MedSecure

Health Document Processing and Storage

[Multi-region Secure VPC Infrastructure]

Presenter: Peter Tran

Manager: Jason Brewer

Cohort: 2

Mentor: Adithya Kesineni

Project Overview

Customer Challenge

A customer with multiple VPCs spread across different AWS regions needs to establish efficient connectivity between these VPCs. The current setup lacks centralized management and scalability. The customer also wants to gain visibility into the traffic patterns between VPCs for security and optimization purposes.

Solution

The VPCs across the entire infrastructure will be connected over Transit Gateway and VPC Peering for private, low-latency communication between them. Centralized management is handled through VPC Flow logs, Athena, and QuickSight to gain visibility into traffic patterns.

Real World Edge Case

A healthcare organization requires a secure medical document processing system that ensures that all PHI data never traverses the public internet and remains completely isolated from internet-facing applications to meet HIPAA requirements.

How this will be accomplished:

- **2 Region** Architecture (us-east-1 & us-west-1)
- **2 VPCs** in each region (Web App VPC & Storage VPC)
- Web application will take health documentation uploads, and **securely transfer** it to the storage VPC **without routing out to the internet**.
- A Lambda function will also perform redaction of PHI of that document to comply with HIPAA standards.

Networking Infrastructure Design

Transit Gateway

The purpose of Transit Gateway in this architecture is to provide a scalable solution for multi-region VPC communication.

When VPCs need to be added, you can easily attach them to the existing Transit Gateways, rather than setting up individual peering connections.

VPC Peering

VPC peering is used in this architecture to connect the VPCs that are in the same region.

The reason why we use this and not just rely on the Transit Gateway, is because it provides very low-latency connection, as well as low-cost.

VPC Flow Logs

VPC Flow Logs are configured in each VPC in this infrastructure and stored in an S3 Bucket.

This allows visibility into traffic patterns, while also allowing us to query traffic history using Athena.

Infrastructure Design (cont.)

Athena

Athena in this architecture is used to query the flow logs stored in S3.

This allows us to gain visibility into traffic, and use SQL to find and parse through certain traffic.

QuickSight

QuickSight is used to created dashboards and visualizations of our network traffic.

It leverages the Athena tables to create them.

Useful to find display metrics such as number of failed requests, top IP talkers, etc.

EC2

EC2 Instances are used to host the web applications that take the document uploads. These uploads are routed to the VPC endpoint of SQS.

All of these instances are hosted in a private subnet, behind a load balancer with health checks configured.

Infrastructure Design (cont.)

CloudFront

CloudFront is used to route users to the closest application EC2 instance to them.

Another important use of CloudFront is to ensure encryption in transit, by forcing HTTP traffic to use HTTPS, as well as allowing for failover in the event that a region goes down.

Simple Storage Service (S3)

S3 is used to store the medical documents in both their original and redacted forms.

S3 is best in this use-case because of its durability, as well as having lifecycle policies to move older documents to a cheaper storage class.

Simple Queue Service (SQS)

SQS helps with making sure no documents are lost in the processing/storing process, especially in high-traffic situations.

SQS ensures private transit of these documents since EC2 is using a private VPC endpoint of SQS with strict endpoint policies.

Concern:
S3 cannot be in a VPC.

S3 is a global, region-scoped service that cannot be strictly placed in a VPC. By default, they are accessible via the public internet. With our requirement, how would we make S3 private to our VPC, and follow best-practice security?

This infrastructure uses a **S3 Gateway VPC endpoint** that creates a private route between the storage VPC and S3 without exposing traffic to the internet. With strict bucket policies, our buckets can be configured to **ONLY be accessible from these endpoints**, effectively putting this bucket inside the VPC.

Reasoning

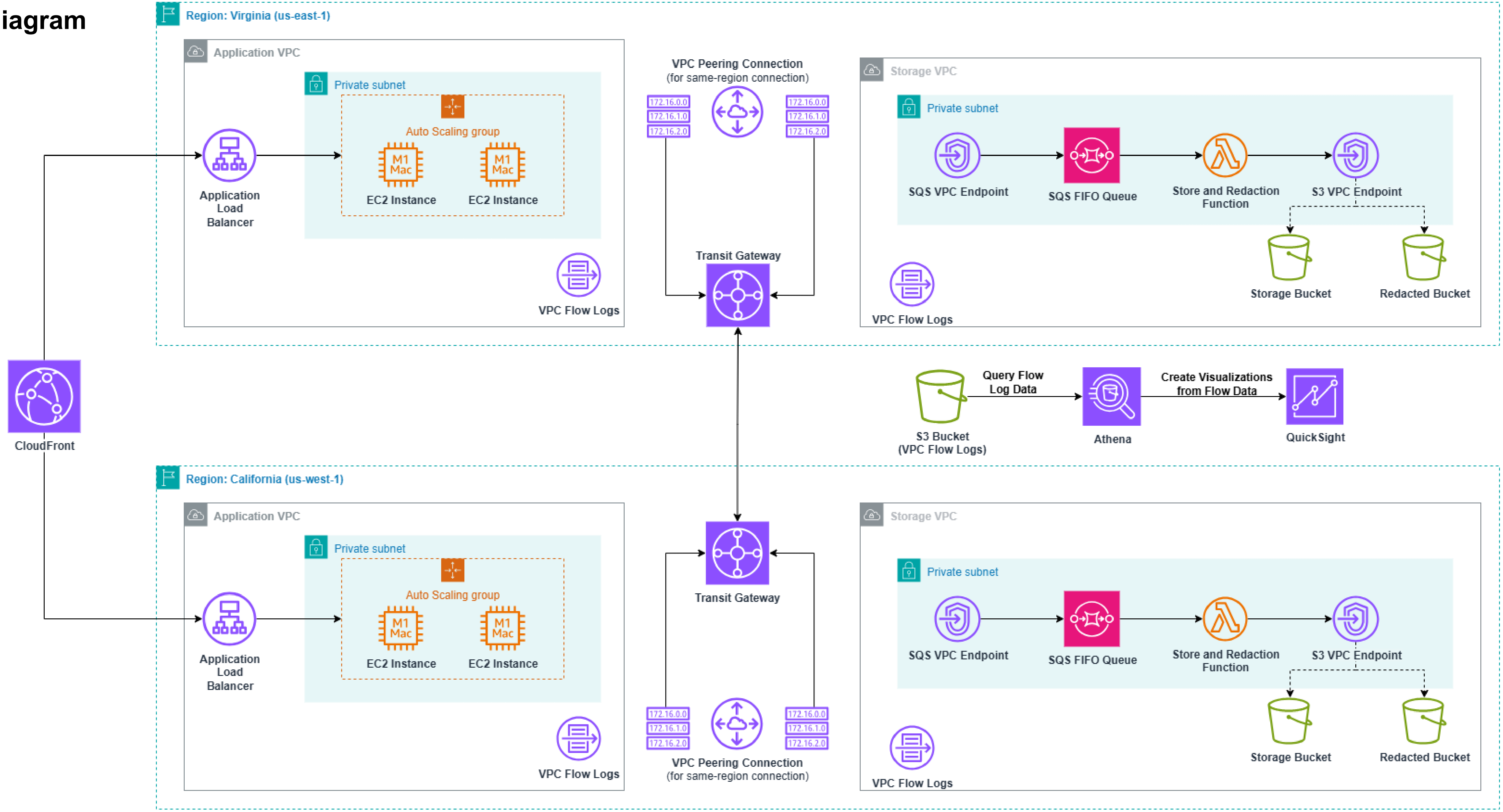
Why Athena? (vs. CloudWatch Logs Insights, OpenSearch)

- Athena seamlessly integrates with QuickSight
- Pay-per-query
- Athena scales for **Petabytes of data** and the Flow Logs will be accumulating up to high storage sizes.

Why QuickSight? (vs. Tableau, Kibana)

- QuickSight is directly integrated with Athena
- Row-level security (not relevant here, but in a real-world situation)
- **Minimal infrastructure overhead**

Architecture Diagram



Why have an Isolated VPC?

An isolated VPC for health document storage is important in this use-case because:

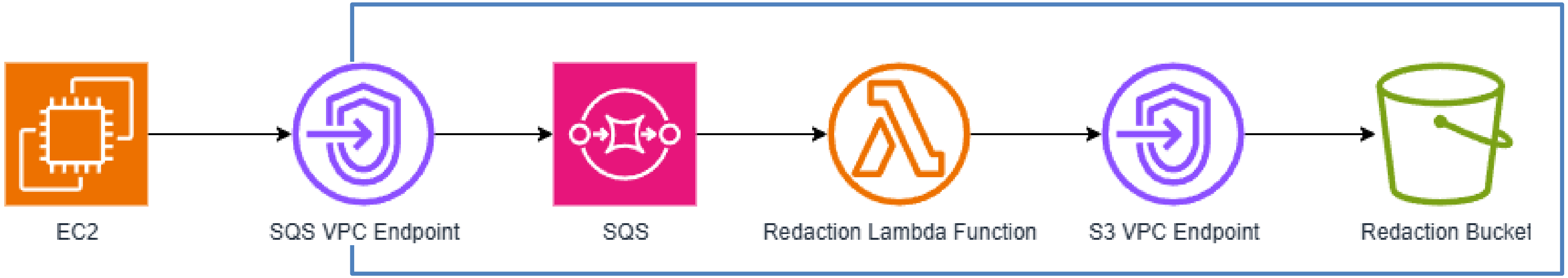
- Isolation prevents SQL injections
- Malware on applications affecting storage
- HIPAA compliance
- No Internet Access

In this scenario, documents are transferred over Transit Gateway (for cross-region) and VPC peering (for same-region) connections.

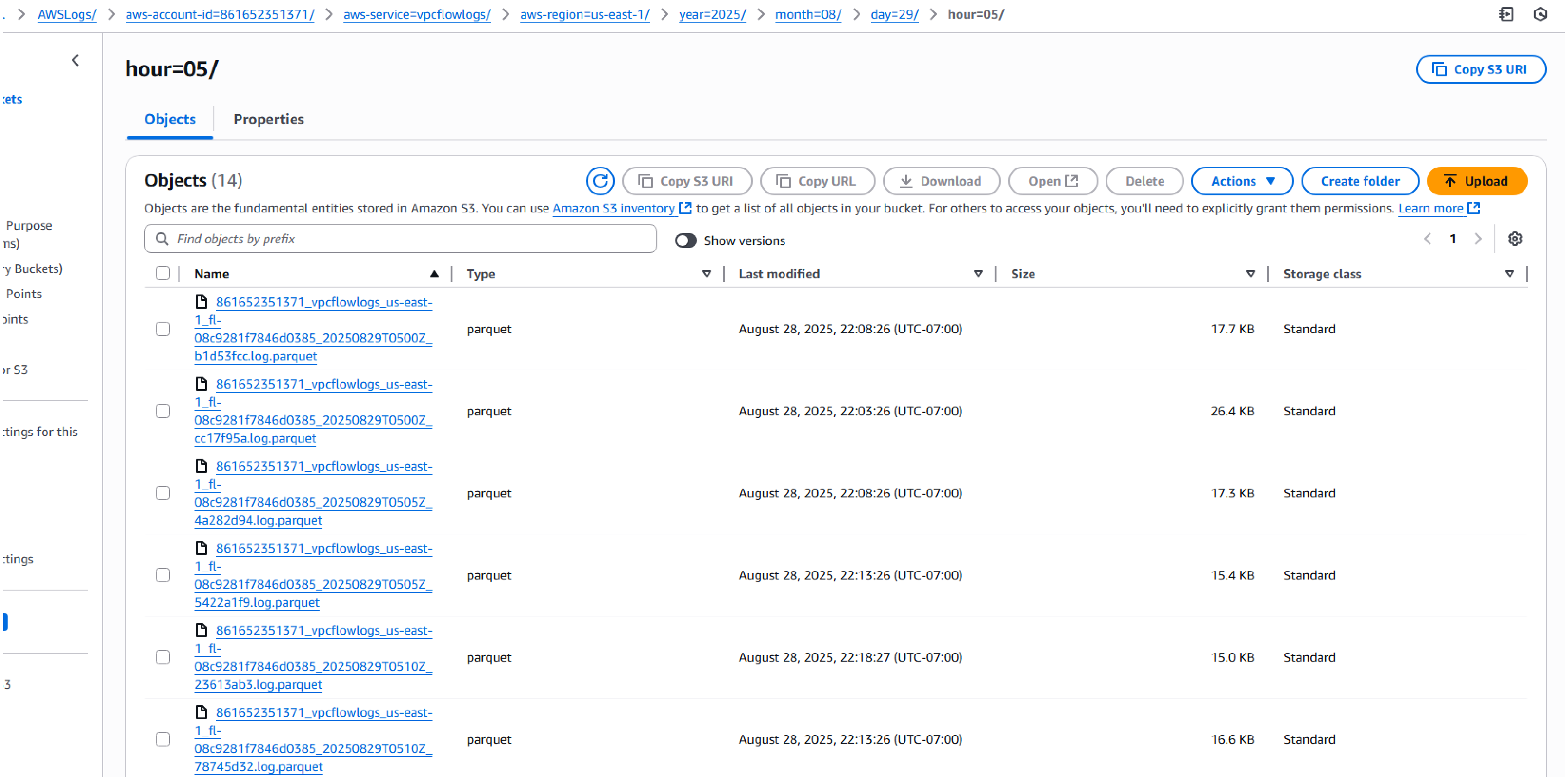
Why have an Isolated VPC?

Our Documents are sent over Transit Gateway / VPC Peering to our storage VPC

- Providing isolation, redaction, and encryption of our documents
- Private storage VPCs have no NAT gateways or Internet Gateways, so no internet connectivity is possible.



VPC Flow Logs Showcase



. > AWSLogs/ > aws-account-id=861652351371/ > aws-service=vpcflowlogs/ > aws-region=us-east-1/ > year=2025/ > month=08/ > day=29/ > hour=05/

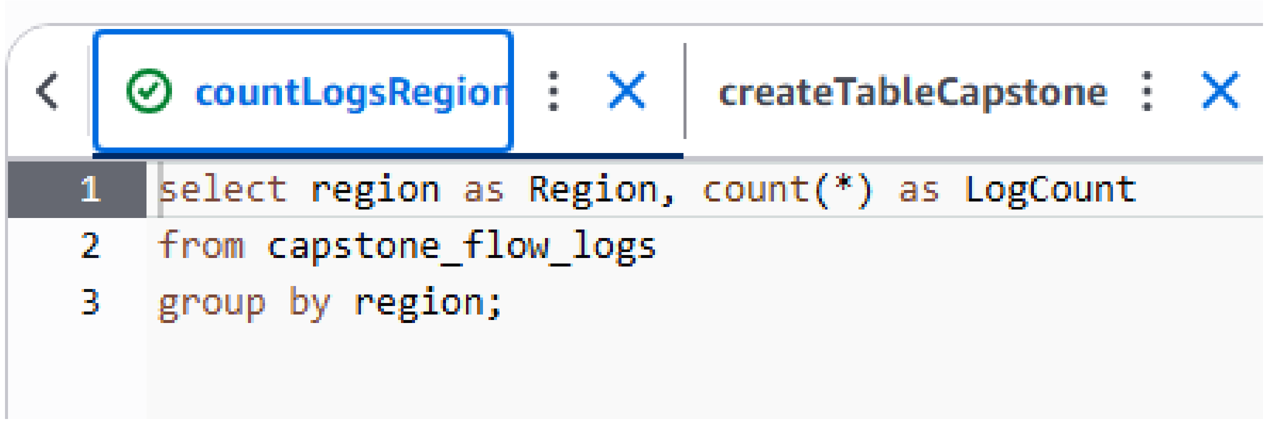
hour=05/

Objects (14)

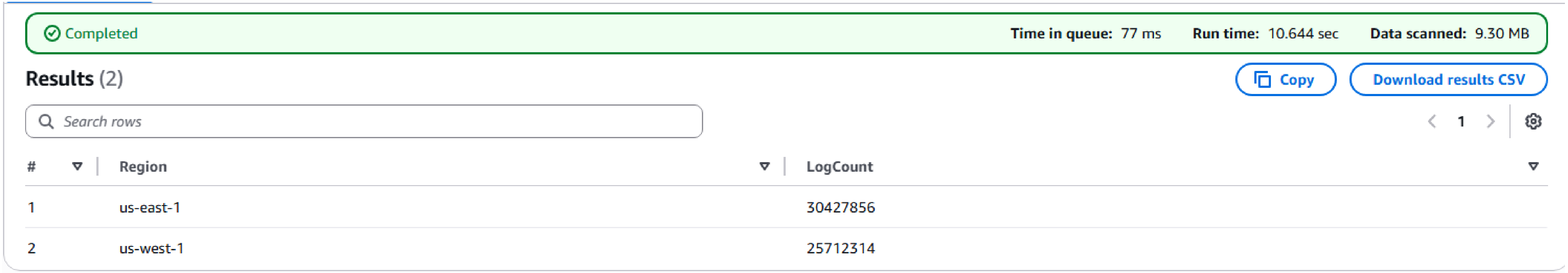
Objects are the fundamental entities stored in Amazon S3. You can use Amazon S3 inventory to get a list of all objects in your bucket. For others to access your objects, you'll need to explicitly grant them permissions. Learn more

	Name	Type	Last modified	Size	Storage class
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0500Z_b1d53fcc.log.parquet	parquet	August 28, 2025, 22:08:26 (UTC-07:00)	17.7 KB	Standard
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0500Z_cc17f95a.log.parquet	parquet	August 28, 2025, 22:03:26 (UTC-07:00)	26.4 KB	Standard
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0505Z_4a282d94.log.parquet	parquet	August 28, 2025, 22:08:26 (UTC-07:00)	17.3 KB	Standard
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0505Z_5422a1f9.log.parquet	parquet	August 28, 2025, 22:13:26 (UTC-07:00)	15.4 KB	Standard
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0510Z_23613ab3.log.parquet	parquet	August 28, 2025, 22:18:27 (UTC-07:00)	15.0 KB	Standard
□	861652351371_vpcflowlogs_us-east-1_fl-08c9281f7846d0385_20250829T0510Z_78745d32.log.parquet	parquet	August 28, 2025, 22:13:26 (UTC-07:00)	16.6 KB	Standard

Athena Query Showcase



```
select region as Region, count(*) as LogCount
from capstone_flow_logs
group by region;
```



Completed

Time in queue: 77 ms Run time: 10.644 sec Data scanned: 9.30 MB

Results (2)

#	Region	LogCount
1	us-east-1	30427856
2	us-west-1	25712314

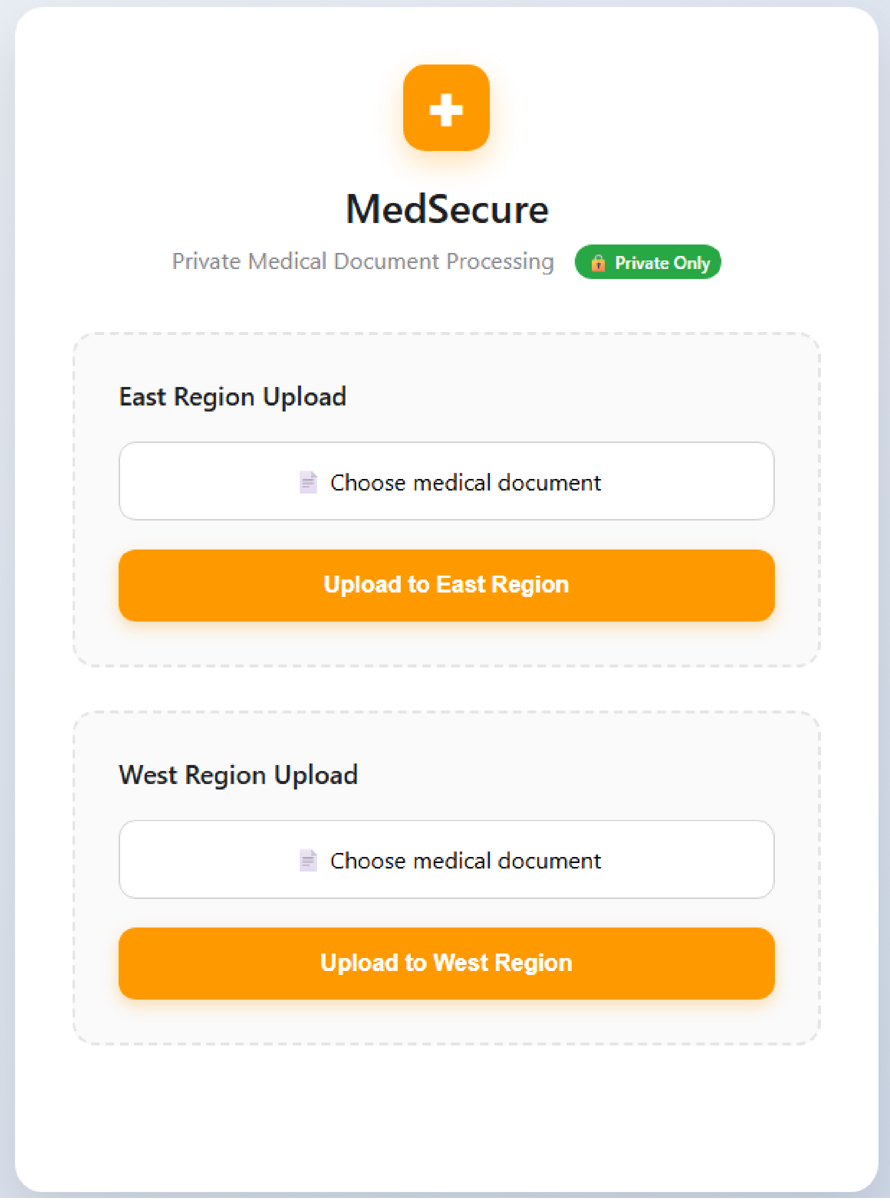
MedSecure

(web application)

This application allows users upload medical documents securely to the storage VPCs in either the East or West region.

Allows the user to browse their files for a medical document, upload it, and it will be securely stored in the storage VPC of that region.

The redaction is done via **Lambda** and **Comprehend Medical**



MedSecure

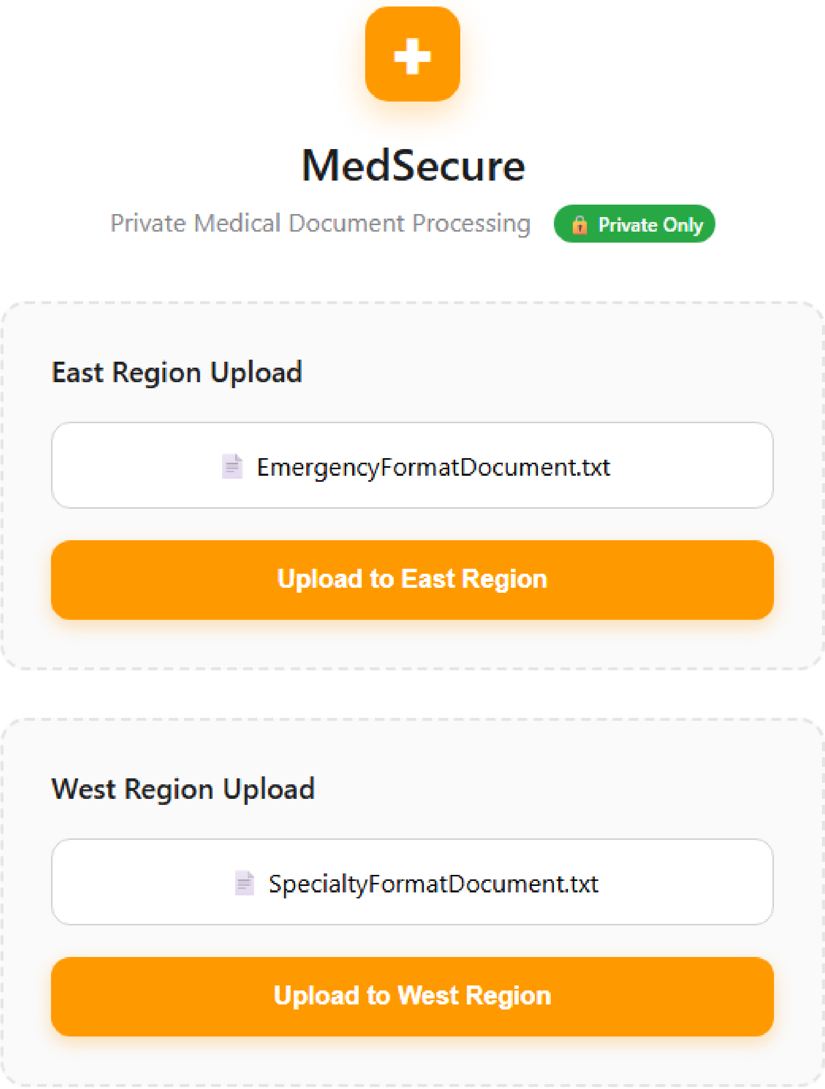
Private Medical Document Processing 🔒 Private Only

East Region Upload

West Region Upload

MedSecure

Work Flow



MedSecure

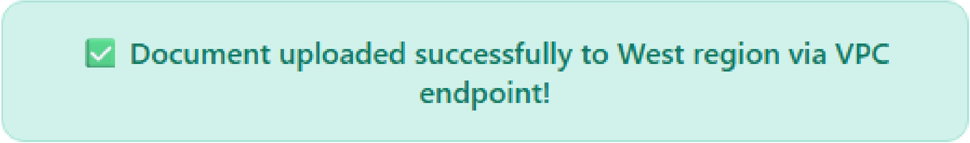
Private Medical Document Processing

East Region Upload

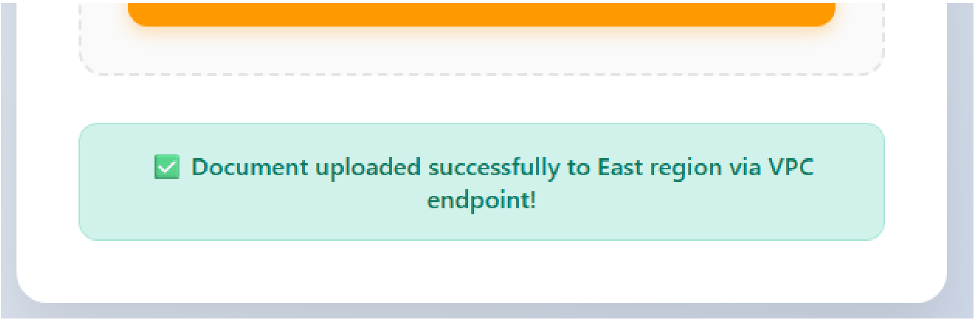
EmergencyFormatDocument.txt

West Region Upload

SpecialtyFormatDocument.txt



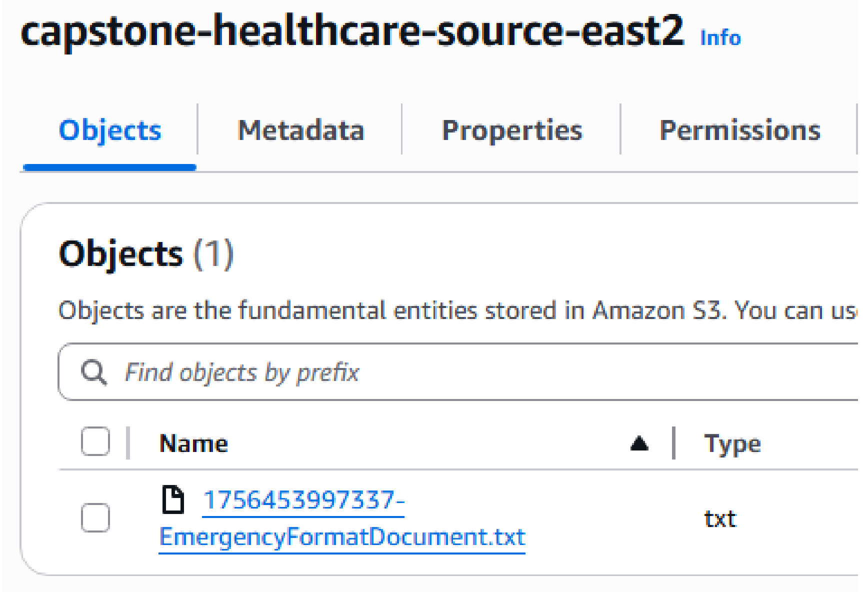
✓ Document uploaded successfully to West region via VPC endpoint!



✓ Document uploaded successfully to East region via VPC endpoint!

MedSecure

Work Flow

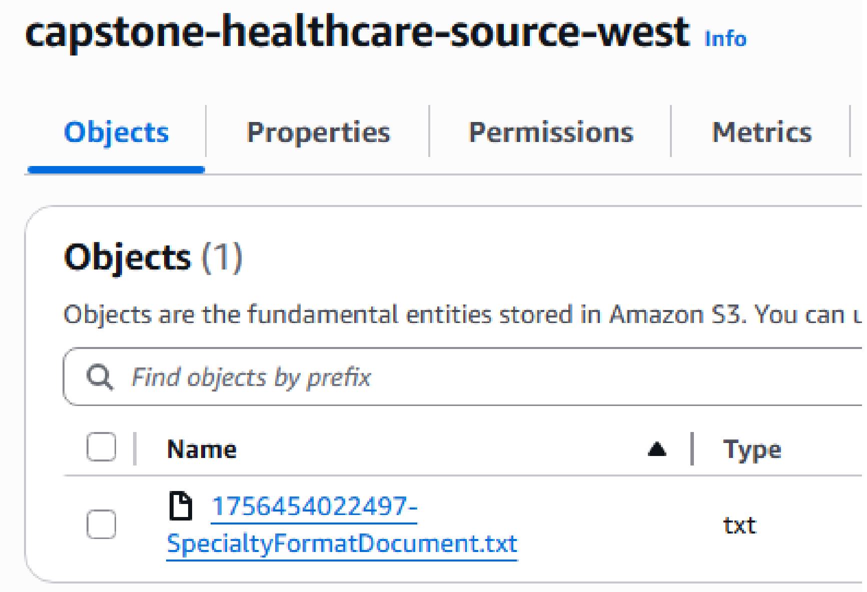


capstone-healthcare-source-east2

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can us

□	Name	Type
□	1756453997337-EmergencyFormatDocument.txt	txt

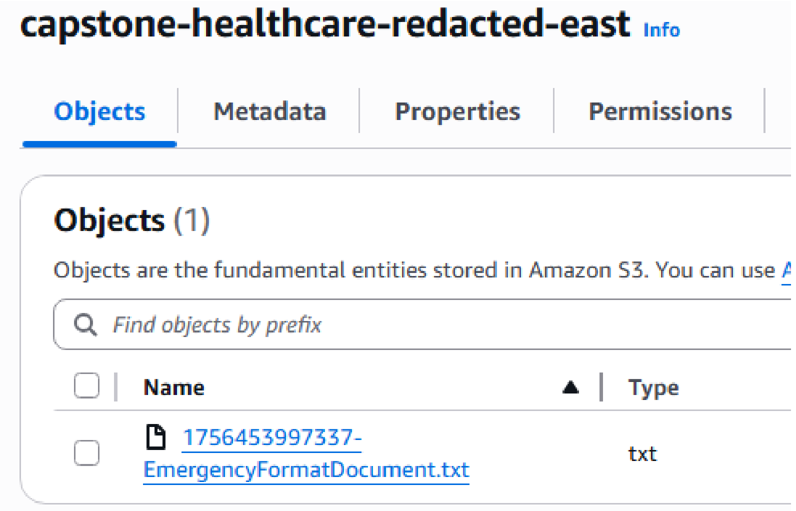


capstone-healthcare-source-west

Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can u

□	Name	Type
□	1756454022497-SpecialtyFormatDocument.txt	txt



capstone-healthcare-redacted-east

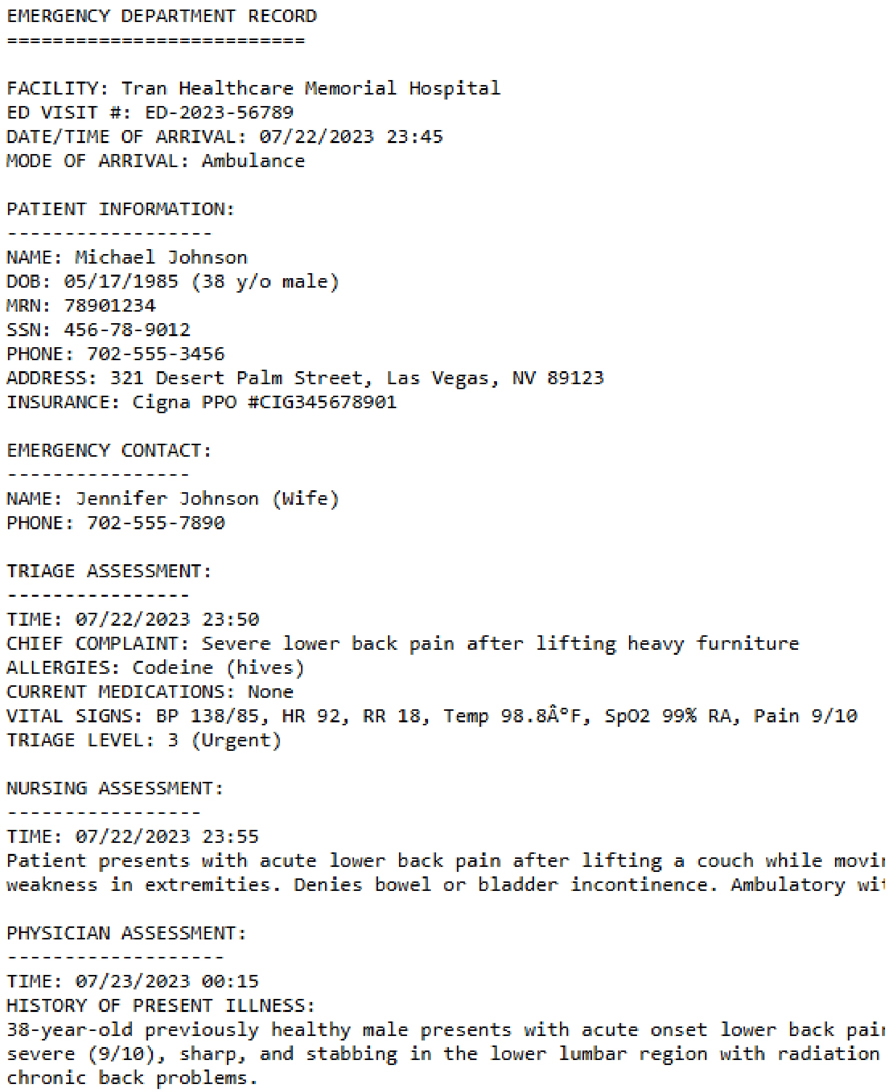
Objects (1)

Objects are the fundamental entities stored in Amazon S3. You can use A

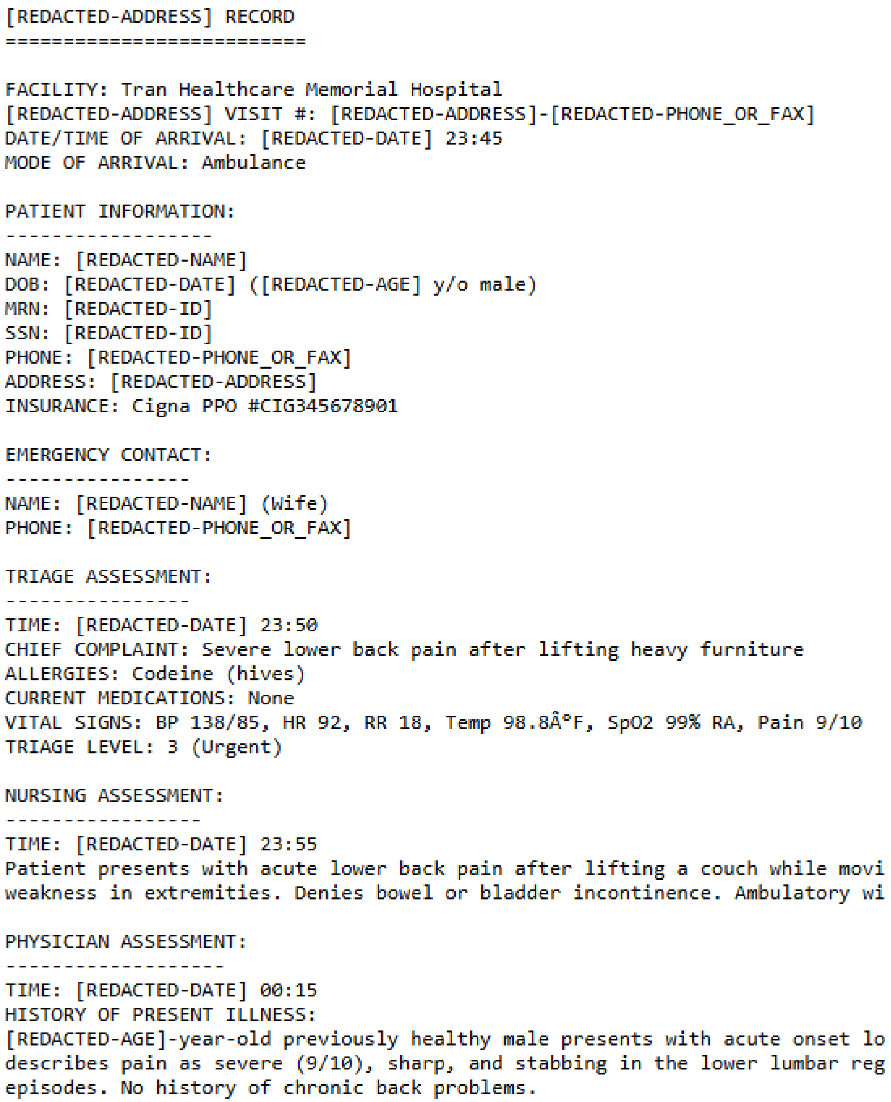
□	Name	Type
□	1756453997337-EmergencyFormatDocument.txt	txt

MedSecure

Work Flow



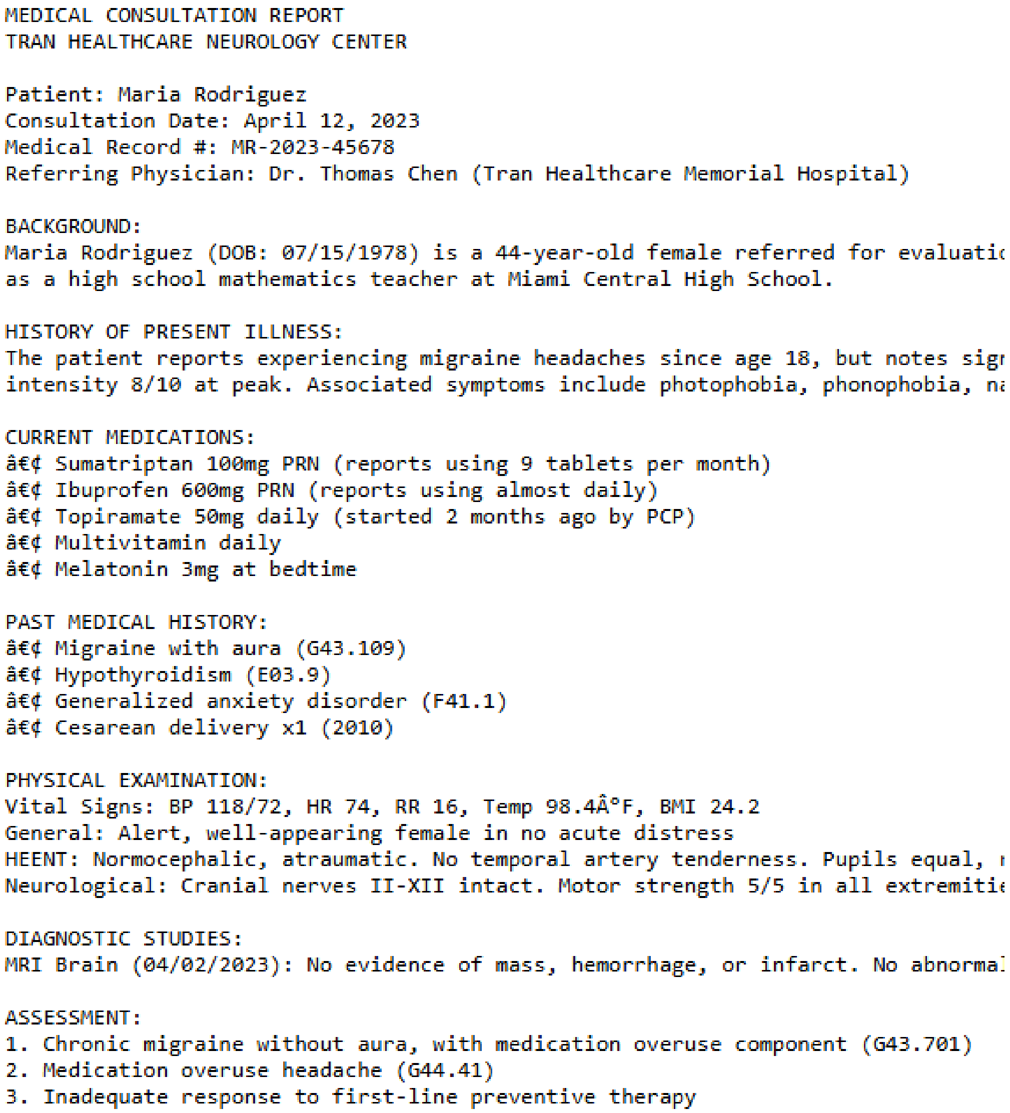
```
EMERGENCY DEPARTMENT RECORD
==========================

FACILITY: Tran Healthcare Memorial Hospital
ED VISIT #: ED-2023-56789
DATE/TIME OF ARRIVAL: 07/22/2023 23:45
MODE OF ARRIVAL: Ambulance

PATIENT INFORMATION:
------------------
NAME: Michael Johnson
DOB: 05/17/1985 (38 y/o male)
MRN: 78901234
SSN: 456-78-9012
PHONE: 702-555-3456
ADDRESS: 321 Desert Palm Street, Las Vegas, NV 89123
INSURANCE: Cigna PPO #CIG345678901

EMERGENCY CONTACT:
----------------
NAME: Jennifer Johnson (Wife)
PHONE: 702-555-7890

TRIAGE ASSESSMENT:
----------------
TIME: 07/22/2023 23:50
CHIEF COMPLAINT: Severe lower back pain after lifting heavy furniture
ALLERGIES: Codeine (hives)
CURRENT MEDICATIONS: None
VITAL SIGNS: BP 138/85, HR 92, RR 18, Temp 98.8Â°F, SpO2 99% RA, Pain 9/10
TRIAGE LEVEL: 3 (Urgent)

NURSING ASSESSMENT:
-----------------
TIME: 07/22/2023 23:55
Patient presents with acute lower back pain after lifting a couch while movi
weakness in extremities. Denies bowel or bladder incontinence. Ambulatory wi

PHYSICIAN ASSESSMENT:
-------------------
TIME: 07/23/2023 00:15
HISTORY OF PRESENT ILLNESS:
38-year-old previously healthy male presents with acute onset lower back pai
severe (9/10), sharp, and stabbing in the lower lumbar region with radiation
chronic back problems.
```



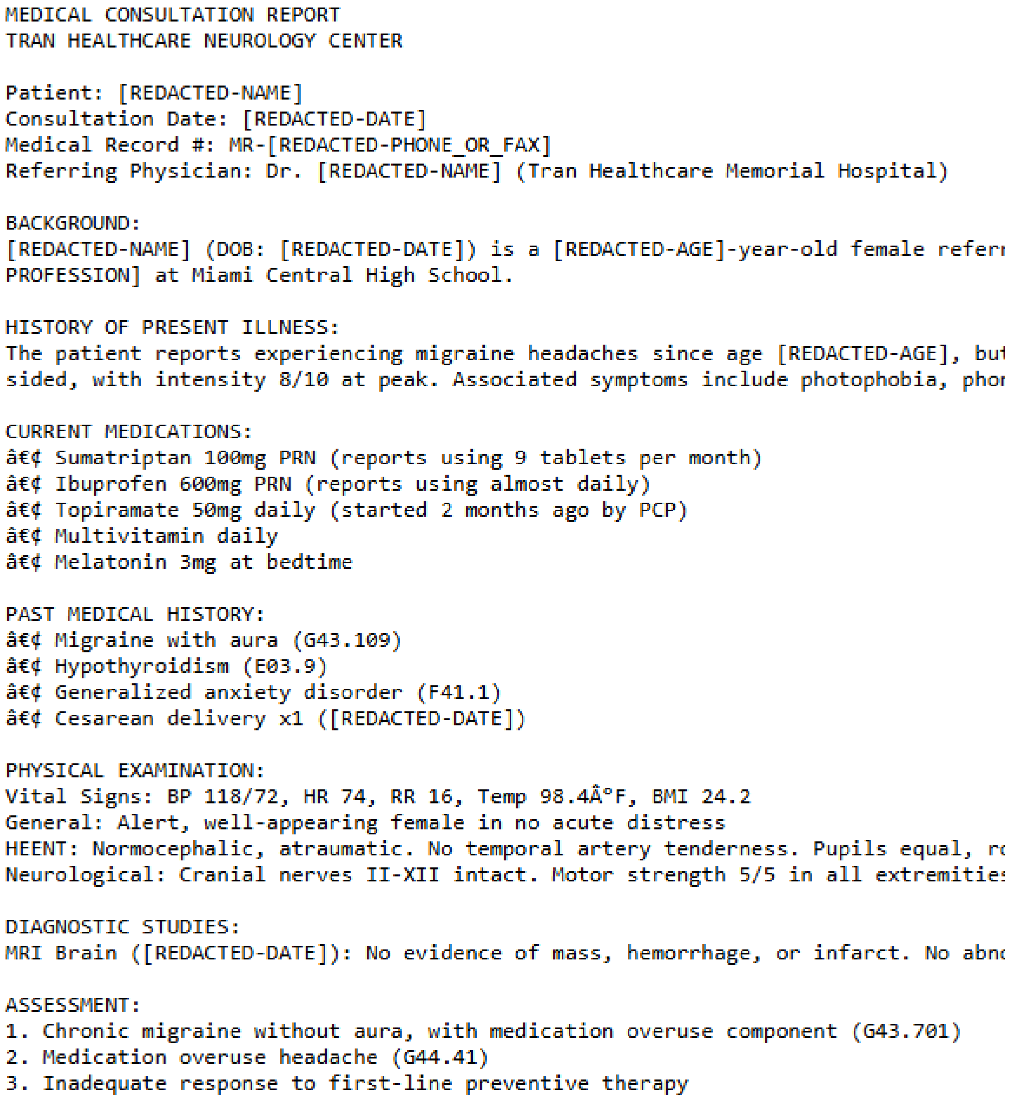
```
[REDACTED-ADDRESS] RECORD
==========================

FACILITY: Tran Healthcare Memorial Hospital
[REDACTED-ADDRESS] VISIT #: [REDACTED-ADDRESS]-[REDACTED-PHONE_OR_FAX]
DATE/TIME OF ARRIVAL: [REDACTED-DATE] 23:45
MODE OF ARRIVAL: Ambulance

PATIENT INFORMATION:
------------------
NAME: [REDACTED-NAME]
DOB: [REDACTED-DATE] ([REDACTED-AGE] y/o male)
MRN: [REDACTED-ID]
SSN: [REDACTED-ID]
PHONE: [REDACTED-PHONE_OR_FAX]
ADDRESS: [REDACTED-ADDRESS]
INSURANCE: Cigna PPO #CIG345678901

EMERGENCY CONTACT:
----------------
NAME: [REDACTED-NAME] (Wife)
PHONE: [REDACTED-PHONE_OR_FAX]

TRIAGE ASSESSMENT:
----------------
TIME: [REDACTED-DATE] 23:50
CHIEF COMPLAINT: Severe lower back pain after lifting heavy furniture
ALLERGIES: Codeine (hives)
CURRENT MEDICATIONS: None
VITAL SIGNS: BP 138/85, HR 92, RR 18, Temp 98.8Â°F, SpO2 99% RA, Pain 9/10
TRIAGE LEVEL: 3 (Urgent)

NURSING ASSESSMENT:
-----------------
TIME: [REDACTED-DATE] 23:55
Patient presents with acute lower back pain after lifting a couch while movi
weakness in extremities. Denies bowel or bladder incontinence. Ambulatory wi

PHYSICIAN ASSESSMENT:
-------------------
TIME: [REDACTED-DATE] 00:15
HISTORY OF PRESENT ILLNESS:
[REDACTED-AGE]-year-old previously healthy male presents with acute onset lo
describes pain as severe (9/10), sharp, and stabbing in the lower lumbar reg
episodes. No history of chronic back problems.
```

MedSecure

Work Flow



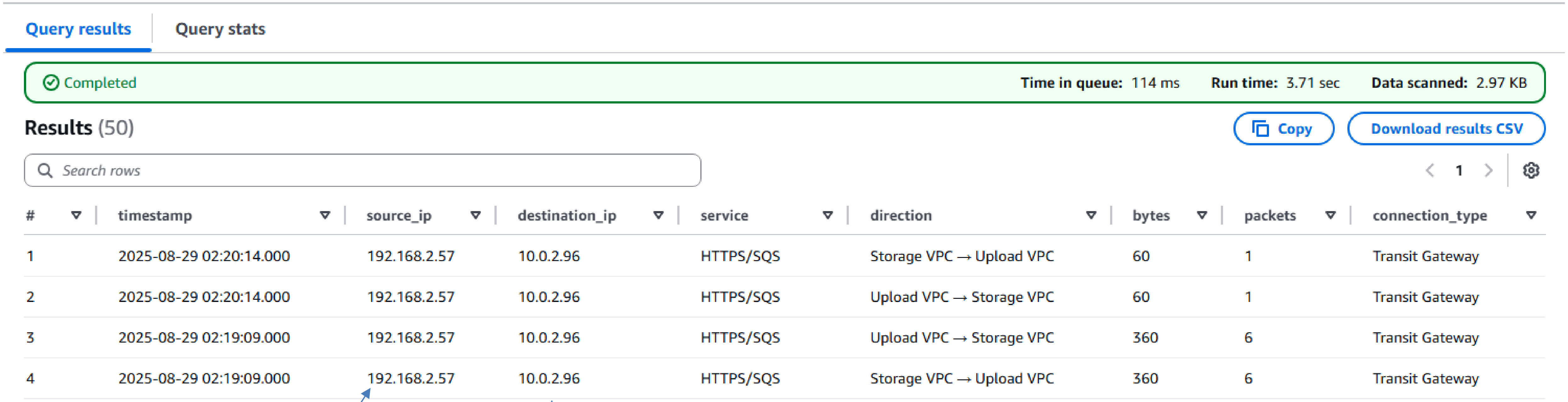
```
MEDICAL CONSULTATION REPORT
TRAN HEALTHCARE NEUROLOGY CENTER

Patient: Maria Rodriguez
Consultation Date: April 12, 2023
Medical Record #: MR-2023-45678
Referring Physician: Dr. Thomas Chen (Tran Healthcare Memorial Hospital)

BACKGROUND:
Maria Rodriguez (DOB: 07/15/1978) is a 44-year-old female referred for evaluati
as a high school mathematics teacher at Miami Central High School.

HISTORY OF PRESENT ILLNESS:
The patient reports experiencing migraine headaches since age 18, but notes sig
intensity 8/10 at peak. Associated symptoms include photophobia, phonophobia, na

CURRENT MEDICATIONS:
â€¢ Sumatriptan 100mg PRN (reports using 9 tablets per month)
â€¢ Ibuprofen 600mg PRN (reports using almost daily)
â€¢ Topiramate 50mg daily (started 2 months ago by PCP)
â€¢ Multivitamin daily
â€¢ Melatonin 3mg at bedtime

PAST MEDICAL HISTORY:
â€¢ Migraine with aura (G43.109)
â€¢ Hypothyroidism (E03.9)
â€¢ Generalized anxiety disorder (F41.1)
â€¢ Cesarean delivery x1 (2010)

PHYSICAL EXAMINATION:
Vital Signs: BP 118/72, HR 74, RR 16, Temp 98.4Â°F, BMI 24.2
General: Alert, well-appearing female in no acute distress
HEENT: Normocephalic, atraumatic. No temporal artery tenderness. Pupils equal,
Neurological: Cranial nerves II-XII intact. Motor strength 5/5 in all extremitie

DIAGNOSTIC STUDIES:
MRI Brain (04/02/2023): No evidence of mass, hemorrhage, or infarct. No abnorma

ASSESSMENT:
1. Chronic migraine without aura, with medication overuse component (G43.701)
2. Medication overuse headache (G44.41)
3. Inadequate response to first-line preventive therapy
```



```
MEDICAL CONSULTATION REPORT
TRAN HEALTHCARE NEUROLOGY CENTER

Patient: [REDACTED-NAME]
Consultation Date: [REDACTED-DATE]
Medical Record #: MR-[REDACTED-PHONE_OR_FAX]
Referring Physician: Dr. [REDACTED-NAME] (Tran Healthcare Memorial Hospital)

BACKGROUND:
[REDACTED-NAME] (DOB: [REDACTED-DATE]) is a [REDACTED-AGE]-year-old female refer
PROFESSION] at Miami Central High School.

HISTORY OF PRESENT ILLNESS:
The patient reports experiencing migraine headaches since age [REDACTED-AGE], bu
sided, with intensity 8/10 at peak. Associated symptoms include photophobia, phon

CURRENT MEDICATIONS:
â€¢ Sumatriptan 100mg PRN (reports using 9 tablets per month)
â€¢ Ibuprofen 600mg PRN (reports using almost daily)
â€¢ Topiramate 50mg daily (started 2 months ago by PCP)
â€¢ Multivitamin daily
â€¢ Melatonin 3mg at bedtime

PAST MEDICAL HISTORY:
â€¢ Migraine with aura (G43.109)
â€¢ Hypothyroidism (E03.9)
â€¢ Generalized anxiety disorder (F41.1)
â€¢ Cesarean delivery x1 ([REDACTED-DATE])

PHYSICAL EXAMINATION:
Vital Signs: BP 118/72, HR 74, RR 16, Temp 98.4Â°F, BMI 24.2
General: Alert, well-appearing female in no acute distress
HEENT: Normocephalic, atraumatic. No temporal artery tenderness. Pupils equal, r
Neurological: Cranial nerves II-XII intact. Motor strength 5/5 in all extremitie

DIAGNOSTIC STUDIES:
MRI Brain ([REDACTED-DATE]): No evidence of mass, hemorrhage, or infarct. No abn

ASSESSMENT:
1. Chronic migraine without aura, with medication overuse component (G43.701)
2. Medication overuse headache (G44.41)
3. Inadequate response to first-line preventive therapy
```

MedSecure

Monitoring uploads with Athena and Flow Logs



Completed **Time in queue:** 114 ms **Run time:** 3.71 sec **Data scanned:** 2.97 KB

Results (50)

#	timestamp	source_ip	destination_ip	service	direction	bytes	packets	connection_type
1	2025-08-29 02:20:14.000	192.168.2.57	10.0.2.96	HTTPS/SQS	Storage VPC → Upload VPC	60	1	Transit Gateway
2	2025-08-29 02:20:14.000	192.168.2.57	10.0.2.96	HTTPS/SQS	Upload VPC → Storage VPC	60	1	Transit Gateway
3	2025-08-29 02:19:09.000	192.168.2.57	10.0.2.96	HTTPS/SQS	Upload VPC → Storage VPC	360	6	Transit Gateway
4	2025-08-29 02:19:09.000	192.168.2.57	10.0.2.96	HTTPS/SQS	Storage VPC → Upload VPC	360	6	Transit Gateway

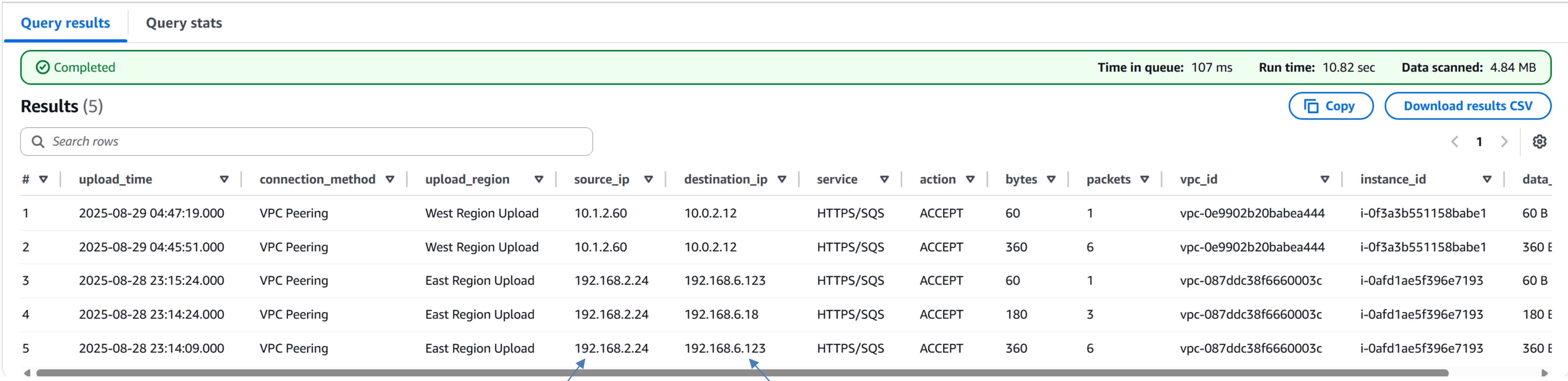
East App CIDR

West Storage SQS Endpoint IP

Note: These are from the Transit Gateway Logs, and the columns are labeled accordingly to better demonstrate the flow of traffic

MedSecure

Monitoring uploads with Athena and Flow Logs



Completed **Time in queue:** 107 ms **Run time:** 10.82 sec **Data scanned:** 4.84 MB

Results (5)

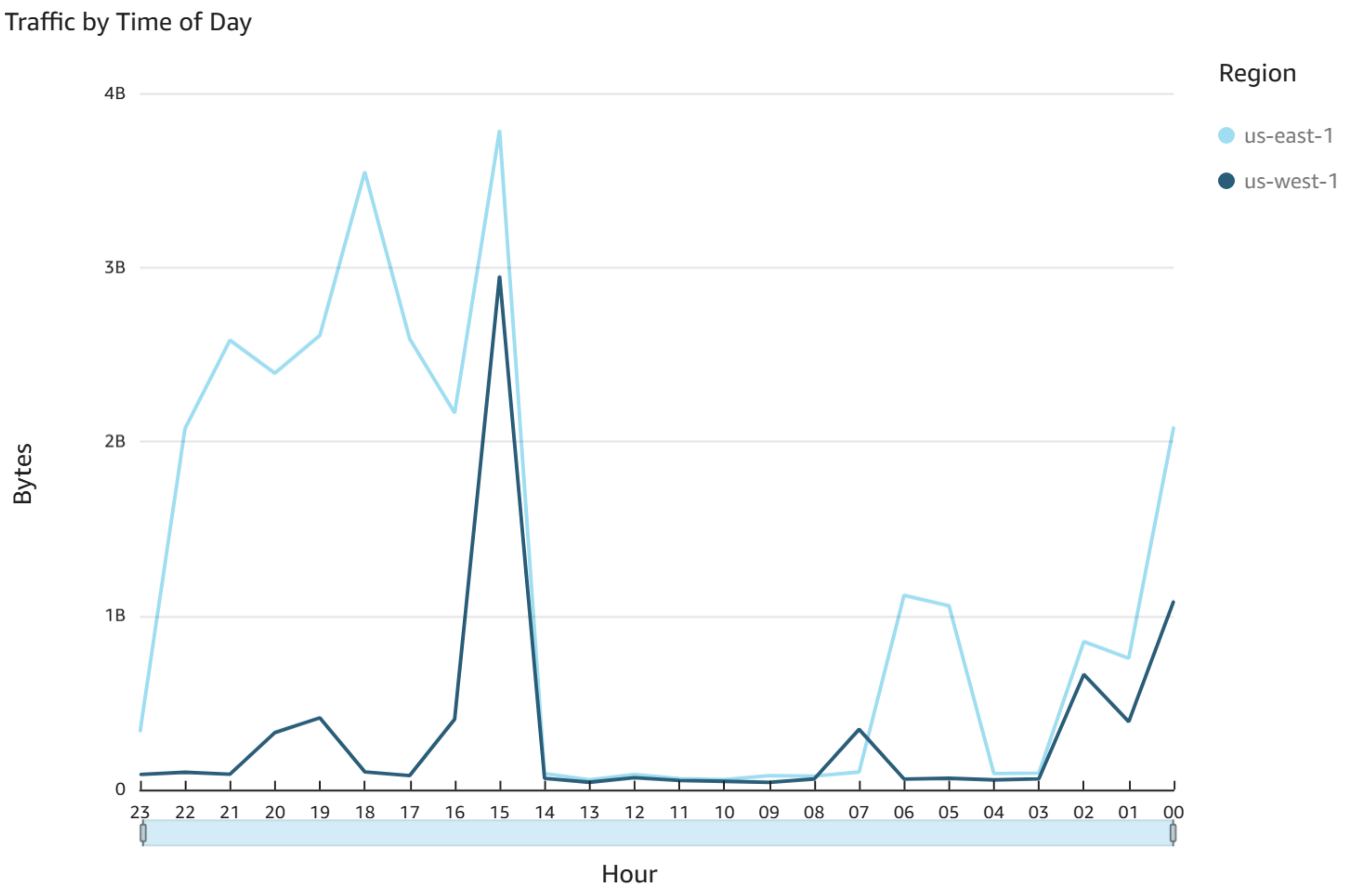
#	upload_time	connection_method	upload_region	source_ip	destination_ip	service	action	bytes	packets	vpc_id	instance_id	data_
1	2025-08-29 04:47:19.000	VPC Peering	West Region Upload	10.1.2.60	10.0.2.12	HTTPS/SQS	ACCEPT	60	1	vpc-0e9902b20babea444	i-0f3a3b551158babe1	60 B
2	2025-08-29 04:45:51.000	VPC Peering	West Region Upload	10.1.2.60	10.0.2.12	HTTPS/SQS	ACCEPT	360	6	vpc-0e9902b20babea444	i-0f3a3b551158babe1	360 B
3	2025-08-28 23:15:24.000	VPC Peering	East Region Upload	192.168.2.24	192.168.6.123	HTTPS/SQS	ACCEPT	60	1	vpc-087ddc38f6660003c	i-0afd1ae5f396e7193	60 B
4	2025-08-28 23:14:24.000	VPC Peering	East Region Upload	192.168.2.24	192.168.6.18	HTTPS/SQS	ACCEPT	180	3	vpc-087ddc38f6660003c	i-0afd1ae5f396e7193	180 B
5	2025-08-28 23:14:09.000	VPC Peering	East Region Upload	192.168.2.24	192.168.6.123	HTTPS/SQS	ACCEPT	360	6	vpc-087ddc38f6660003c	i-0afd1ae5f396e7193	360 B

East App CIDR

East Storage SQS Endpoint IP(s)

QuickSight Visualizations

Visualizing Traffic to make Conclusions



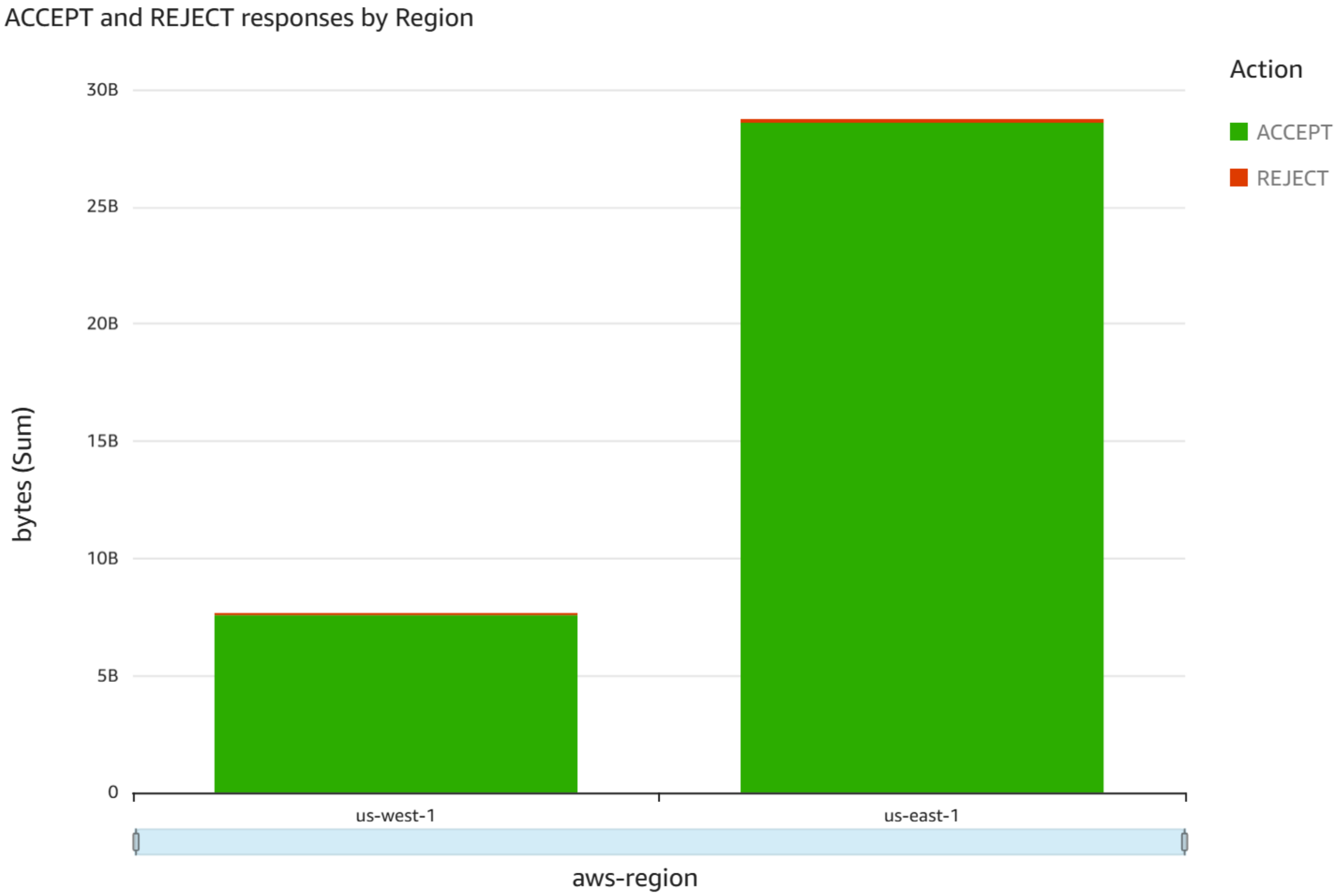
This visual shows off the traffic in each region depending on the time of Day.

Useful for seeing when your high traffic times are, to help you make:

- Scaling decisions
- Infrastructure changes
- Etc.

QuickSight Visualizations

Visualizing Traffic to make Conclusions

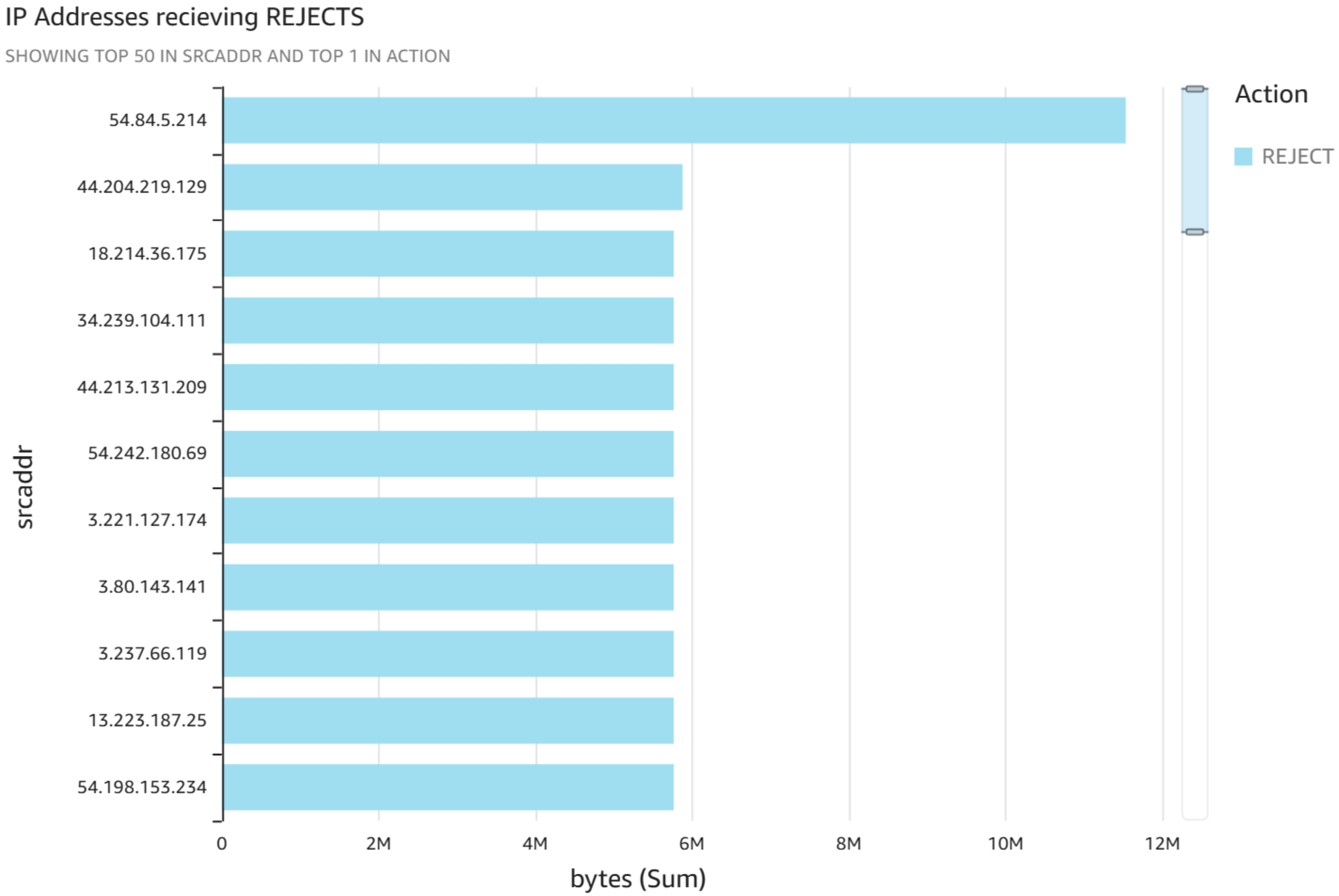


This visual shows off the accepted and rejecting traffic across the two regions.

This will be useful in scenarios where a region is having issues, and you can see which one is rejecting an abnormal amount of traffic.

QuickSight Visualizations

Visualizing Traffic to make Conclusions



This visual shows off which IP addresses are getting REJECT responses.

This will be useful in scenarios where you want to pinpoint what IP addresses are getting rejected.

(ex. View attacker IP addresses when they get rejected)

Frequently Asked Questions

Aren’t the files going over the public internet when uploaded from the client’s device?

This service uses CloudFront, with enforced HTTPS connection to the web application, which will ensure encrypted data in transit. In the event of needed additional security, the private AWS VPC infrastructure can be integrated with either the company’s existing VPN or AWS’s Client VPN service.

What the experience of a healthcare provider using this service?

The healthcare provider will be able to seamlessly upload their documents to this portal over the private AWS network. The documentation’s redaction and storage in the storage VPC of each region is handled automatically upon client upload.

How is using MedSecure more effective than simply storing the documents on-site?

MedSecure and the infrastructure behind it are crucial for security. This is because the documents are completely isolated from internet traffic, while still allowing customers to upload documents from wherever they’re operating from. This solution is also more scalable, allowing Tran Healthcare to expand out to more regions whenever needed.

Who can use MedSecure?

MedSecure can be configured to only allow private IPs leveraging either a company VPN, or a Client VPN on AWS. Authentication of users can also be configured for Cognito and IAM Identity center authentication. For the purposes of the demo, this service is available to all.

How does the redaction work?

The redaction of healthcare documents is handled by AWS Lambda as well as Comprehend Medical. This is crucial for redacting documents because no matter what format of document, Comprehend Medical leverages artificial intelligence to detect PHI, and redact it for secure storage.

Any Questions?

Contact Me!

email: petertran004@gmail.com

in : in/petertran004