Contents

[**Giai đoạn 1: Bảo mật dữ liệu trong Amazon S3** 2](#_Toc202377558)

[**Nhiệm vụ 1.1: Tạo một thùng, áp dụng chính sách xô và truy cập kiểm tra** 2](#_Toc202377559)

[**Bước 1. Truy cập S3 Console** 2](#_Toc202377560)

[**Bước 2: Xác định unique-ID** 2](#_Toc202377561)

[**Bước 3: Tạo bucket mới** 3](#_Toc202377562)

[**Bước 4: Tải file myfile.txt lên** 3](#_Toc202377563)

[**Bước 5: Áp dụng Bucket Policy** 4](#_Toc202377564)

[**Bước 6 Đăng nhập với Paulo để kiểm tra quyền truy cập** 6](#_Toc202377565)

[**Bước 7: Kiểm tra quyền của Mary** 7](#_Toc202377566)

[**Nhiệm vụ 1.2: Bật versioning và object-level logging cho bucket** 9](#_Toc202377567)

[**Bước 1: Đăng nhập đúng với IAM role Voclabs như yêu cầu** 9](#_Toc202377568)

[**Bước 2: Bật Versioning cho bucket data-bucket** 9](#_Toc202377569)

[**Bước 3: Bật Server Access Logging** 10](#_Toc202377570)

[**Bước 4: Kiểm tra Bucket Policy của s3-objects-access-log** 11](#_Toc202377571)

[**Nhiệm vụ 1.3: Bật tính năng S3 Inventory cho bucket data-bucket** 11](#_Toc202377572)

[**Bước 1 Vào tab Management** 11](#_Toc202377573)

[**Bước 2. Ở phần Create inventory configuration** 12](#_Toc202377574)

[**Nhiệm vụ 1.4: Xác nhận rằng tính năng versioning hoạt động như mong muốn.** 13](#_Toc202377575)

[**Bước 1 Tạo file customers.csv** 13](#_Toc202377576)

[**Bước 2: Đăng nhập AWS bằng user paulo và upload file** 14](#_Toc202377577)

[**Bước 3: Kiểm tra Versioning** 14](#_Toc202377578)

[**Bước 4: Chỉnh sửa và upload phiên bản mới** 15](#_Toc202377579)

[**Bước 5: Kiểm tra lại các phiên bản** 15](#_Toc202377580)

[**Bước 6: Đăng nhập bằng user mary để test quyền truy cập** 15](#_Toc202377581)

[**Nhiệm vụ 1.5: Kiểm tra object-level logging và dùng Athena để truy vấn log** 16](#_Toc202377582)

[**Bước 1: Kiểm tra log đã ghi vào S3** 16](#_Toc202377583)

[**Bước 2: Tạo Athena table để dễ phân tích** 16](#_Toc202377584)

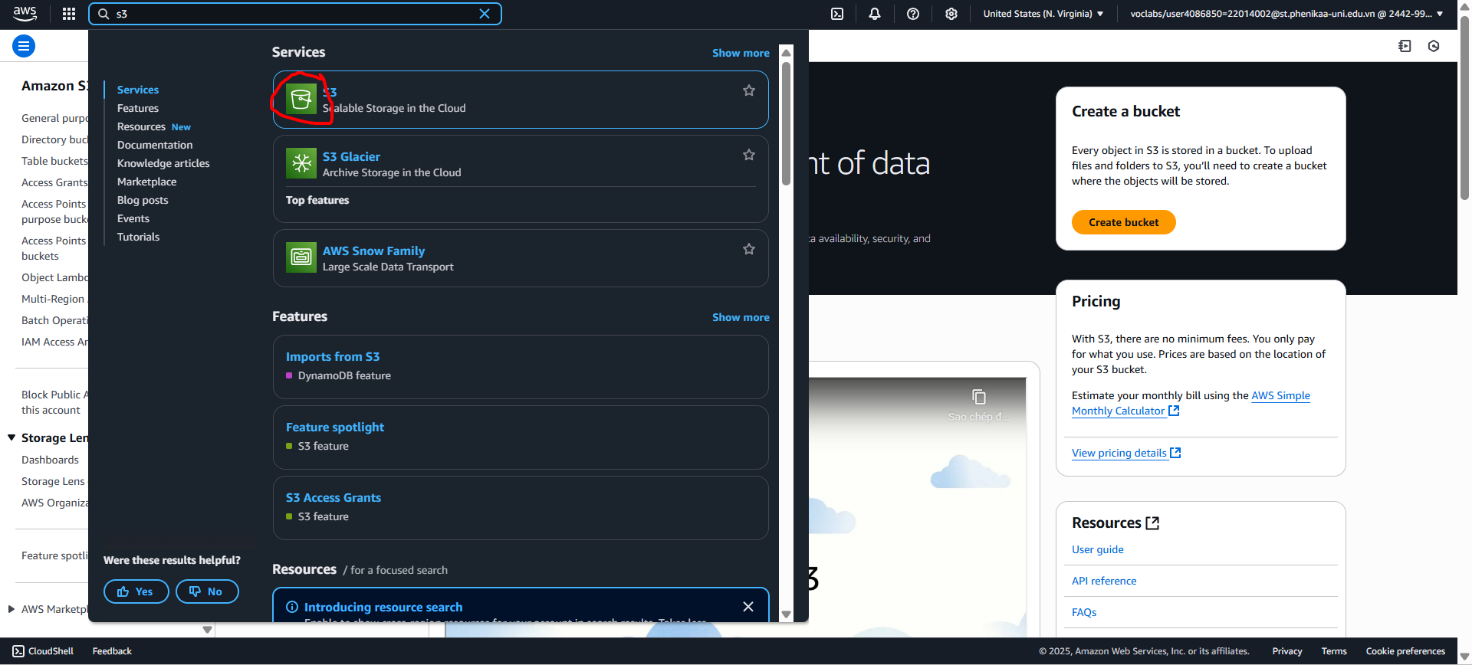
# 

# **Giai đoạn 1: Bảo mật dữ liệu trong Amazon S3**

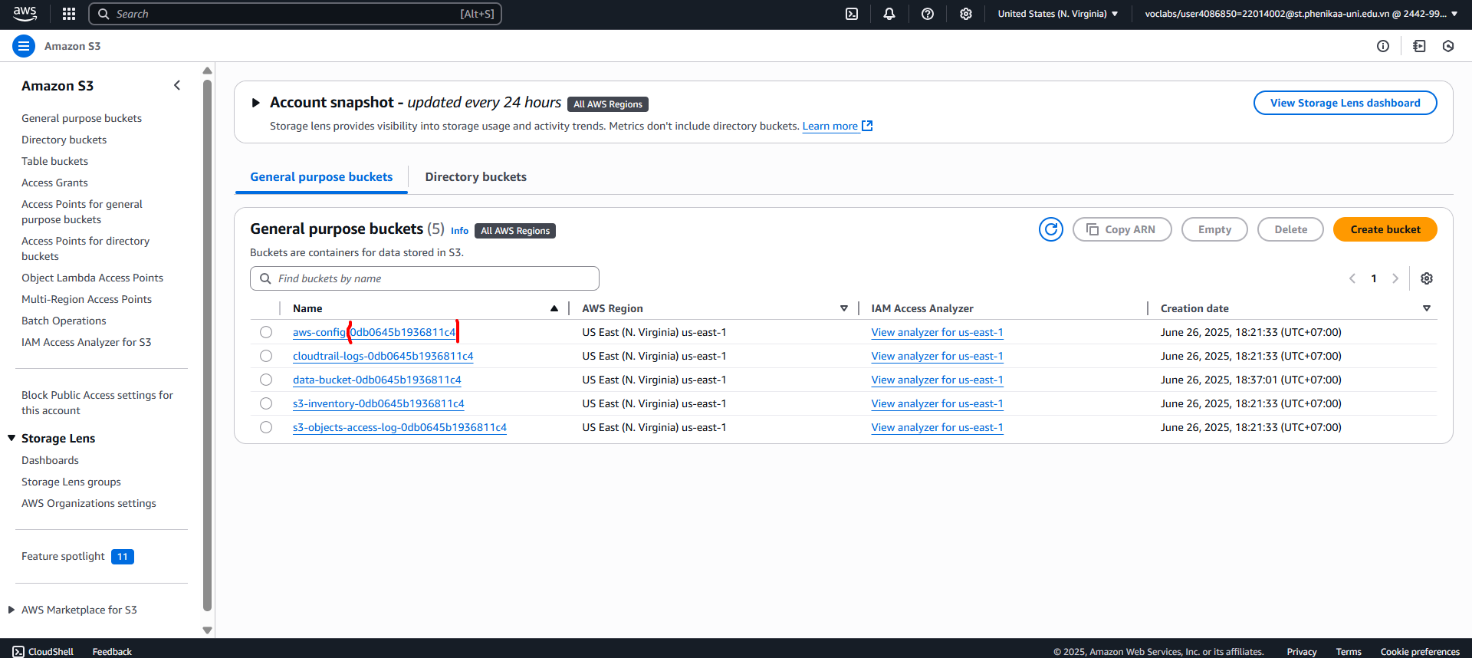
*Trong giai đoạn này, bạn được thử thách để bắt đầu triển khai cài đặt bảo mật trong tài khoản AWS. Bạn đã được yêu cầu bảo mật dữ liệu PII của khách hàng được lưu trữ trong Amazon S3. Đội ngũ lãnh đạo của AnyCompany Financial đã nghe về các vi phạm dữ liệu gần đây tại các công ty khác và muốn bảo vệ dữ liệu khách hàng khỏi truy cập trái phép. Công ty muốn làm như sau: Hạn chế quyền truy cập vào xô cho một số người quản lý tài khoản nhất định, những người trong nhóm người quản lý tài khoản. Bật phiên bản trên các thùng S3 và tất cả các đối tượng trong đó. Bật đăng nhập đối tượng trong tất cả các thùng S3. Mã hóa tất cả các nhóm bằng cách sử dụng mã hóa phía máy chủ với các khóa được quản lý của Amazon S3 (SSE-S3). Thực hiện kho hàng Amazon S3 để giữ kho lưu trữ của tất cả các tệp được lưu trữ trong Amazon S3. Đến cuối giai đoạn này, bạn sẽ có kiến ​​trúc được hiển thị trong sơ đồ sau:*

## **Nhiệm vụ 1.1: Tạo một thùng, áp dụng chính sách xô và truy cập kiểm tra**

### **Bước 1. Truy cập S3 Console**



### **Bước 2: Xác định unique-ID**



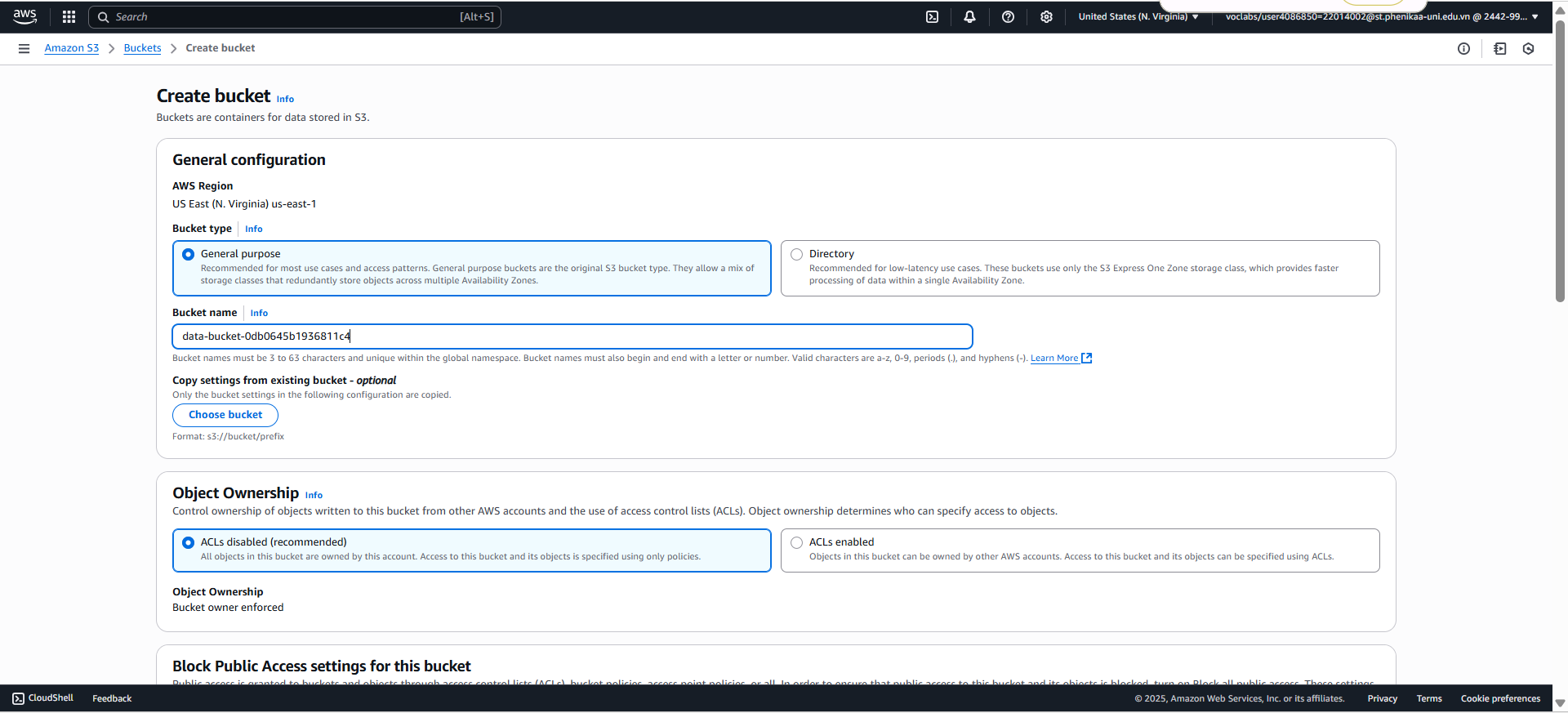
### **Bước 3: Tạo bucket mới**

Nhấn **Create bucket**

Đặt tên bucket:  
👉 data-bucket-0db0645b1936811c4

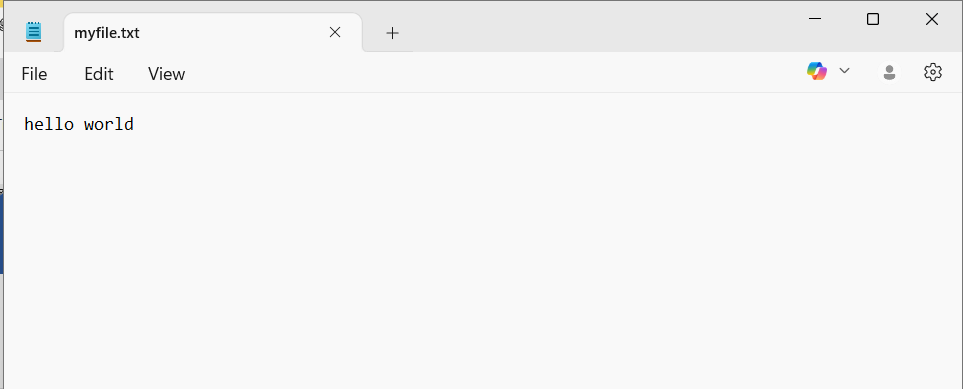
Region: us-east-1

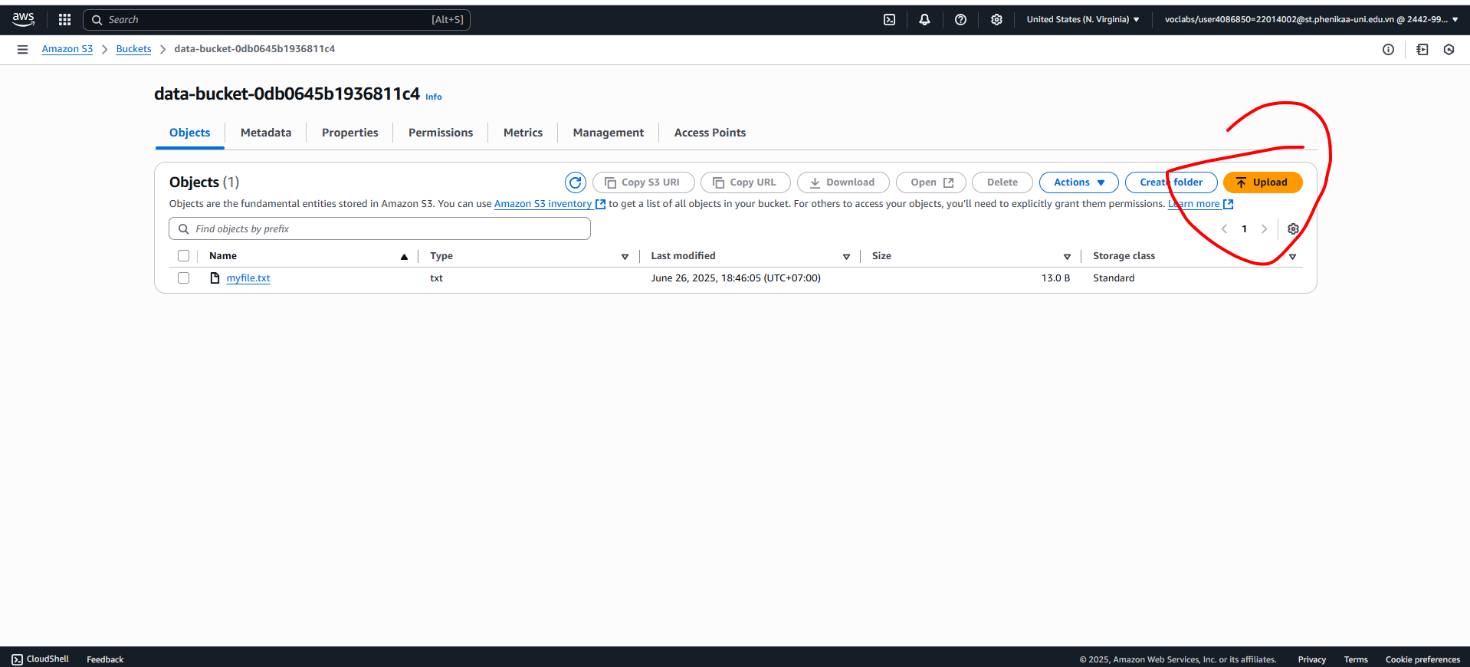
Giữ tất cả cài đặt mặc định → **Create bucket**



### **Bước 4: Tải file myfile.txt lên**

* Tạo file myfile.txt với nội dung: hello world
* Vào bucket data-bucket-0db0645b1936811c4 → Chọn **Upload**
* Chọn file myfile.txt từ máy → Nhấn **Upload**





### **Bước 5: Áp dụng Bucket Policy**

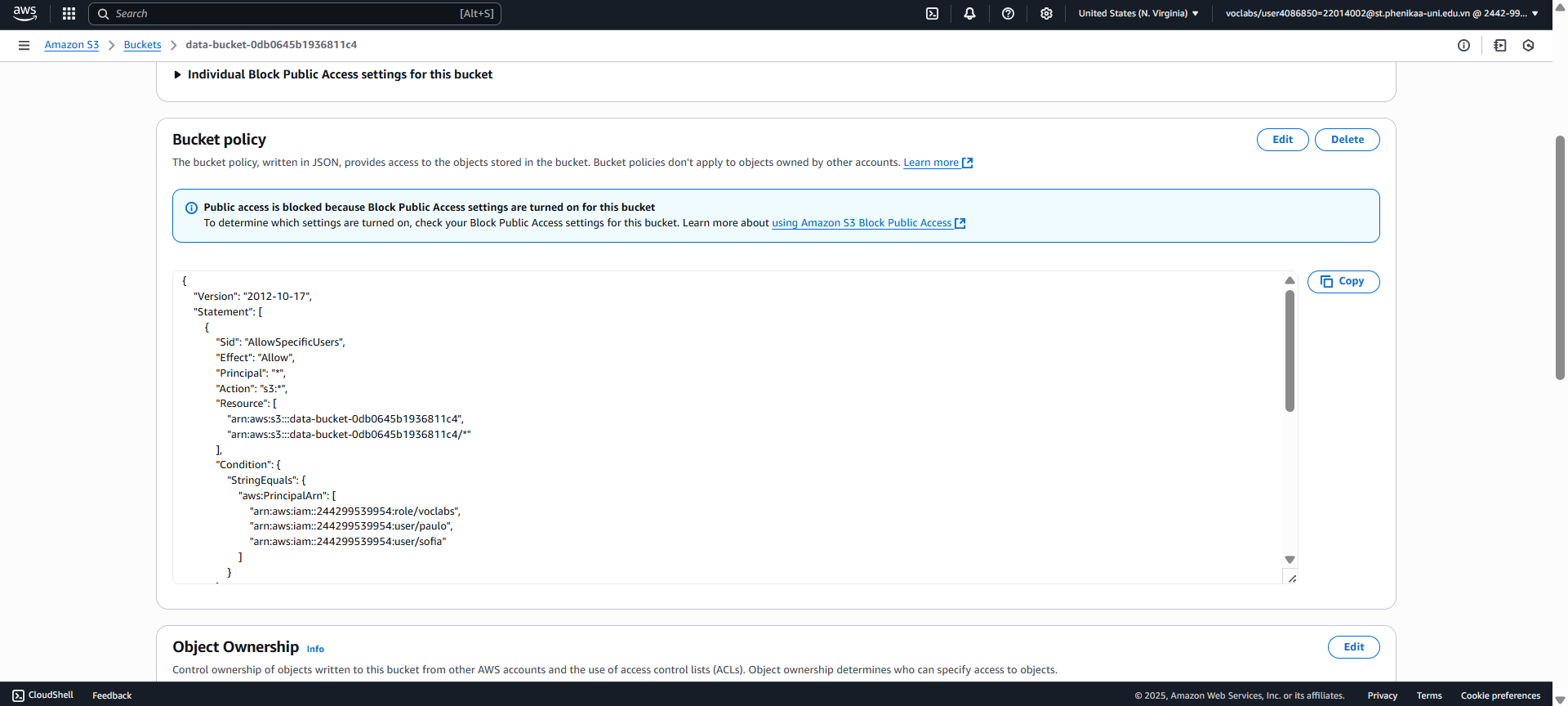
 Quay lại bucket data-bucket-0db0645b1936811c4

 Chọn tab **Permissions**

 Kéo xuống phần **Bucket policy**

 Dán đoạn JSON trên vào

 Nhấn **Save**



{

"Version": "2012-10-17",

"Statement": [

{

"Sid": "AllowSpecificUsers",

"Effect": "Allow",

"Principal": "\*",

"Action": "s3:\*",

"Resource": [

"arn:aws:s3:::data-bucket-0db0645b1936811c4",

"arn:aws:s3:::data-bucket-0db0645b1936811c4/\*"

],

"Condition": {

"StringEquals": {

"aws:PrincipalArn": [

"arn:aws:iam::244299237547:role/voclabs",

"arn:aws:iam::244299237547:user/paulo",

"arn:aws:iam::244299237547:user/sofia"

]

}

}

},

{

"Sid": "DenyOthers",

"Effect": "Deny",

"Principal": "\*",

"Action": "s3:\*",

"Resource": [

"arn:aws:s3:::data-bucket-0db0645b1936811c4",

"arn:aws:s3:::data-bucket-0db0645b1936811c4/\*"

],

"Condition": {

"StringNotEquals": {

"aws:PrincipalArn": [

"arn:aws:iam::244299237547:role/voclabs",

"arn:aws:iam::244299237547:user/paulo",

"arn:aws:iam::244299237547:user/sofia"

]

}

}

}

]

}

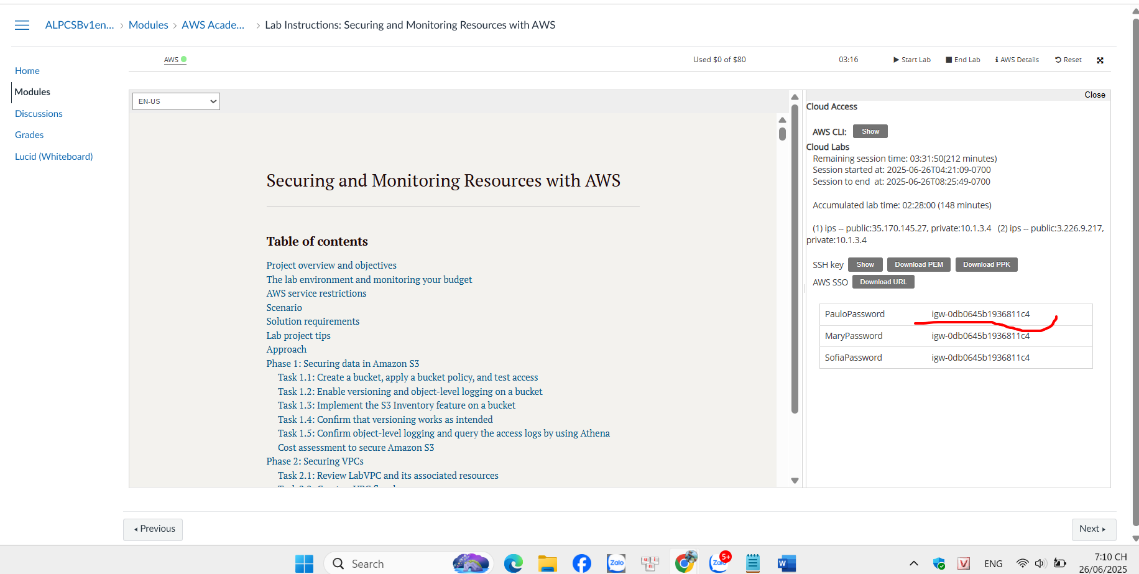
### **Bước 6 Đăng nhập với Paulo để kiểm tra quyền truy cập**

1. **Mở cửa sổ trình duyệt ẩn danh (Incognito Window)**.
2. Truy cập đường dẫn đăng nhập IAM:

https://244299539954.signin.aws.amazon.com/console

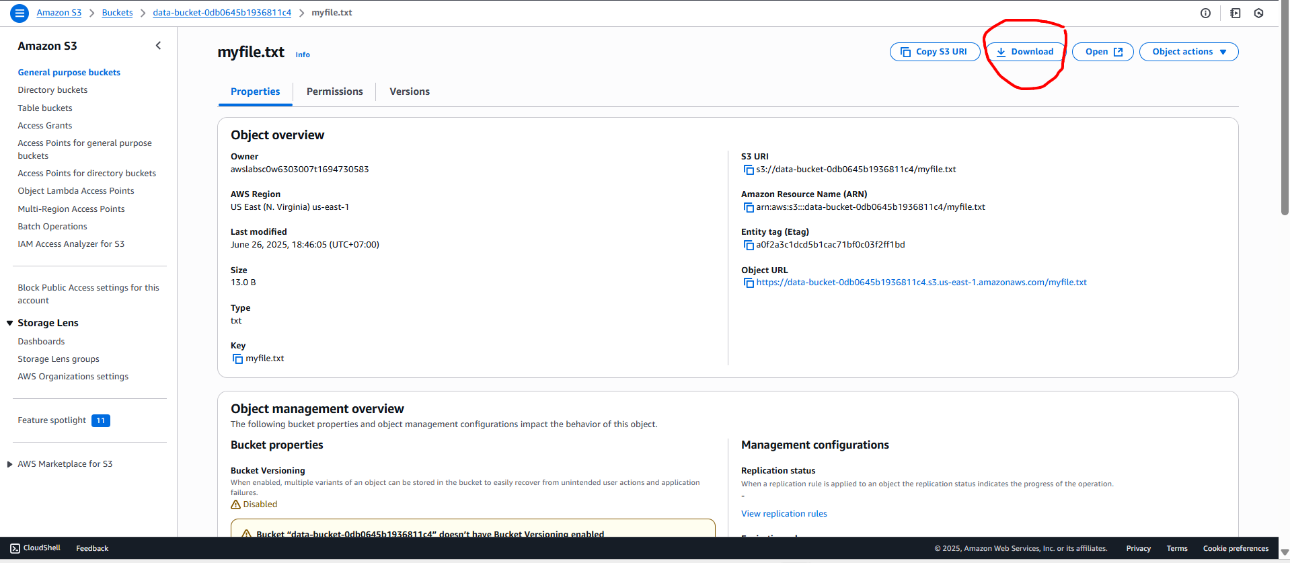
*(Nếu không biết link chính xác, bạn có thể tìm trong IAM Console khi đang đăng nhập với vai trò voclabs)*

1. **Đăng nhập với user Paulo**
   * Username: paulo
   * Password: có thể được cấp trong phần **AWS Details** trên giao diện lab

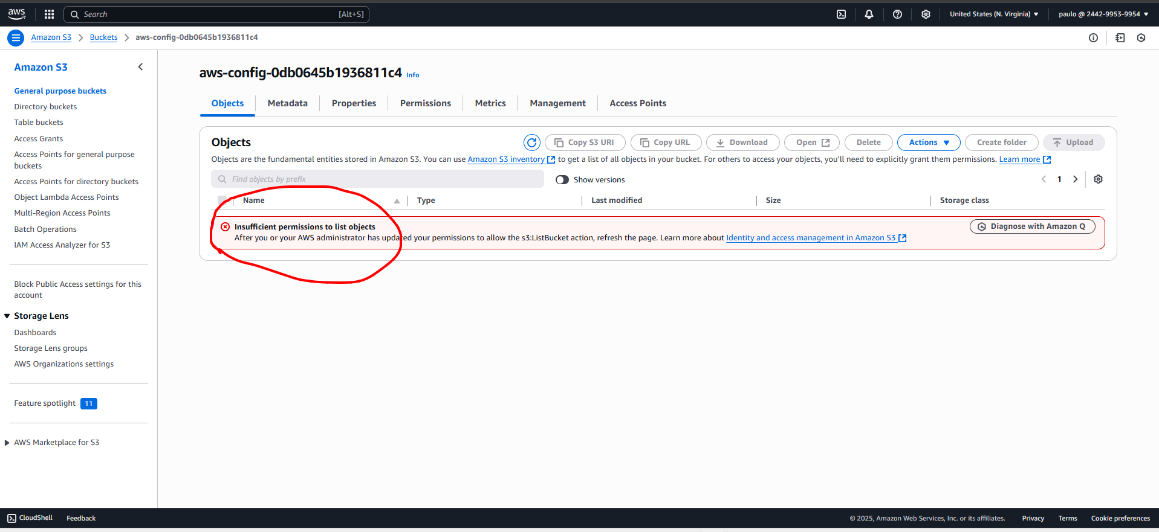


**✅ Kiểm tra sau khi đăng nhập:**

1. Vào **S3 Console tương tự như bước 1**
2. Kiểm tra:
   * Có thấy được bucket data-bucket-0db0645b1936811c4 không?
   * Có tải được file myfile.txt về không?



* + Không truy cập được các bucket khác (ví dụ như cloudtrail-logs-...)



### **Bước 7: Kiểm tra quyền của Mary**

**Mục tiêu:**

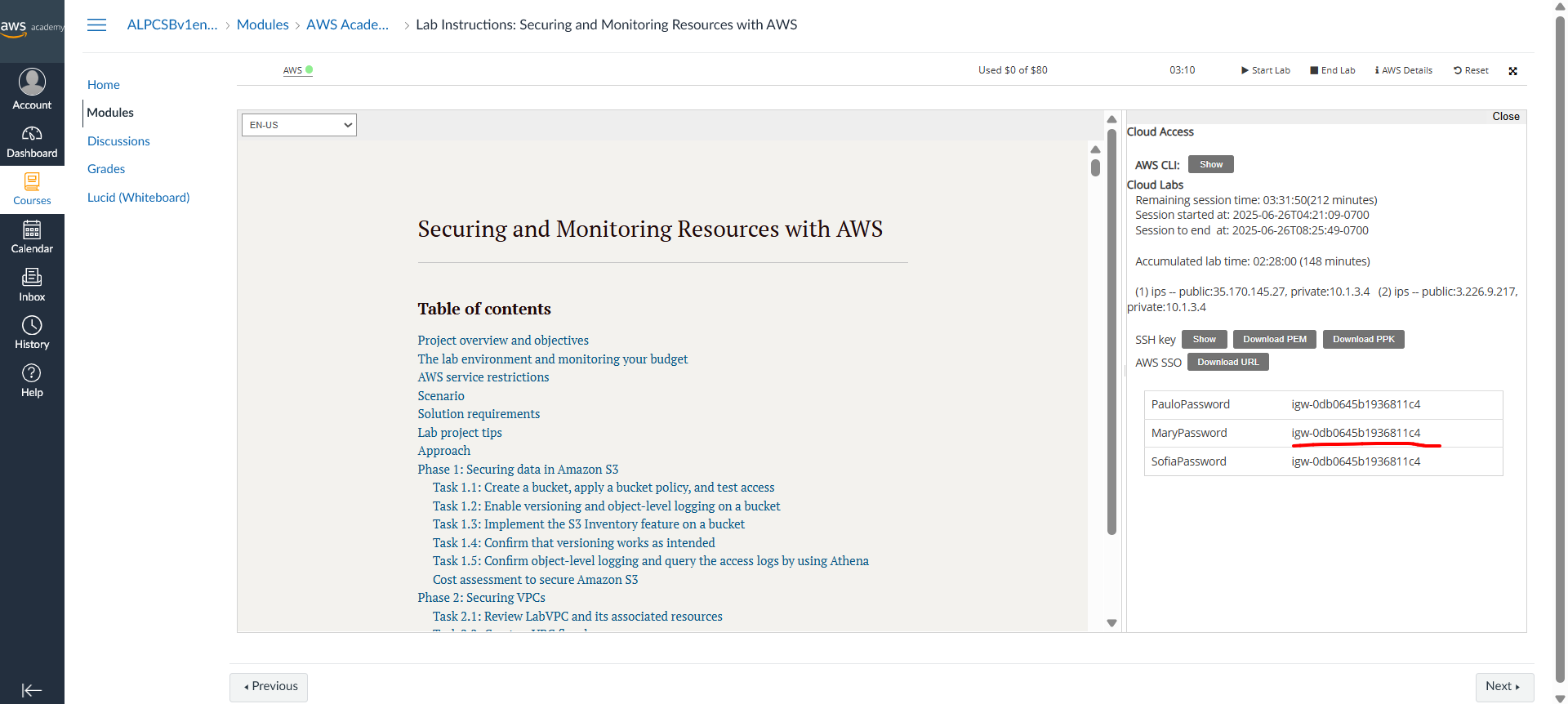
* Mary **thấy lỗi "Access: Error"** với data-bucket
* Các bucket khác hiển thị “Buckets and objects not public”
* Khi click vào bất kỳ bucket nào → **lỗi Insufficient permissions**

**Các bước thực hiện:**

1. **Đăng xuất khỏi Paulo**
   * Trên góc trên bên phải → chọn **Sign Out**
2. **Đăng nhập lại trong cùng cửa sổ ẩn danh**

https://244299539954.signin.aws.amazon.com/console

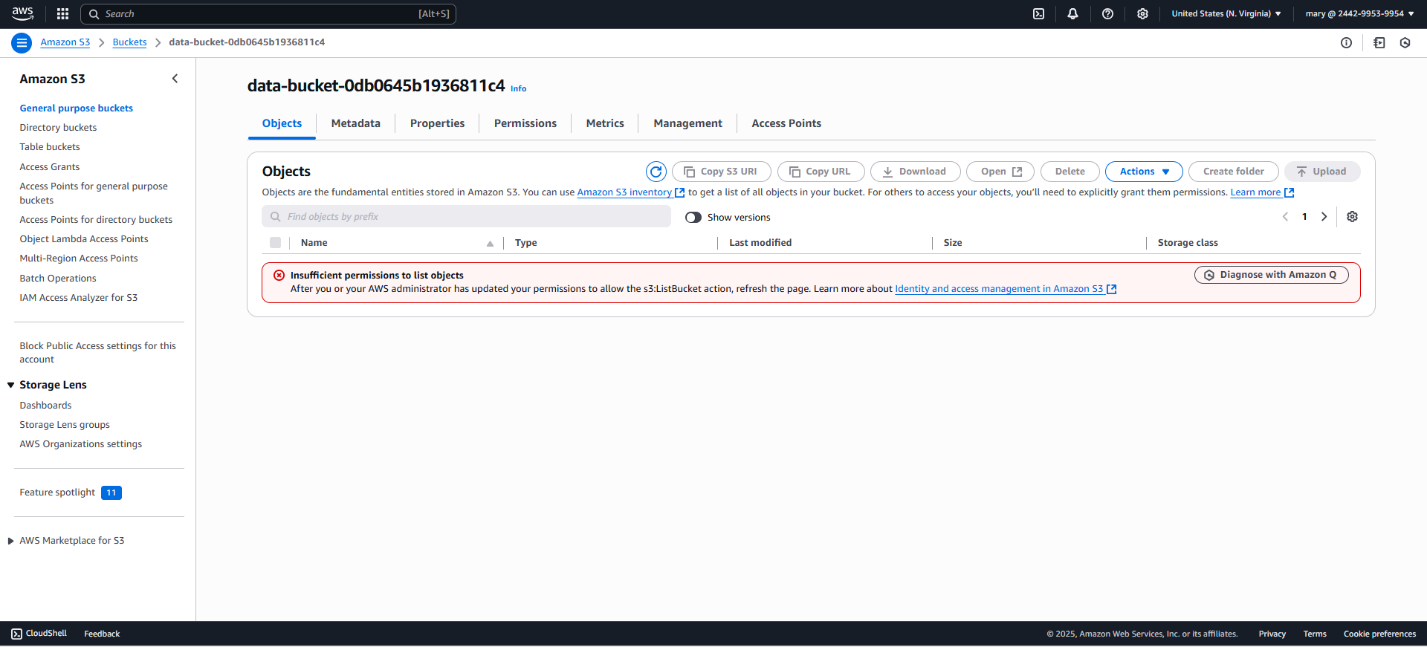
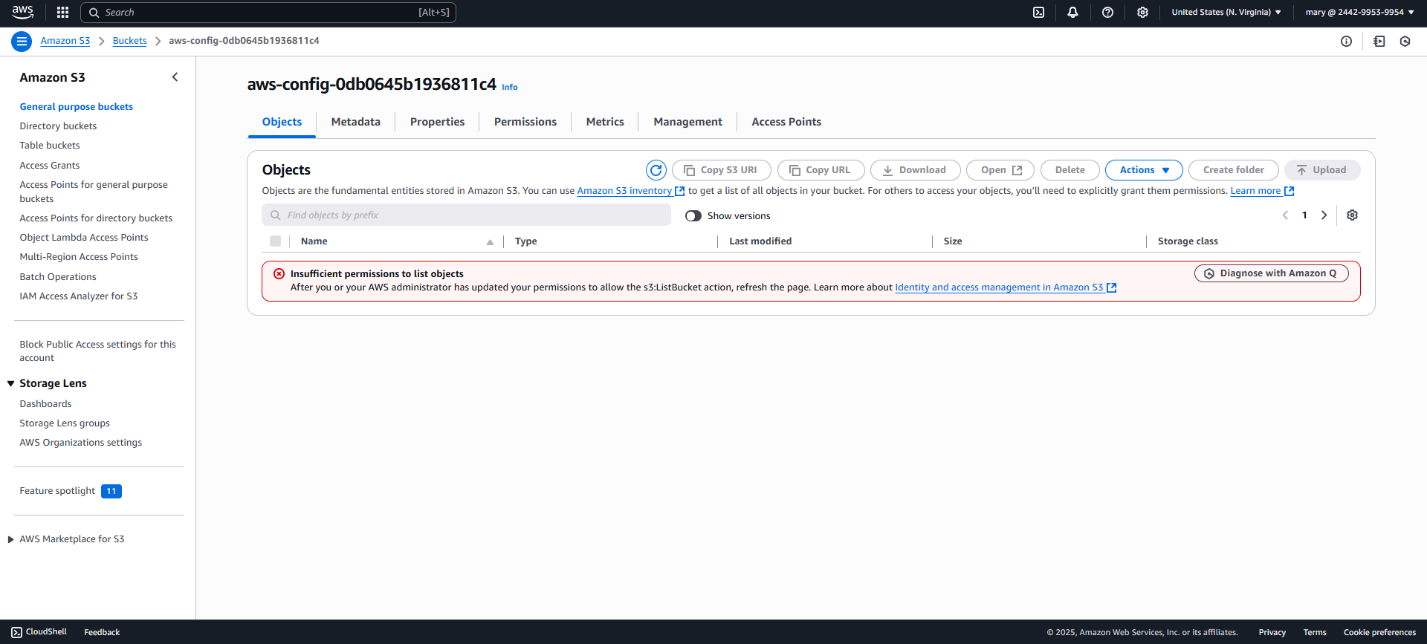
* + Sử dụng:
    - Username: mary
    - Password: từ phần **AWS Details** (giống như lúc dùng cho Paulo)



1. Vào lại **S3 Console**
   * Kiểm tra danh sách bucket

Kết quả mong đợi

| **Bucket** | **Trạng thái hiển thị với Mary** |
| --- | --- |
| data-bucket-0db0645b1936811c4 | ❌ Access: Error |
| Các bucket khác | ⚠️ Buckets and objects not public |
| Khi click vào bất kỳ bucket | ❌ Insufficient permissions to list objects |

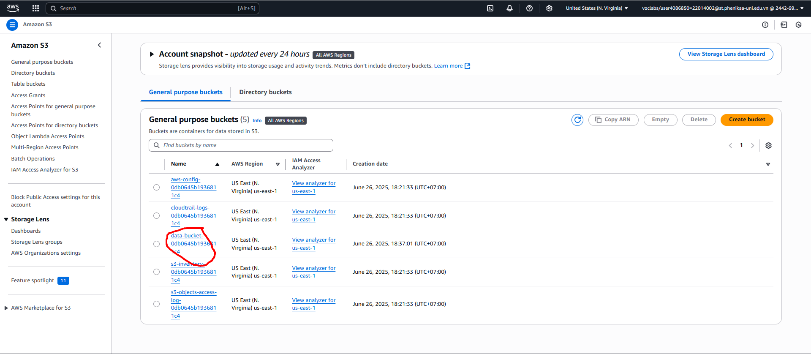


## **Nhiệm vụ 1.2: Bật versioning và object-level logging cho bucket**

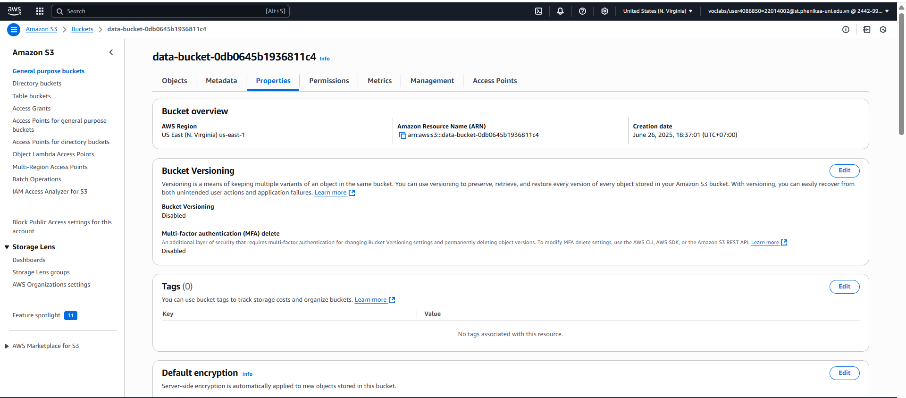
### **Bước 1: Đăng nhập đúng với IAM role Voclabs như yêu cầu**

### **Bước 2: Bật Versioning cho bucket data-bucket**

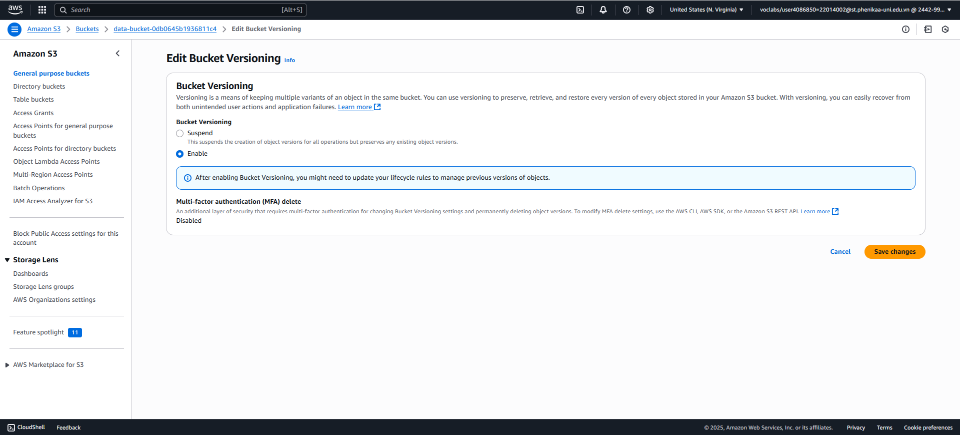
1. Vào **S3** trong AWS Console.



1. Chọn bucket có tên **data-bucket**.
2. Chuyển đến tab **"Properties"**.



1. Tìm phần **"Bucket Versioning"**, nhấn **"Edit"**.
2. Chọn **Enable** → Nhấn **Save changes**.



### **Bước 3: Bật Server Access Logging**

 Vẫn trong bucket data-bucket, tab **Properties**.

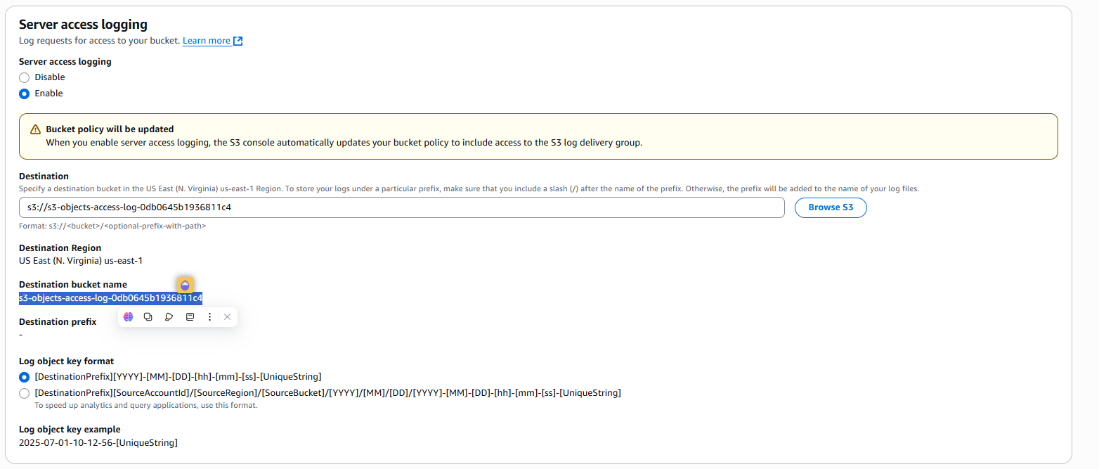
 Cuộn xuống phần **"Server access logging"** → Nhấn **"Edit"**.

 Chọn **Enable**.

 Trong phần **Target bucket**, chọn bucket đích là: s3-objects-access-log.

 (Chú ý: dấu gạch chéo / ở cuối để phân loại thư mục)

 Nhấn **Save changes**.

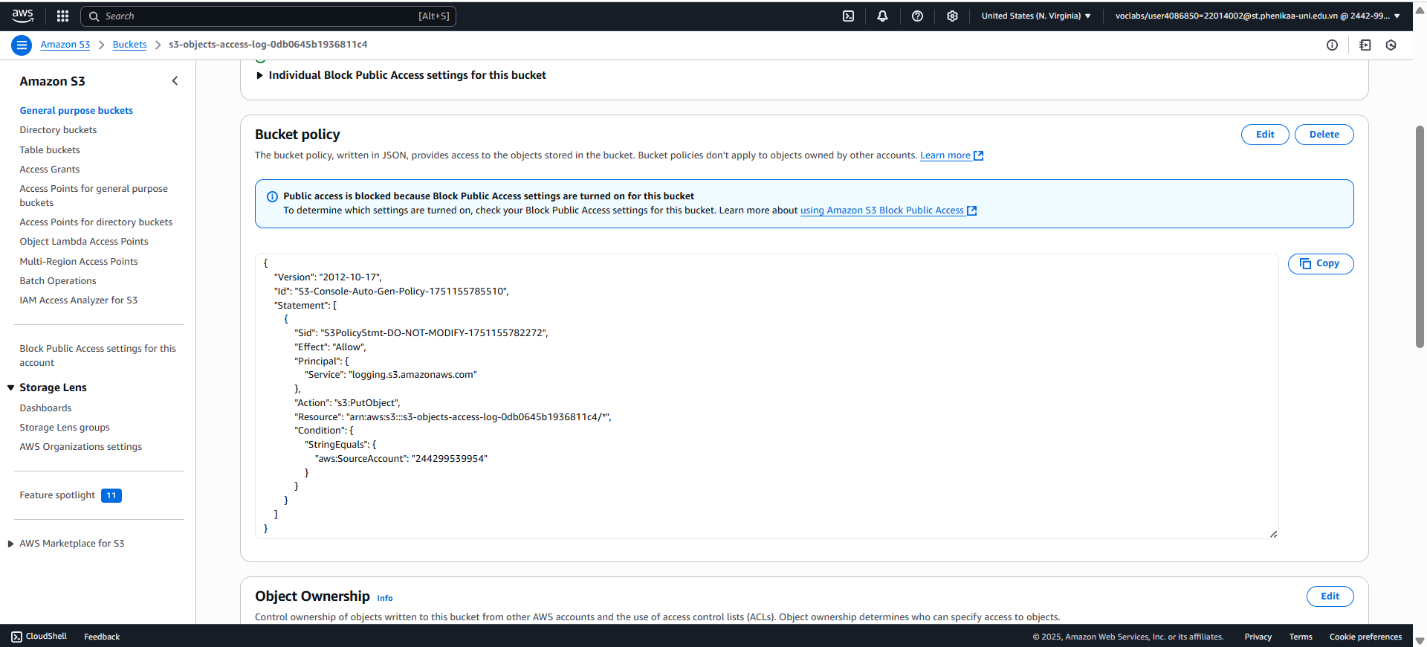


### **Bước 4: Kiểm tra Bucket Policy của s3-objects-access-log**

 Vào lại **bucket s3-objects-access-log**.

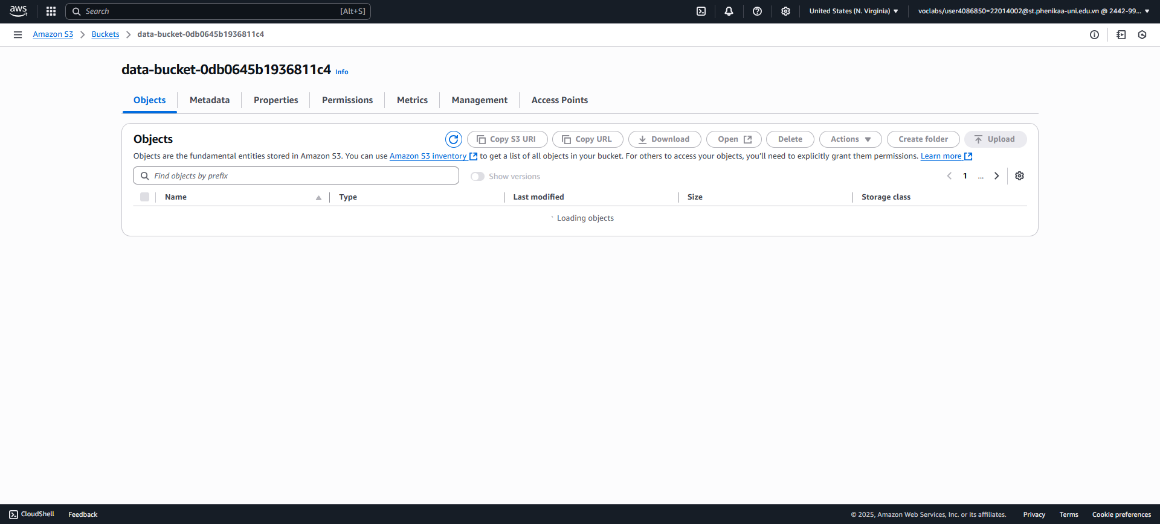
 Chuyển đến tab **Permissions** > **Bucket Policy**.

 Kiểm tra xem có đoạn policy nào giống như sau không:



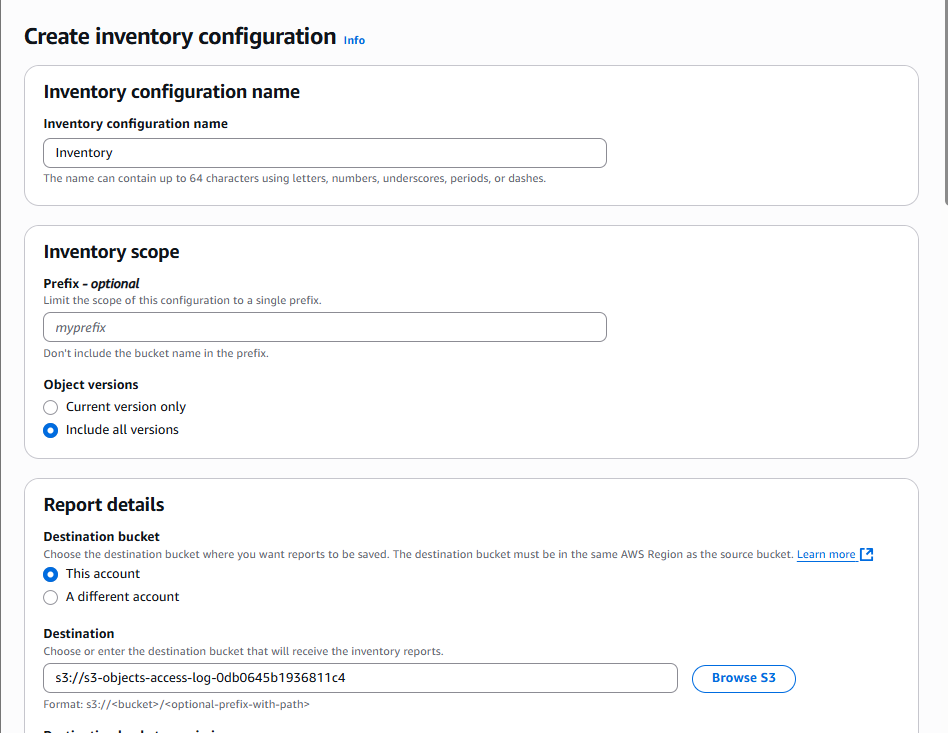
## **Nhiệm vụ 1.3: Bật tính năng S3 Inventory cho bucket data-bucket**

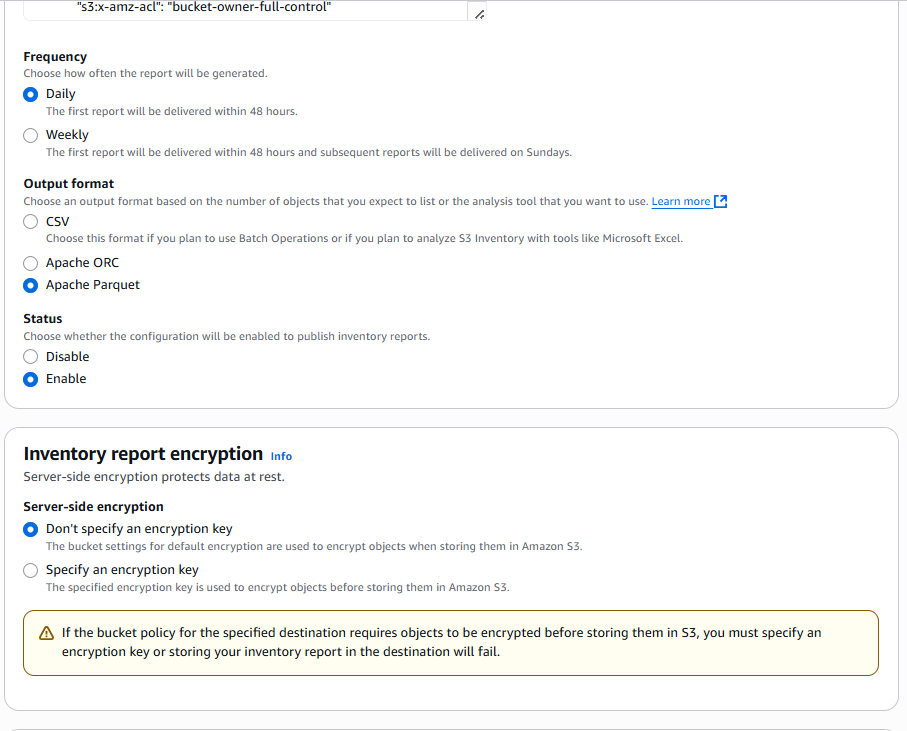
### **Bước 1 Vào tab Management**

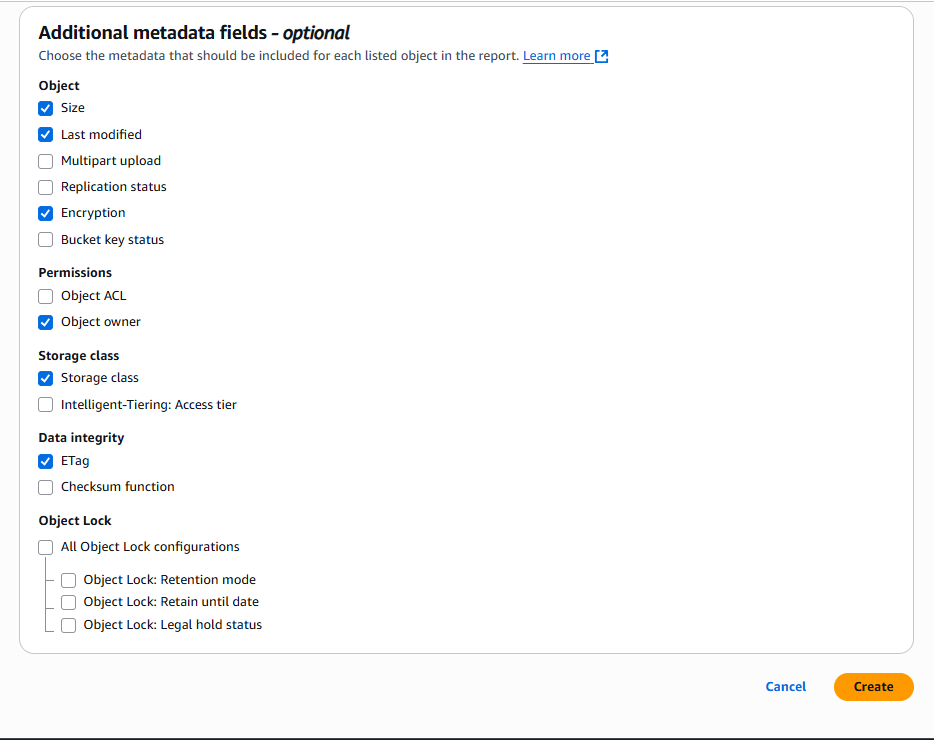


1. Truy cập **S3 Console**
2. Chọn bucket **data-bucket**
3. Chuyển sang tab **Management**
4. Ở mục **Inventory configurations**, nhấn **Create inventory configuration**

### **Bước 2. Ở phần Create inventory configuration**







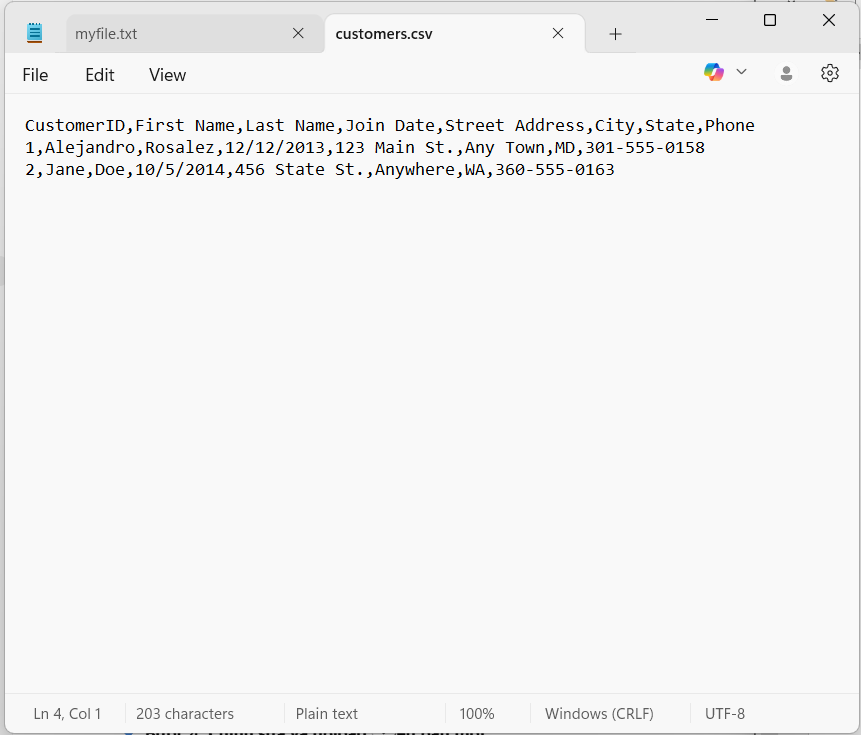
## **Nhiệm vụ 1.4: Xác nhận rằng tính năng versioning hoạt động như mong muốn.**

### **Bước 1 Tạo file customers.csv**

CustomerID,First Name,Last Name,Join Date,Street Address,City,State,Phone

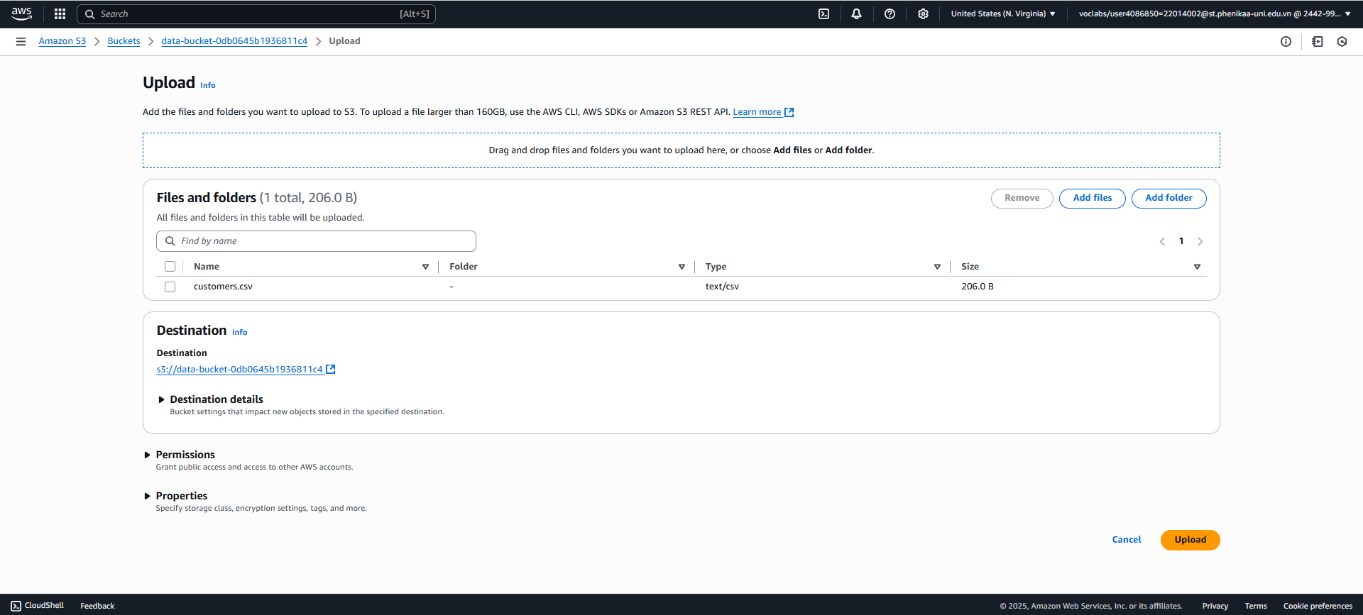
1,Alejandro,Rosalez,12/12/2013,123 Main St.,Any Town,MD,301-555-0158

2,Jane,Doe,10/5/2014,456 State St.,Anywhere,WA,360-555-0163

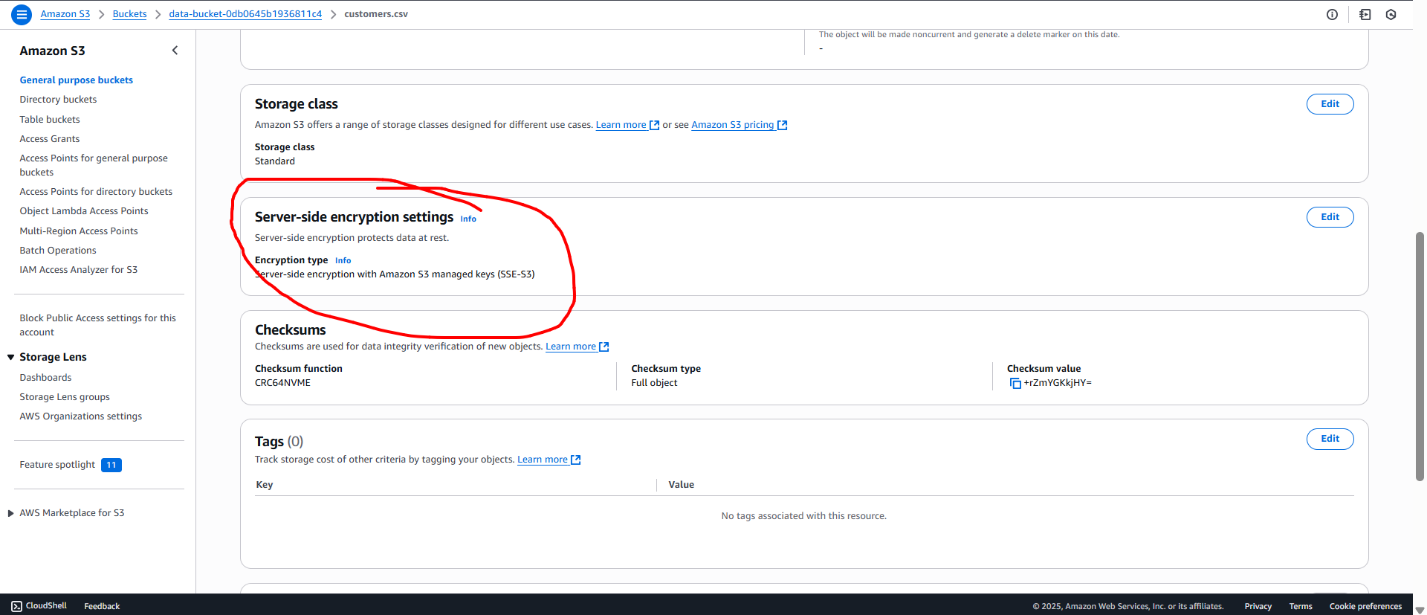


### **Bước 2: Đăng nhập AWS bằng user paulo và upload file**

1. Mở trình duyệt riêng tư (Incognito)
2. Đăng nhập bằng tài khoản AWS với user **paulo ( Như bước 6 của task 1.1)**
3. Vào S3 Console → mở bucket **data-bucket**
4. Nhấn **Upload** → Chọn file customers.csv vừa tạo



1. Trong phần "Properties" của file sau khi upload xong:
   * Kiểm tra encryption: phải hiển thị là **SSE-S3 (Amazon S3 managed keys)**



### **Bước 3: Kiểm tra Versioning**

1. Trong data-bucket, tìm file customers.csv
2. Click vào file → Chọn tab **Versions**
3. Lúc này bạn sẽ thấy **1 phiên bản** (hiện tại)

### **Bước 4: Chỉnh sửa và upload phiên bản mới**

1. Mở lại file customers.csv trên máy
2. Thêm **2 dòng sau** vào cuối file:

*3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122*

*4,Li,Juan,6/29/2011,1323 22nd Ave.,Anytown,NY,914-555-0149*

1. Lưu lại và **upload lần 2** vào data-bucket (đè lên file cũ)

3,John,Stiles,9/20/2016,1980 8th St.,Nowhere,NY,914-555-0122

4,Li,Juan,6/29/2011,1323 22nd Ave.,Anytown,NY,914-555-0149

### **Bước 5: Kiểm tra lại các phiên bản**

1. Mở tab **Versions** của customers.csv lại
2. Lúc này sẽ thấy có **2 phiên bản**:
   * Một bản 2 dòng
   * Một bản 4 dòng
3. Click mở từng phiên bản:
   * Phiên bản mới → kiểm tra thấy đủ **4 dòng**
   * Phiên bản cũ → kiểm tra chỉ có **2 dòng**

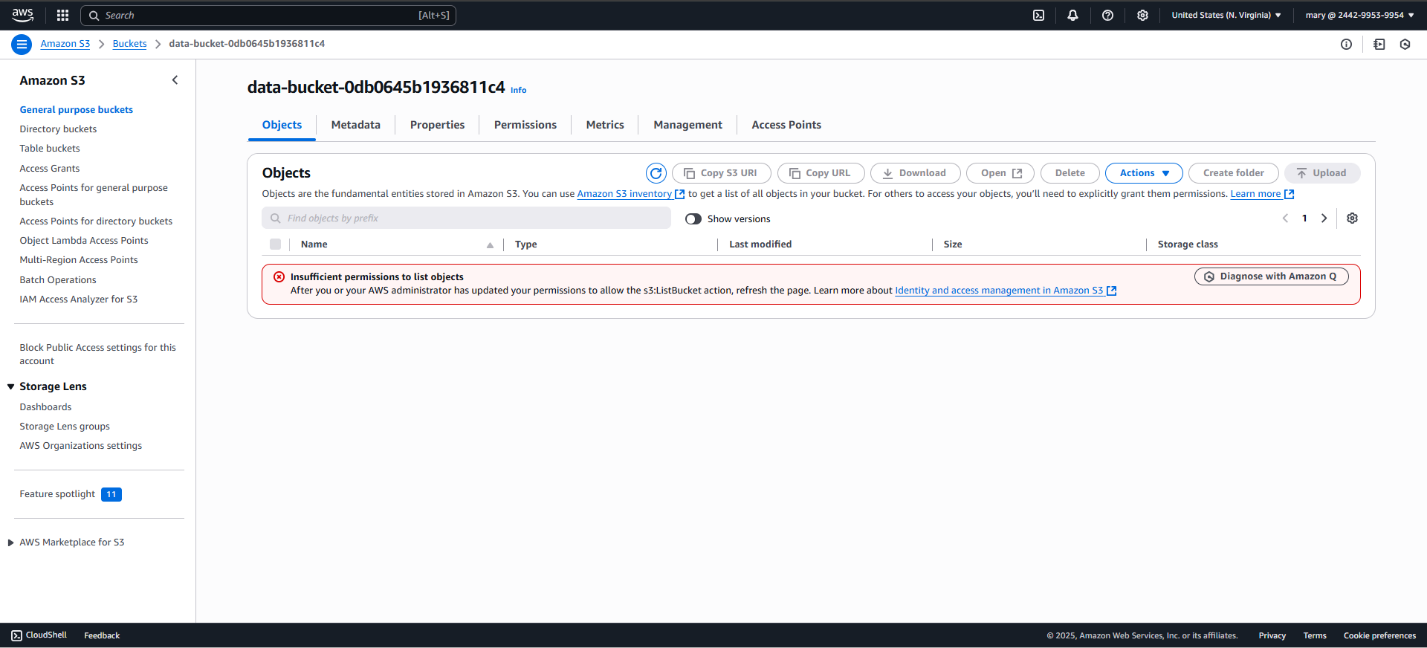
✅ Như vậy là **versioning hoạt động đúng**!

### **Bước 6: Đăng nhập bằng user mary để test quyền truy cập**

1. **Đăng xuất user paulo** và đóng cửa sổ incognito
2. Mở cửa sổ incognito mới, đăng nhập bằng user **mary (Bước 7 task 1.1)**
3. Vào S3 console → thử truy cập bucket data-bucket

⛔️ **Kết quả mong đợi**:

* Bạn **không truy cập được** bucket hoặc file trong đó
* Hành động này sẽ được ghi lại trong **object-level log**



## **Nhiệm vụ 1.5: Kiểm tra object-level logging và dùng Athena để truy vấn log**

### **Bước 1: Kiểm tra log đã ghi vào S3**

1. Vào bucket s3-objects-access-log-<ID>
2. Kiểm tra trong mục Objects đã có file log chưa

❗ Nếu chưa có, đợi khoảng 10–15 phút

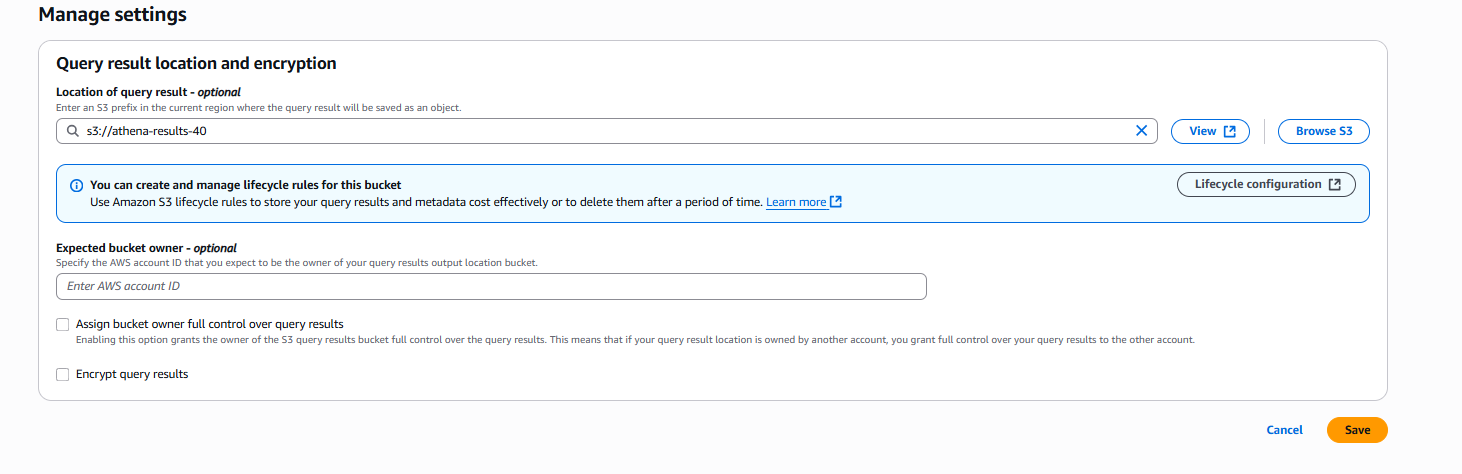
1. Tải file log .log về và mở ra, bạn sẽ thấy nội dung dạng dòng text ghi lại từng truy cập.

### **Bước 2: Tạo Athena table để dễ phân tích**

1. Tạo bucket mới đặt tên theo dạng:

athena-results-<random-number> ( ví dụ athena-results-40)

1. Vào dịch vụ **Athena** > chọn **Query Editor**
   * Chọn Set up query result location và dán đường dẫn bucket mới tạo



1. Dán đoạn SQL sau vào editor:

*CREATE EXTERNAL TABLE `default.bucket\_logs`(*

*`bucketowner` STRING,*

*`bucket\_name` STRING,*

*`requestdatetime` STRING,*

*`remoteip` STRING,*

*`requester` STRING,*

*`requestid` STRING,*

*`operation` STRING,*

*`key` STRING,*

*`request\_uri` STRING,*

*`httpstatus` STRING,*

*`errorcode` STRING,*

*`bytessent` BIGINT,*

*`objectsize` BIGINT,*

*`totaltime` STRING,*

*`turnaroundtime` STRING,*

*`referrer` STRING,*

*`useragent` STRING,*

*`versionid` STRING,*

*`hostid` STRING,*

*`sigv` STRING,*

*`ciphersuite` STRING,*

*`authtype` STRING,*

*`endpoint` STRING,*

*`tlsversion` STRING,*

*`accesspointarn` STRING,*

*`aclrequired` STRING)*

*ROW FORMAT SERDE*

*'org.apache.hadoop.hive.serde2.RegexSerDe'*

*WITH SERDEPROPERTIES (*

*'input.regex'='([^ ]\*) ([^ ]\*) \\[(.\*?)\\] ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) (\"[^\"]\*\"|-) (-|[0-9]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) (\"[^\"]\*\"|-) ([^ ]\*)(?: ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*) ([^ ]\*))?.\*$')*

*STORED AS INPUTFORMAT*

*'org.apache.hadoop.mapred.TextInputFormat'*

*OUTPUTFORMAT*

*'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'*

*LOCATION*

*'s3://s3-objects-access-log-<UNIQUE-ID>/'*

👉 **Chỉnh lại <UNIQUE-ID>** cho đúng với tên bucket log bạn có.

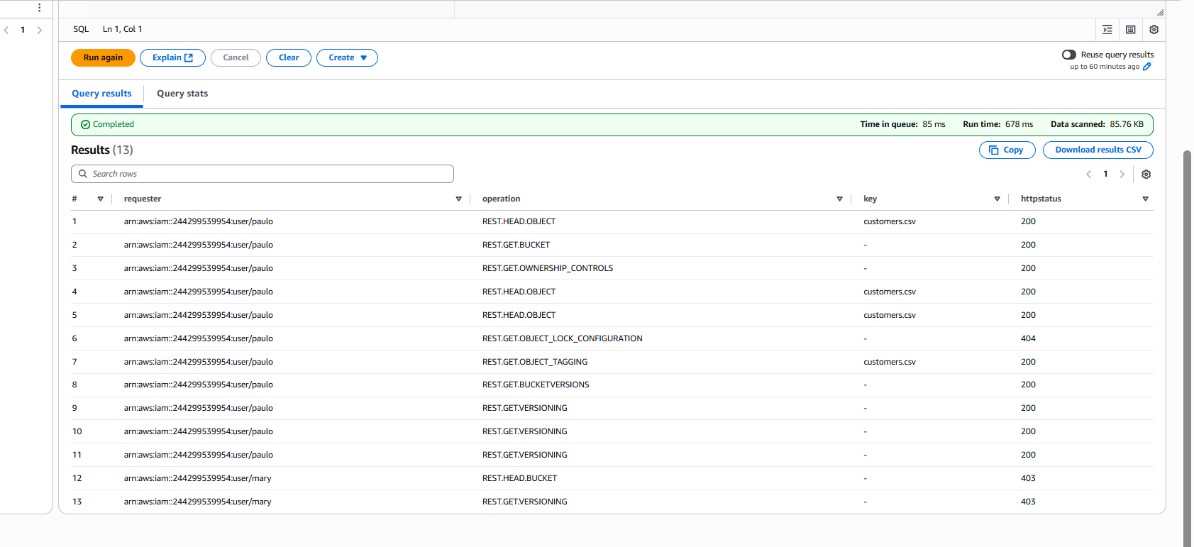
4. Dán câu truy vấn sau vào Query Editor

*SELECT requester, operation, key, httpstatus*

*FROM "default"."bucket\_logs"*

*WHERE requester LIKE 'arn:aws:iam%';*

**Ra kết quả**

****