS), 53 (DNS), 123 (NTP), and 2049 (NFS). These ports enable the activation of the File Gateway appliance. They also enable connectivity from the Linux server to the NFS share that you will create on the File Gateway.

For additional information about the ports used by Storage Gateway, refer to the [Storage Gateway documentation](https://docs.aws.amazon.com/storagegateway/latest/userguide/Requirements.html#networks).

* + - Also select the security group with **OnPremSshAccess** in the name

**Note**: This security group is configured to allow Secure Shell (SSH) connections on port 22.

* + - Verify that both security group now appear as selected (details on each will appear in boxes in the console).

**Tip**: You may need to choose **Show all selected** to see them both.

2. Configure the storage settings for the gateway.
   * In the Configure storage panel, notice there is already an entry to create one 80GiB root volume.
   * Choose **Add new volume**
   * Set the size of the EBS volume to 150GiB

1. Finish creating the gateway.
   * In the *Summary* panel on the right, keep the number of instances set to 1, and choose **Launch instance**

A Success message displays.

* + Choose **View all instances**

Your *File Gateway Appliance* instance will take a few minutes to initialize.

1. Monitor the status of the deployment and wait for **Status Checks** to complete.

**Tip:** Choose the refresh button to more quickly learn the status of the instance.

1. Select your File Gateway instance, then in the *Details* tab below, locate the **Public IPv4 address** and copy it.

You will use this IP address when you complete the File Gateway deployment.

1. Return to the **AWS Storage Gateway** tab in your browser. It should still be at the **Set up gateway on Amazon EC2** screen.

1. Check the box next to *I completed all the steps above and launched the EC2 instance*, then choose **Next**

1. Configure the **Step 2: Connect to AWS** settings:
   * In the Gateway connection options:
     + For **IP address**, paste in the **IPv4 Public IP** address that you copied from your File Gateway Appliance instance
   * For the **Service endpoint**, select **Publicly accessible**.
   * Choose **Next**

1. In the **Step 3: Review and activate** settings screen choose **Activate gateway**

1. Configure the **Step 4: Configure gateway** settings:
   * CloudWatch log group: **Deactivate logging**
   * CloudWatch alarms: **No Alarm**

A *Successfully activated gateway File Gateway Appliance* message displays.

In the Configure cache storage panel, you will see that a message the local disks are loading.

* + Wait for the local disks status to show that it finished processing (approximately 1 minute).
  + Choose **Configure**

1. Start creating a file share.
   * Wait for **File Gateway** status to change to **Running**.
   * From the left side panel, choose **File shares**.
   * Choose **Create file share**.
2. On the **Create file share** screen, configure these settings:
   * **Gateway:** Select the name of the File Gateway that you just created (which should be *File Gateway* *Appliance*)
   * **File share protocol:** **NFS**
   * **Amazon S3 bucket name**: Choose the name of the source bucket that you created in the US East (Ohio) us-east-2 Region in Task 1.
   * Choose **Customize configuration**
   * For **File share name** use share and choose **Next**.
3. On the **Amazon S3 storage settings** screen, configure these settings:
   * **Storage class for new objects:** *S3 Standard*
   * **Object metadata**:
     + *Guess MIME type*
     + *Gateway files acccessible to S3 bucket owner*
     + Enable *Requester Pays*
   * **Access your S3 bucket**: *Use an existing IAM role*
   * **IAM role:** Paste the *FgwIamPolicyARN*, which you can retrieve by following these instructions –
     + Choose the Details dropdown menu above these instructions
     + Select Show
     + Copy the *FgwIamPolicyARN* value
   * Choose **Next**
4. In the File access settings screen, accept the default settings.

**Note:** You might get a warning message that the file share is accessible from anywhere. For this lab, you can safely disregard this warning. In a production environment, you should always create policies that are as restrictive as possible to prevent unwanted or malicious connections to your instances.

* + Choose **Next**

1. Scroll to the bottom of the *Review and create* screen, then select **Create**

Monitor the status of the deployment and wait for **Status** to change to *Available*, which takes less than a minute.

**Note:** You can choose the refresh button occasionally to notice more quickly when the status has changed.

1. Select the file share that you just created by choosing the link.

1. At the bottom of the screen, note the command to mount the file share on Linux. You will need it for the next task.

**Task 5: Mounting the file share to the Linux instance and migrating the data**

Before you can migrate data to the NFS share that you created, you must first mount the share. In this task, you will mount the NFS share on a Linux server, then copy data to the share.

1. Connect to the **On-Prem Linux Server** instance.

**Microsoft Windows users**

These instructions are specifically for Microsoft Windows users. If you are using macOS or Linux, [skip to the next section](https://labs.vocareum.com/web/2656082/2537354.0/ASNLIB/public/docs/lang/en_us/README.html#ssh-MACLinux).

1. Above these instructions that you are currently reading, choose the Details dropdown menu, and then select Show

A **Credentials** window opens.

1. Choose the **Download PPK** button and save the **labsuser.ppk** file.

**Note:** Typically, your browser saves the file to the **Downloads** directory.

1. Note the **OnPremLinuxInstance** address, if it is displayed.

1. Exit the **Details** panel by choosing the **X**.

1. To use SSH to access the EC2 instance, you must use **PuTTY**. If you do not have PuTTY installed on your computer, [download PuTTY](https://the.earth.li/~sgtatham/putty/latest/w64/putty.exe).

1. Open **putty.exe**.

1. To keep the PuTTY session open for a longer period of time, configure the PuTTY timeout:
   * Choose **Connection**
   * **Seconds between keepalives**: 30

1. Configure your PuTTY session by using the following settings.
   * Choose **Session**
     + **Host Name (or IP address):** Paste the **OnPremLinuxInstance** for the instance you noted earlier
       - Alternatively, return to the Amazon EC2 console and choose **Instances**
       - Select the instance you want to connect to
       - In the *Description* tab, copy the **IPv4 Public IP** value
   * Back in PuTTY, in the **Connection** list, expand **SSH**
   * Choose **Auth** and expand **Credentials**
   * Under **Private key file for authentication:** Choose **Browse**
   * Browse to and select the **labsuser.ppk** file that you downloaded
   * To select the file, choose **Open**
   * Choose **Open** again
2. To trust and connect to the host, choose **Accept**.

1. When you are prompted with **login as**, enter: ec2-user

This action connects you to the EC2 instance.

1. Microsoft Windows users: [Click here to skip ahead to the next task.](https://labs.vocareum.com/web/2656082/2537354.0/ASNLIB/public/docs/lang/en_us/README.html#ssh-after)

**macOS  and Linux  Users**

These instructions are specifically for macOS or Linux users. If you are a Windows user, [skip ahead to the next task.](https://labs.vocareum.com/web/2656082/2537354.0/ASNLIB/public/docs/lang/en_us/README.html#ssh-after)

1. Above these instructions that you are currently reading, choose the Details dropdown menu, and then select Show

A **Credentials** window opens.

1. Choose the **Download PEM** button and save the **labsuser.pem** file.

1. Note the **OnPremLinuxInstance** address, if it is displayed.

1. Exit the **Details** panel by choosing the **X**.

1. Open a terminal window, and change directory to the directory where the *labsuser.pem* file was downloaded by using the cd command.

For example, if the *labsuser.pem* file was saved to your **Downloads** directory, run this command:

cd ~/Downloads

1. Change the permissions on the key to be read-only, by running this command:

chmod 400 labsuser.pem

1. Run the following command (replace **<public-ip>** with the **OnPremLinuxInstance** address that you copied earlier).
   * Alternatively, to find the IP address of the on-premises instance, return to the Amazon EC2 console and select **Instances**
   * Select the **On-Prem Linux Server** instance that you want to connect to
   * In the **Details** tab, copy the **Public IPv4 address** value

ssh -i labsuser.pem ec2-user@<public-ip>

2. When you are prompted to allow the first connection to this remote SSH server, enter yes.

Because you are using a key pair for authentication, you are not prompted for a password.

**You should now be connected to the instance.**

1. On the Linux instance, to view the data that exists on this server, enter the following command:

ls /media/data

You should see 20 image files in the .png format.

1. Create the directory that will be used to synchronize data with your S3 bucket by using the following command:

sudo mkdir -p /mnt/nfs/s3

1. Mount the file share on the Linux instance by using the command that you located in the Storage Gateway file shares details screen at the end of the last task.

sudo mount -t nfs -o nolock,hard <File-Gateway-appliance-private-IP-address>:/share /mnt/nfs/s3

Notice that the command starts with sudo and ends with /mnt/nfs/s3.

For example:

sudo mount -t nfs -o nolock,hard 10.10.1.33:/share /mnt/nfs/s3

1. Verify that the share was mounted correctly by entering the following command:

df -h

The output of the command should be similar to the following example:

[ec2-user@ip-10-10-1-210 ~]$ df -h

Filesystem                  Size  Used Avail Use% Mounted on

devtmpfs                    483M   64K  483M   1% /dev

tmpfs                       493M     0  493M   0% /dev/shm

/dev/xvda1                  7.8G  1.1G  6.6G  14% /

10.10.1.33:/share  8.0E     0  8.0E   0% /mnt/nfs/s3

1. Now that you created the mount point, you can copy the data that you want to migrate to Amazon S3 into the share by using this command:

cp -v /media/data/\*.png /mnt/nfs/s3

**Task 6: Verifying that the data is migrated**

You have finished configuring the gateway and copying data into the NFS share. Now, you will verify that the configuration works as intended.

1. In the **Services** search box, search for and choose **S3** to open the S3 console.

1. Select the bucket that you created in the *US East (Ohio)* Region.
   * Verify that the 20 image files are listed.

**Note:** You might need to choose the refresh icon in the S3 console.

2. Return to the **Buckets** page and select the bucket that you created in the *US West (Oregon)* Region.
   * Verify that the images files were replicated to this bucket, based on the policy that you created earlier.

**Note:** S3 Object replication can take up to 15 minutes to complete. Keep refreshing until you see the replicated objects.

Congratuations, you successfully migrated data to Amazon S3 by using AWS Storage Gateway in File Gateway mode! After your data is stored in Amazon S3, you can act on it like native Amazon S3 data. In this lab, you created a replication policy to copy the data to a secondary Region. You could also perform other operations, such as configuring a lifecycle policy. For example, you could migrate infrequently used data automatically from S3 Standard to Amazon Simple Storage Service Glacier for long-term storage, which can reduce costs.

**Submitting your work**

1. At the top of these instructions, choose **Submit** to record your progress and when prompted, choose **Yes**.

1. If the results don't display after a couple of minutes, return to the top of these instructions and choose Grades

**Tip**: You can submit your work multiple times. After you change your work, choose **Submit** again. Your last submission is what will be recorded for this lab.

1. To find detailed feedback on your work, choose Details followed by **View Submission Report**.

**Lab complete**

Congratulations! You have completed the lab.

1. Choose End Lab at the top of this page, and then select **Yes** to confirm that you want to end the lab.

A panel should appear with this message: *DELETE has been initiated... You may close this message box now.*​

1. Select the **X** in the top right corner to close the panel.

*©2022 Amazon Web Services, Inc. and its affiliates. All rights reserved. This work may not be reproduced or redistributed, in whole or in part, without prior written permission from Amazon Web Services, Inc. Commercial copying, lending, or selling is prohibited.*