Configuration Manual

MSc Research Project
MSc in Cloud Computing

Mihir Patel

Student ID: 23231408

School of Computing
National College of Ireland

Supervisor: Abubakr Siddig

Student Name:	Mihir Patel
Student ID:	23231408
Programme:	MSc in Cloud Computing
Year:	2024
Module:	MSc Research Project
Supervisor:	Abubakr Siddig
Submission Due Date:	12/12/2024
Project Title:	Configuration Manual
Word Count:	2273
Page Count:	8

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Mihir Patel
Date:	11th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Mihir Patel
23231408

1 Account Setup:

Step 1: Create AWS free tier account at https://signin.aws.amazon.com/signup?request$_t ype =$ *register*

Step 2: Create free Google Cloud account at https://cloud.google.com/free

2 Setup Storage Buckets

Step 3: To create the AWS S3 bucket follow the below steps

1. Login to the AWS account created in step 1 and navigate to S3 console in the us-east-1 region and click on create bucket

2. Keep Bucket Type as General purpose and enter bucket name

3. Select the ACLs disabled radio button under object ownership

4. Keep the Block all public access checkbox checked

5. Select enable under Bucket Versioning

6. Provide key-value pair under tags and do not change anything under Default Encryption and Advanced Settings and click on create bucket

Step 4: To create the Cloud Storage bucket in Google cloud follow below steps

1. Login to Google cloud account created in step 2 and navigate to Cloud Storage Dashboard and click on the create bucket button

2. Enter bucket name and click continue

3. Select Region radio button and select us-east1 under location type and click on continue

4. Select standard option under storage class type and click on continue

5. Keep default selection under control access to objects and click on continue

6. Keep default settings under protect object data section and click on create button

3 Setup DataSync Agent in AWS VPC and Google Cloud

Step 5: Configure HMAC key for the Google Cloud Storage bucket

1. Navigate to Service Account Dashboard and select the existing project

2. Enter the service account name and ID and click on create and continue

3. Select Storage object viewer as a IAM role and click on continue

4. Click on the Done button

5. Navigate to the Cloud Storage Settings page

6. Click on Interoperability tab and click on create key for service account

7. Select the service account created above and click on create key

8. Note down the secret key and access key

Step 6: To create vpc follow below steps

1. Navigate to VPC console in the us-east-1 region and click on create vpc button

2. Select vpc and more option under resources to create

3. Enter Name tag and keep the default values for IPv4 CIDR block and IPv6 CIDR block

4. Keep the tenancy as default

5. Keep the remaining options as default and click on create vpc button

Step 7: Create vpc service endpoint using below steps

1. Navigate to VPC console in the us-east-1 region and click on Endpoints from the left navigation panel

2. Click on create endpoint button

3. Enter relevant name in the Name tag field

4. Select AWS Services option under Type

5. Search for datasync in the Services and select 'com.amazonaws.us-east-1.datasync'

6. Select the vpc created in step 6 under vpc

7. Uncheck the 'Enable DNS name' checkbox under the Additional settings section

8. Select the enabled availability zone from the subnet section, us-east-1 in this case and select the subnet of the vpc created in step 5 from the subnetid dropdown

9. Click on the create endpoint button

Step 8: Configure Security group for the vpc as below

1. Navigate to vpc console and click on the Security Groups from the left navigation panel and click on create security group button

2. Enter security group name and description

3. Select vpc created in step 6 from the vpc dropdown

4. Add inbound rule by allowing All Traffic from source as Anywhere-IPv4

5. Add outbound rule by allowing All Traffic to destination as Anywhere-IPv4

6. Click on create security group button

7. Navigate to the vpc service endpoint created in step 7 and modify the security group to add the above created security grouo to the servcice endpoint

Step 9: Create EC2 instance to host the agent

1. Open AWS Cloudshell and run the 'aws ssm get-parameter –name /aws/service/datasync/ami –region vpc-region' command, replace vpc-region with us-east-1 and notedown the value property from the response

2. Launch the url 'https://console.aws.amazon.com/ec2/v2/home?region=vpc-regionLaunchInstance id' in new tab, replace vpc-region with us-east-1 and ami-id with the value propeerty from the above step

3. Select the m5.2xlarge instance type which is supported for the datasync agent

4. Select existing or new key-value pair

5. Choose vpc and subnet configured in step 6 under netwrok settings

6. Enable the auto-assign public-ip and click on launch instance, instance should be in running state

Step 10: Activate the agent

1. Navigate to DataSync console and click on agents fro the left navigation panel

2. Click on create agent button

3. Select Amazon EC2 from the hypervisor dropdown

4. Select 'VPC endpoints using AWS PrivateLink' from the endpoint type dropdown

5. Choose the vpc endpoint created in step 7 from vpc endpoint dropdown

6. Choose the subnet for the vpc created in step 6

7. Choose the security group created in step 8

8. Enter the public ipv4 of the EC2 instance and click on Get Key, activation key should be retrieved successfully

9. Enter agent name and clic on create agent

Step 11: Configure the agent in the Google Cloud

1. Navigate to Google Cloud CLI installer and install it in local machine

2. Navigate to DataSync console and click on agents from the left navigation panel

3. Click on create agent button

4. Choose VMware ESXi for hypervisor and download the agent image

5. Open terminal in local machine and move to the folder where agent image is kept and run the 'unzip AWS-DataSync-Agent-VMWare.zip' command

6. Run 'tar -xvf aws-datasync-2.0.1655755445.1-x86_64.xfs.gpt.ova' to extract the content of .ova file, replace the aws-datasync-2.0.1655755445.1-x86_64.xfs.gpt.ova with the file name unzipped in the machine

7. Open Googlwe Cloud CLI and run the 'gcloud compute images import aws-datasync-2-test –source-file INCOMPLETE-aws-datasync-2.0.1655755445.1-x86_64.xfs.gpt-disk1.vmdk –os centos-7' to import the agent's .vmdk file into google cloud, replace the .vmdk file name with appropriate file name

8. Navigate to Compute Engine dashboard and click on create instance button

9. Select us-east1 region and e2-standard-8 as a machine type

10. Navigate to custom images section under the OS ans Storage tab and select the above imported agent image and select the service account created in step 5

11. After the instance is running, note down the ip of the instance

12. Navigate back to the create agent screen in the AWS DataSync console

13. For endpoint type select Public service endpoints in US(N.Virginia)

14. Enter ip of the VM instance and click on Get Key, key should get activated successfully

15. Enter the agent name and click on create agent

4   Setup DataSync Transfer Tasks

Step 12: Configure the Google Cloud Storage bucket location for the DataSync Transfer tasks with agent deployed in AWS VPC

1. Navigate to DataSync console and click on Locations from the left navigation panel

2. Click on create location

3. Choose Object Storage location tyoe

4. Choose the agent created in step 10

5. Enter 'storage.googleapis.com' in the server field

6. Enter the name of the Cloud Storage bucket in the bucket name field

7. Expand Additional Settings and select server protocol HTTPS and enter server port 443

8. Under the authentication section enter the HMAC access key and secret key created in the step 5 and click on create location

Step 13: Configure the AWS S3 bucket location for the DataSync Transfer tasks

1. Navigate to DataSync console and click on Locations from the left navigation panel

2. Click on create location

3. Choose AWS S3 location tyoe

4. Enter the S3 URI of the S3 bucket created in step 3 or browse the S3 bucket

5. Select Standard under S3 storage class type

6. Use the autogenerated IAM role

7. Click on create location

Step 14: Create DataSync transfer task to migrate data from Google Cloud Storage to AWS S3 with agent deployed in AWS VPC

1. Navigate to DataSync console and click on the 'Tasks' from the left navigation panel

2. Click on create task button

3. Select choose existing location as source location and select the Google Cloud Storage location created in the step 12 in the us-east-1 region and click on Next

4. Select choose existing location for destination location and select the AWS S3 bucket location setup in step 13 and click on Next

5. Enter the Name and keep the task mode as Basic

6. Select Everything in the content to scan field

7. Select Transfer all data for transfer mode and select Verify all data for Verification field

8. Keep the schedule frequency and Not scheduled

9. Use the auto-generated CloudWatch log group and click on Next, verify the review page and click on create task button

Step 15: Create DataSync transfer task to migrate data from AWS S3 to Google Cloud Storage with agent deployed in AWS VPC

1. Navigate to DataSync console and click on the 'Tasks' from the left navigation panel

2. Click on create task button

3. Select choose existing location as source location and select the AWS S3 location created in the step 13 in the us-east-1 region and click on Next

4. Select choose existing location for destination location and select the Google Cloud Storage bucket location setup in step 12 and click on Next

5. Enter the Name and keep the task mode as Basic

6. Select Everything in the content to scan field

7. Select Transfer all data for transfer mode and select Verify all data for Verification field

8. Keep the schedule frequency and Not scheduled

9. Use the auto-generated CloudWatch log group and click on Next, verify the review page and click on create task button

Step 16: Configure the Google Cloud Storage bucket location for the DataSync Transfer tasks with agent deployed in google cloud

1. Navigate to DataSync console and click on Locations from the left navigation panel

2. Click on create location

3. Choose Object Storage location tyoe

4. Choose the agent created in step 11

5. Enter 'storage.googleapis.com' in the server field

6. Enter the name of the Cloud Storage bucket in the bucket name field

7. Expand Additional Settings and select server protocol HTTPS and enter server port 443

8. Under the authentication section enter the HMAC access key and secret key created in the step 5 and click on create location

Step 17: Create DataSync transfer task to migrate data from Google Cloud Storage to AWS S3 with agent deployed in google cloud

1. Navigate to DataSync console and click on the 'Tasks' from the left navigation panel

2. Click on create task button

3. Select choose existing location as source location and select the Google Cloud Storage location created in the step 16 in the us-east-1 region and click on Next

4. Select choose existing location for destination location and select the AWS S3 bucket location setup in step 13 and click on Next

5. Enter the Name and keep the task mode as Basic

6. Select Everything in the content to scan field

7. Select Transfer all data for transfer mode and select Verify all data for Verification field

8. Keep the schedule frequency and Not scheduled

9. Use the auto-generated CloudWatch log group and click on Next, verify the review page and click on create task button

Step 18: Create DataSync transfer task to migrate data from AWS S3 to Google Cloud Storage with agent deployed in google cloud

1. Navigate to DataSync console and click on the 'Tasks' from the left navigation panel

2. Click on create task button

3. Select choose existing location as source location and select the AWS S3 location created in the step 13 in the us-east-1 region and click on Next

4. Select choose existing location for destination location and select the Google Cloud Storage bucket location setup in step 16 and click on Next

5. Enter the Name and keep the task mode as Basic

6. Select Everything in the content to scan field

7. Select Transfer all data for transfer mode and select Verify all data for Verification field

8. Keep the schedule frequency and Not scheduled

9. Use the auto-generated CloudWatch log group and click on Next, verify the review page and click on create task button

5   Setup Storage Trabsfer Service task

Step 19: Create AWS IAM role for Storage Transfer job to authenticate with AWS S3 bucket

1. Navigate to IAM console and click on roles.

2. Click on create role

3. Select custom trust policy for trusted entity type

4. Follow the steps under Authenticate Using Federated Identity from the Storage trasfer service guide and after IAM role creation note down the ARN of the role

Step 20: Configure the transfer job in Storage Transfer Service to migrate data from AWS S3 to Google Cloud Storage

1. Navigate to Storage Transfer Dashboard and click on create transfer job button

2. Choose source type as AWS S3 and destination type as Google Cloud Storage and click on Next Step

3. Enter the S3 bucket's name created in step 3 in the Bucket field of source section

4. Choose AWS IAM role for identity federation

5. Enter ARN of AWS IAM role created in step 19 in AWS IAM role ARN field and click on next step

6. Browse Cloud Storage bucket created in step 4 and select it as a destination location and click on next step

7. Select Run on demand for job to run

8. Set Storage class to Standard for Storage class field

9. Choose if different option for when to overwrite field

10. Choose never for when to delete field

11. Enable Cloud logging and click on create button