Virus Scanning in AWS S3 with Hash Checking and VirusTotal Integration

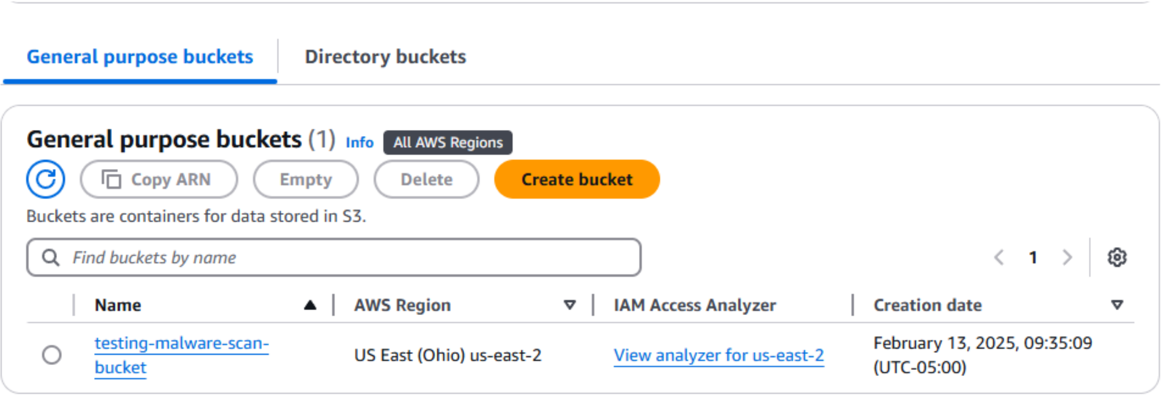
Introduction

The objective of this project is to implement a script that checks the hash of files uploaded to an S3 bucket against the VirusTotal API. This ensures that potentially malicious files can be identified and handled appropriately. For this, I utilized various AWS functionalities, including Lambda functions, S3 buckets, and GuardDuty.

Implementation

Step 1: Setting Up the S3 Bucket

- Created an S3 bucket named testing-malware-scan-bucket.
- Uploaded multiple test files to simulate a real-world scenario.

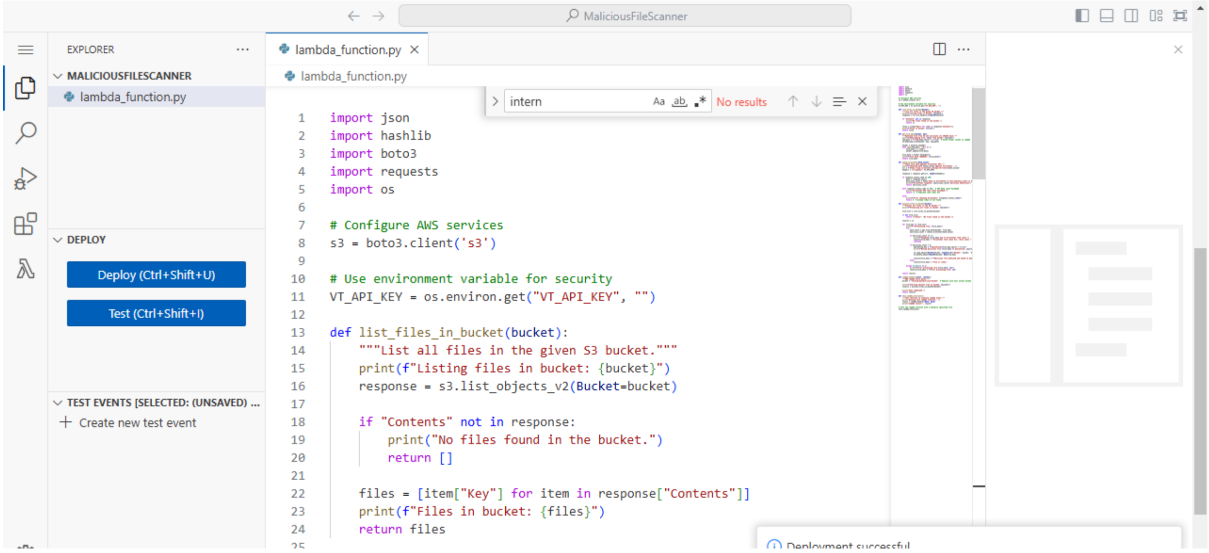


Buckets are containers for data stored in S3.

	Name	AWS Region	IAM Access Analyzer	Creation date
○	testing-malware-scan-bucket	US East (Ohio) us-east-2	View analyzer for us-east-2	February 13, 2025, 09:35:09 (UTC-05:00)

Step 2: Writing the Lambda Function

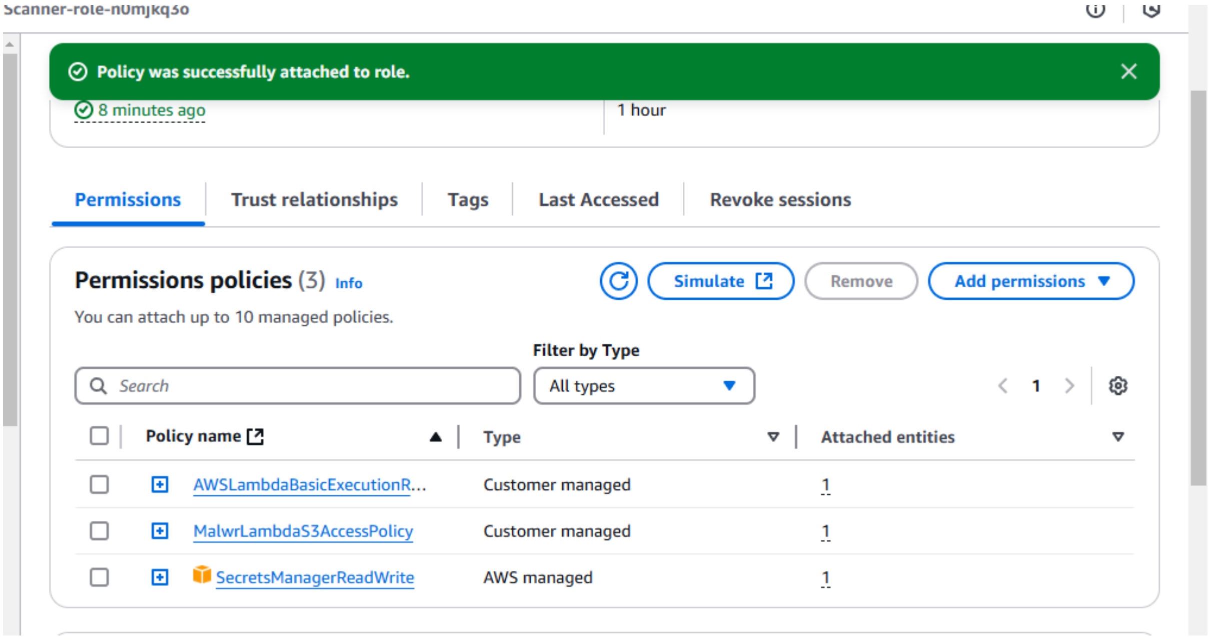
- Developed a Python script to:
  - List all files in the S3 bucket.
  - Compute the SHA-256 hash of each file.
  - Query the VirusTotal API to check if the file is flagged as malicious.
- Handled API rate limits and logging mechanisms for better tracking.



```python
import json
import hashlib
import boto3
import requests
import os

# Configure AWS services
s3 = boto3.client('s3')

# Use environment variable for security
VT_API_KEY = os.environ.get("VT_API_KEY", "")

def list_files_in_bucket(bucket):
    """List all files in the given S3 bucket."""
    print(f"Listing files in bucket: {bucket}")
    response = s3.list_objects_v2(Bucket=bucket)

    if "Contents" not in response:
        print("No files found in the bucket.")
        return []

    files = [item["Key"] for item in response["Contents"]]
    print(f"Files in bucket: {files}")
    return files
```

Step 3: Configuring IAM Permissions

- Ensured that the Lambda function had proper permissions to:
    - Read objects from the S3 bucket.
    - Write logs to CloudWatch.
- Adjusted IAM roles to avoid 403 Forbidden errors when accessing S3 objects.



Policy was successfully attached to role.

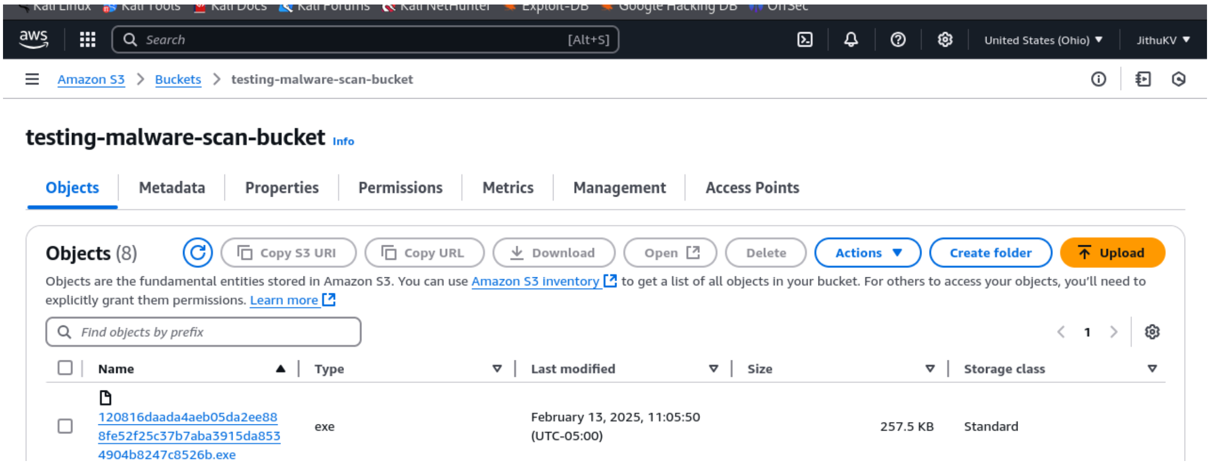
Permissions policies (3)

You can attach up to 10 managed policies.

□	Policy name	Type	Attached entities
□	AWSLambdaBasicExecutionR...	Customer managed	1
□	MalwrLambdaS3AccessPolicy	Customer managed	1
□	SecretsManagerReadWrite	AWS managed	1

Step 4: Testing the Function

- Deployed the Lambda function and executed test cases.
- Verified logs in CloudWatch to confirm file hashes and VirusTotal responses.
- Addressed issues such as missing file permissions and API request errors.



testing-malware-scan-bucket

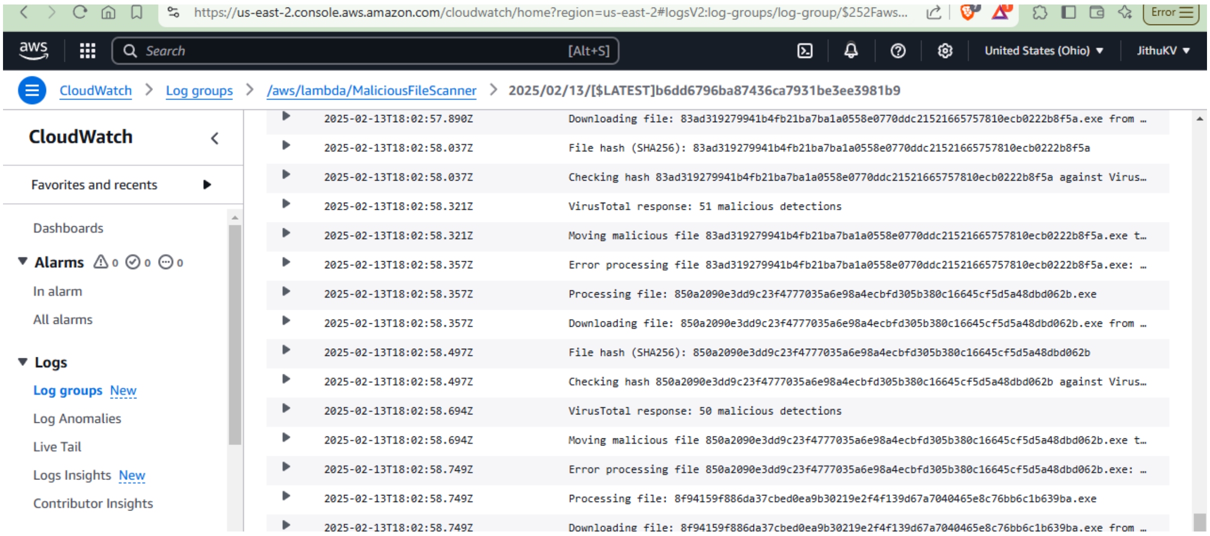
Objects (8)

Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

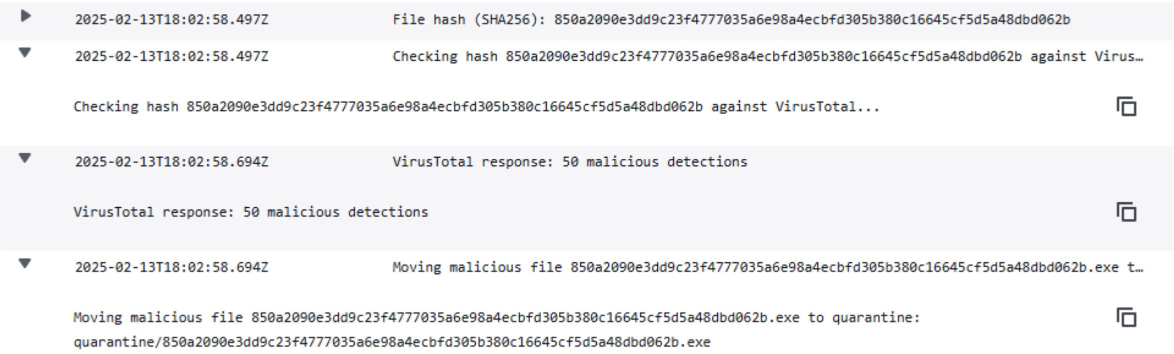
□	Name ▲	Type	Last modified	Size	Storage class
□	120816daada4aeb05da2ee888fe52f25c37b7aba3915da8534904b8247c8526b.exe	exe	February 13, 2025, 11:05:50 (UTC-05:00)	257.5 KB	Standard

Results

- Successfully retrieved file hashes and checked them against VirusTotal.
- Encountered and handled API rate limits appropriately.
- Detected potential malicious files and logged them for further action.

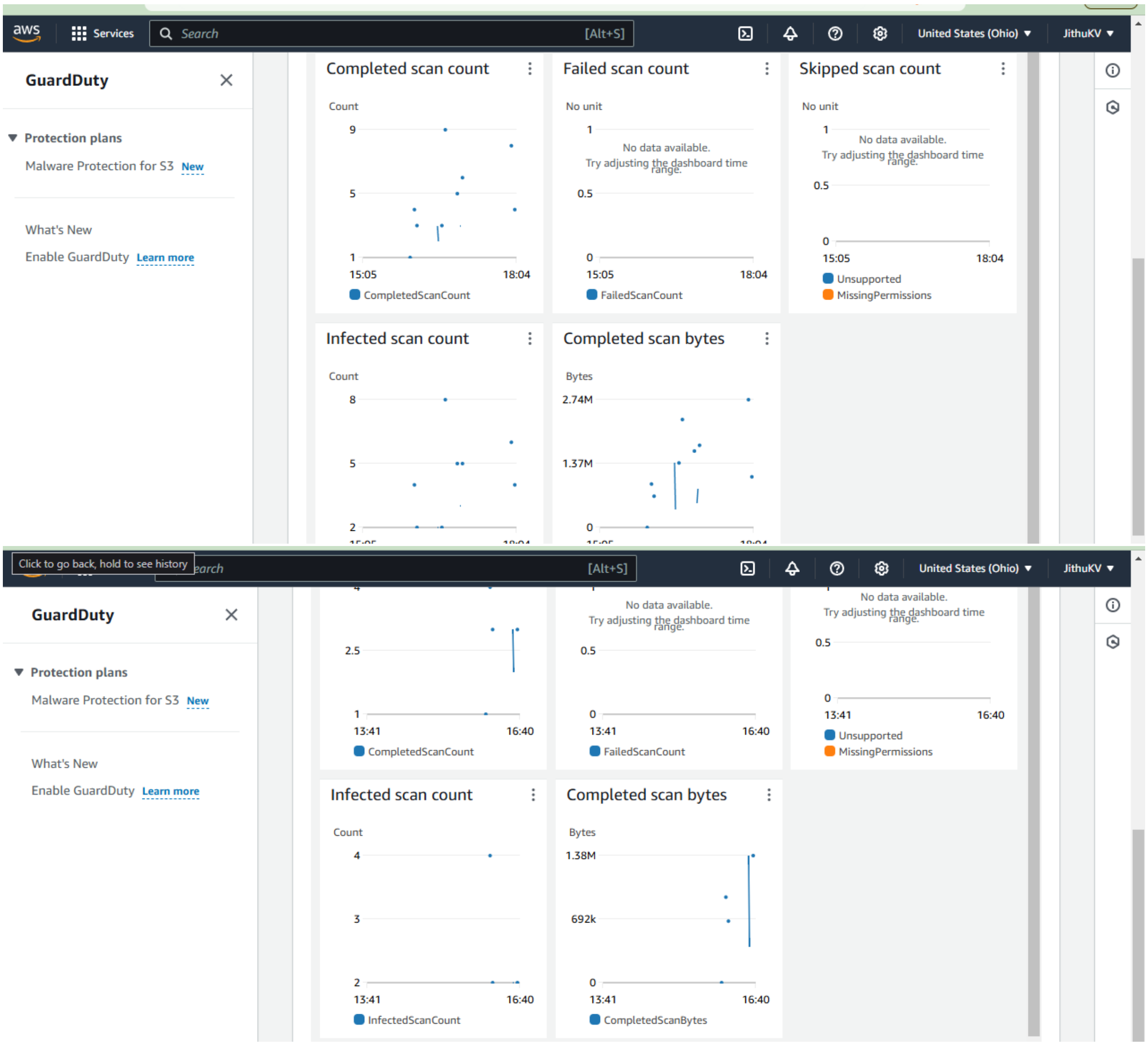

https://us-east-2.console.aws.amazon.com/cloudwatch/home?region=us-east-2#logsV2:log-groups/log-group/$252Faws...
Error
aws
Search
[Alt+S]
United States (Ohio) ▼
JithuKV ▼
CloudWatch > Log groups > /aws/lambda/MaliciousFileScanner > 2025/02/13/[$LATEST]b6dd6796ba87436ca7931be3ee3981b9
CloudWatch
Favorites and recents
Dashboards
Alarms 0 0 0
In alarm
All alarms
Logs
Log groups New
Log Anomalies
Live Tail
Logs Insights New
Contributor Insights

```
2025-02-13T18:02:57.890Z    Downloading file: 83ad319279941b4fb21ba7ba1a0558e0770ddc21521665757810ecb0222b8f5a.exe from …
2025-02-13T18:02:58.037Z    File hash (SHA256): 83ad319279941b4fb21ba7ba1a0558e0770ddc21521665757810ecb0222b8f5a
2025-02-13T18:02:58.037Z    Checking hash 83ad319279941b4fb21ba7ba1a0558e0770ddc21521665757810ecb0222b8f5a against Virus…
2025-02-13T18:02:58.321Z    VirusTotal response: 51 malicious detections
2025-02-13T18:02:58.321Z    Moving malicious file 83ad319279941b4fb21ba7ba1a0558e0770ddc21521665757810ecb0222b8f5a.exe t…
2025-02-13T18:02:58.357Z    Error processing file 83ad319279941b4fb21ba7ba1a0558e0770ddc21521665757810ecb0222b8f5a.exe: …
2025-02-13T18:02:58.357Z    Processing file: 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe
2025-02-13T18:02:58.357Z    Downloading file: 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe from …
2025-02-13T18:02:58.497Z    File hash (SHA256): 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b
2025-02-13T18:02:58.497Z    Checking hash 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b against Virus…
2025-02-13T18:02:58.694Z    VirusTotal response: 50 malicious detections
2025-02-13T18:02:58.694Z    Moving malicious file 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe t…
2025-02-13T18:02:58.749Z    Error processing file 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe: …
2025-02-13T18:02:58.749Z    Processing file: 8f94159f886da37cbed0ea9b30219e2f4f139d67a7040465e8c76bb6c1b639ba.exe
2025-02-13T18:02:58.749Z    Downloading file: 8f94159f886da37cbed0ea9b30219e2f4f139d67a7040465e8c76bb6c1b639ba.exe from …
```



```
▶  2025-02-13T18:02:58.497Z    File hash (SHA256): 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b
▼  2025-02-13T18:02:58.497Z    Checking hash 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b against Virus…

   Checking hash 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b against VirusTotal...

▼  2025-02-13T18:02:58.694Z    VirusTotal response: 50 malicious detections

   VirusTotal response: 50 malicious detections

▼  2025-02-13T18:02:58.694Z    Moving malicious file 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe t…

   Moving malicious file 850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe to quarantine:
   quarantine/850a2090e3dd9c23f4777035a6e98a4ecbfd305b380c16645cf5d5a48dbd062b.exe
```

Guard duty malware agent is able to detect the malicious files too



Conclusion

This project demonstrates an automated approach to scanning files in an S3 bucket for potential malware. Using AWS services like Lambda and GuardDuty, along with VirusTotal integration, ensures a proactive security posture. Future improvements could include integrating SNS for real-time alerts and optimizing API requests for large-scale file scanning.