

Enterprise Centralized Logging and Observability Architecture on AWS.

1. Problem Statement

An organization operating multiple AWS accounts lacked centralized visibility into logs and security findings. Logs were dispersed across accounts, limiting detection capabilities. A unified logging and SIEM integration architecture was required.

I implemented this solution to ensure that critical logs from distributed workloads are consistently collected, securely stored, retained according to compliance requirements, and analyzed through a centralized SIEM platform.

2. Objectives

The primary objectives of this implementation were to:

- Centralize AWS service and application logs
- Improve security monitoring and incident investigation
- Support regulatory and compliance requirements
- Enable real-time detection and alerting
- Ensure secure and durable log storage
- Reduce operational blind spots

3. Logging Architecture Overview

The centralized logging platform was built using:

- AWS CloudTrail
- Amazon CloudWatch Logs
- Amazon Kinesis Data Firehose
- AWS Lambda
- Amazon OpenSearch
- Amazon S3

All service and application logs are routed through CloudWatch and streamed into a secure analytics platform.

4. CloudTrail Configuration

CloudTrail was configured as a multi-region, organization-level trail.

The screenshot shows the AWS CloudTrail console with a green success message at the top: "Trail successfully created". Below it, the "Trails" table lists one entry:

Name	Home region	Multi-region trail	ARN	Insights	Organization trail	S3 bucket	Log file prefix	CloudWatch Logs log group	Status
CloudsecTrail	US East (N. Virginia)	Yes	arn:aws:cloudtrail:us-east-1:879381257906:trail/CloudsecTrail	Enabled	Yes	aws-cloudtrail-logs-879381257906-dd20b7e0	-	arn:aws:logs:us-east-1:879381257906:log-group:aws-cloudtrail-logs-879381257906-9efc6266-*	Logging

The configuration includes:

- Management event logging
- CloudTrail Insights for anomaly detection
- Centralized S3 storage
- Streaming to CloudWatch Logs

This ensured full visibility into account activity across regions.

5. Log Retention and Lifecycle Management

CloudWatch Log Groups were configured with defined retention periods to prevent indefinite storage.

CloudTrail logs stored in S3 are managed using lifecycle policies that:

- Retain logs in S3 Standard for 90 days

The screenshot shows the CloudWatch Log Management interface. On the left, there's a sidebar with navigation links like CloudWatch, Favorites and recents, AI Operations, GenAI Observability, Application Signals (APM), Infrastructure Monitoring, Logs (Log Management New, Log Anomalies, Live Tail, Logs Insights, Contributor Insights), and Metrics (All metrics, Explorer). The main area is titled "Log groups (1)" and shows a single log group named "f6266" which is Standard type, configured, and has its retention set to off for 3 months. There are buttons for Actions, View in Logs Insights, Start tailing, and Create log group.

- Archive older logs to S3 Glacier

- Delete logs after one year

The screenshot shows the Amazon S3 Lifecycle rule configuration for a bucket named "aws-cloudtrail-logs-879381257906-dd20b7e0". The lifecycle rule is named "cloudseclogretentioncycle". It is enabled and applies to the entire bucket. The rule defines actions for objects based on their age: moving them to "Glacier Instant Retrieval" at Day 90 and deleting them at Day 365. There are sections for Prefix, Object tags, Minimum object size, and Maximum object size. A large watermark of the letters "A" and "YU" is overlaid on the bottom left of the screenshot.

This approach aligns with common compliance standards such as PCI-DSS while controlling storage costs.

6. Centralized SIEM Platform (Amazon OpenSearch)

Amazon OpenSearch was deployed as the centralized log analytics and SIEM platform.

The screenshot shows the Amazon OpenSearch Service console with the 'cloudssecdomain' domain selected. The 'Cluster configuration' tab is active. Key details shown include:

- Name:** cloudssecdomain
- Domain ARN:** arn:aws:es:us-east-1:879381257906:domain/cloudssecdomain
- Deployment option(s):** 3-AZ with standby
- Domain processing status:** Active
- Configuration change status:** Completed
- Cluster health:** Green
- Version:** OpenSearch 3.5 (latest)
- Service software version:** OpenSearch_3_3_R20251121-P3 (latest)
- OpenSearch Dashboards URL (IPv4):** https://search-cloudssecdomain-ueat4lzgysjf2ppqmbxoc72rm.us-east-1.es.amazonaws.com/_dashboards
- Domain endpoint:** https://search-cloudssecdomain-ueat4lzgysjf2ppqmbxoc72rm.us-east-1.es.amazonaws.com

The domain was configured with:

- No public internet exposure
- VPC-based deployment
- Fine-grained access control
- IAM-based authentication
- Encrypted storage
- Node-to-node encryption

Access was restricted using security groups to authorized ingestion services.

7. Secure Log Ingestion Pipeline

7.1 CloudWatch Subscription Filters

Subscription filters were created to forward all CloudTrail events to Kinesis Data Firehose.

No filtering was applied to ensure complete audit coverage.

7.2 Kinesis Data Firehose Configuration

Firehose was configured with:

- OpenSearch as primary destination

- S3 backup destination
- VPC delivery enabled

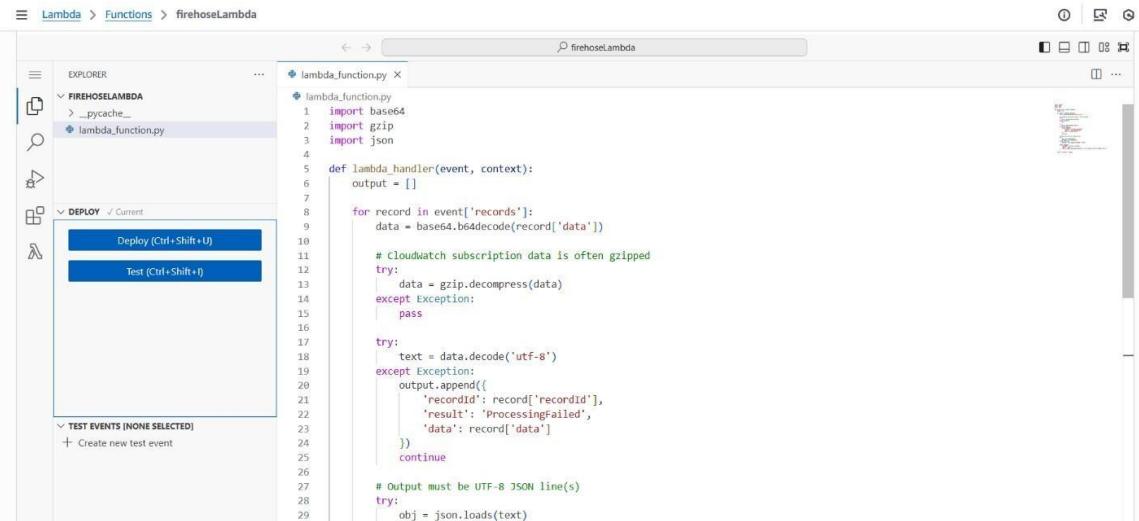
The screenshot shows the Amazon Data Firehose console interface. At the top, a blue banner displays the message: "You have stopped sending demo data to your Firehose stream PUT-OPS-KKey6." Below this, the stream name "PUT-OPS-KKey6" is shown with an "Info" link. On the left, a sidebar titled "Amazon Data Firehose" includes links for "Firehose streams", "Resources" (with "What's new", "Developer guide", and "API reference"), and "Monitoring". The main content area is titled "PUT-OPS-KKey6" and contains "Firehose stream details". It shows the status as "Active", the destination as "Amazon OpenSearch Service", and the ARN as "arn:aws:firehose:us-east-1:879381257906:deliverystream/PUT-OPS-KKey6". A "Data transformation" section indicates "Enabled". Below this, a "Test with demo data" button is available. At the bottom, tabs for "Monitoring" (selected), "Configuration", "Destination error logs", and "Backup error logs" are present. The "Firehose stream metrics" section shows three metrics: "Incoming bytes" (Bytes), "Incoming put requests" (Count), and "Incoming records" (Count). A timestamp "3h" is shown next to the metrics.

An IAM role was created with permissions for:

- Writing to OpenSearch
- Delivering data to S3
- Managing network interfaces

7.3 Lambda Log Transformation

A Lambda function was integrated into Firehose to perform log transformation.



```
lambda_function.py
1 import base64
2 import gzip
3 import json
4
5 def lambda_handler(event, context):
6     output = []
7
8     for record in event['records']:
9         data = base64.b64decode(record['data'])
10
11     # CloudWatch subscription data is often gzipped
12     try:
13         data = gzip.decompress(data)
14     except Exception:
15         pass
16
17     try:
18         text = data.decode('utf-8')
19     except Exception:
20         output.append({
21             'recordId': record['recordId'],
22             'result': 'ProcessingFailed',
23             'data': record['data']
24         })
25     continue
26
27     # Output must be UTF-8 JSON line(s)
28     try:
29         obj = json.loads(text)
```

Functions included:

- Normalizing log formats
- Adding metadata
- Ensuring indexing compatibility

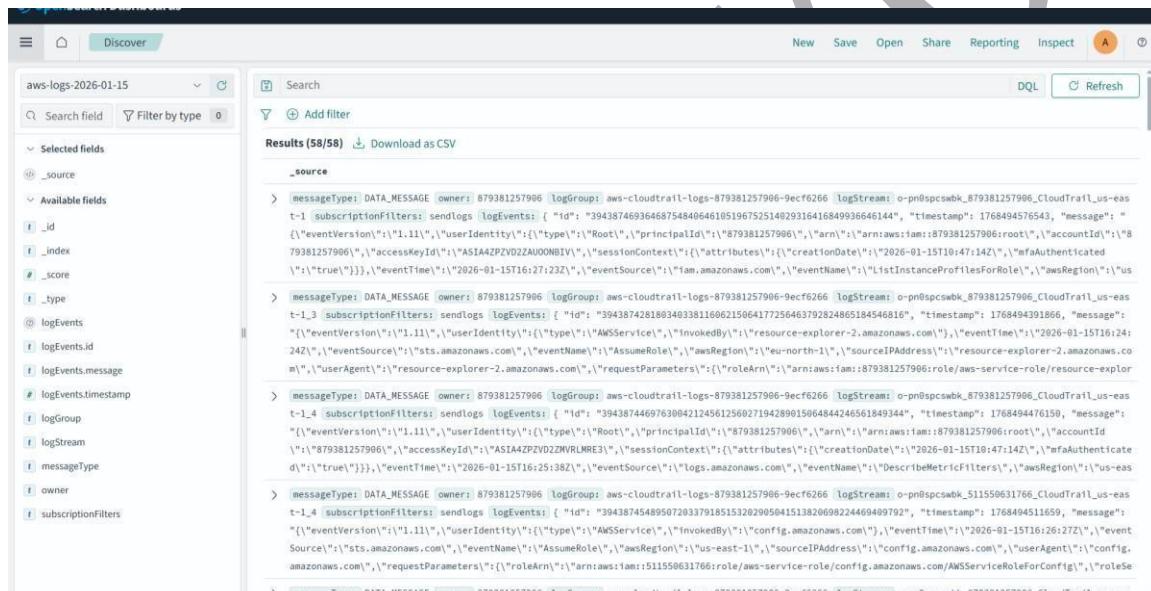
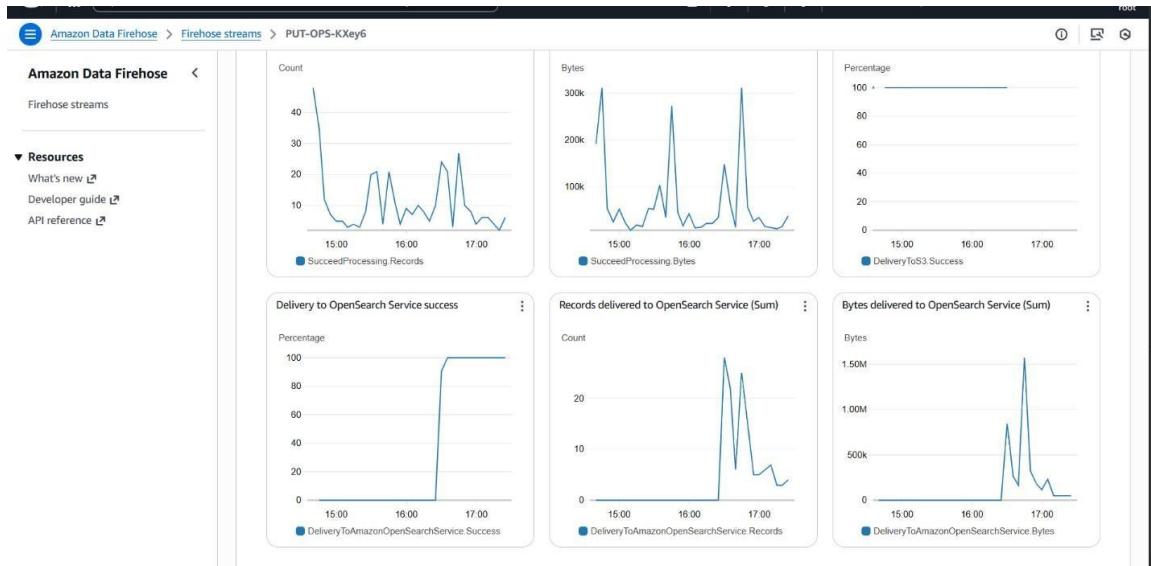
This improved searchability and consistency.

8. Pipeline Validation and Testing

The logging pipeline was validated end-to-end.

After resolving VPC network restrictions, log ingestion was confirmed in OpenSearch.

Indexed records were verified to be searchable and complete.



9. Workload Log Integration

9.1 AWS Lambda Logging

Lambda workloads were verified to emit logs to CloudWatch.

Test invocations confirmed correct log group and stream creation.

The screenshot shows the CloudWatch Log Management interface. The left sidebar has sections for CloudWatch, Favorites and recents, Alarms, AI Operations, GenAI Observability, Application Signals (with a 'New' badge), Infrastructure Monitoring, and Logs. Under Logs, there are sub-options for Log Management (with a 'New' badge), Log Anomalies, Live Tail, Logs Insights, and Contributor Insights. The main area is titled 'Log events' and contains a search bar with placeholder 'Filter events - press enter to search'. Below the search bar are time range buttons: Clear, 1m, 30m, 1h, 12h, Custom, and UTC timezone. A 'Display' dropdown menu is open. The log table has columns for 'Timestamp' and 'Message'. The first few log entries are:

- 2026-01-15T18:20:37.669Z END RequestId: bac6f639-8fa5-49d1-8c9c-d30d849c8aab
- 2026-01-15T18:20:37.669Z REPORT RequestId: bac6f639-8fa5-49d1-8c9c-d30d849c8aab Duration: 38.07 ms Billed Duration: 39 ms Memory Size: 128 MB Max Memory Use: 128 MB
- 2026-01-15T18:22:48.077Z START RequestId: 176329b2-0b0c-4428-ab46-d821c67c2a2a Version: \$LATEST
- 2026-01-15T18:22:48.138Z END RequestId: 176329b2-0b0c-4428-ab46-d821c67c2a2a
- 2026-01-15T18:22:48.139Z REPORT RequestId: 176329b2-0b0c-4428-ab46-d821c67c2a2a Duration: 52.01 ms Billed Duration: 53 ms Memory Size: 128 MB Max Memory Use: 128 MB
- 2026-01-15T18:23:56.161Z START RequestId: fcf351bb-c644-4e67-b81f-13a3a53038e0 Version: \$LATEST
- 2026-01-15T18:23:56.194Z END RequestId: fcf351bb-c644-4e67-b81f-13a3a53038e0
- 2026-01-15T18:23:56.194Z REPORT RequestId: fcf351bb-c644-4e67-b81f-13a3a53038e0 Duration: 33.01 ms Billed Duration: 34 ms Memory Size: 128 MB Max Memory Use: 128 MB
- 2026-01-15T18:23:56.194Z REPORT RequestId: fcf351bb-c644-4e67-b81f-13a3a53038e0 Duration: 33.01 ms Billed Duration: 34 ms Memory Size: 128 MB Max Memory Use: 45 MB

At the bottom right of the log table is a 'Back to top' button.

9.2 ECS Container Logging

ECS Fargate workloads were configured using the awslogs driver.

Application logs were verified in dedicated CloudWatch log groups.

This ensured container workloads were centrally monitored.

The screenshot shows the CloudWatch Log Management interface. The left sidebar has sections for CloudWatch, Favorites and recents, Alarms, AI Operations, GenAI Observability, Application Signals (with a 'New' badge), Infrastructure Monitoring, and Logs. Under Logs, there are sub-options for Log Management (with a 'New' badge), Log Anomalies, Live Tail, Logs Insights, and Contributor Insights. The main area is titled 'Log events' and contains a search bar with placeholder 'Filter events - press enter to search'. Below the search bar are time range buttons: Clear, 1m, 30m, 1h, 12h, Custom, and UTC timezone. A 'Display' dropdown menu is open. The log table has columns for 'Timestamp' and 'Message'. The log entries are:

- 2026-01-15T19:05:40.214Z 10-listen-on-ipv6-by-default.sh: info: Getting the checksum of /etc/nginx/conf.d/default.conf
- 2026-01-15T19:05:40.221Z 10-listen-on-ipv6-by-default.sh: info: Enabled listen on IPv6 in /etc/nginx/conf.d/default.conf
- 2026-01-15T19:05:40.222Z /docker-entrypoint.sh: Sourcing /docker-entrypoint.d/15-local-resolvers.envsh
- 2026-01-15T19:05:40.222Z /docker-entrypoint.sh: Launching /docker-entrypoint.d/20-envsubst-on-templates.sh
- 2026-01-15T19:05:40.234Z /docker-entrypoint.sh: Launching /docker-entrypoint.d/30-tune-worker-processes.sh
- 2026-01-15T19:05:40.235Z /docker-entrypoint.sh: Configuration complete; ready for start up
- 2026-01-15T19:05:40.239Z 2026/01/15 19:05:40 [notice] 1#1: using the "epoll" event method
- 2026-01-15T19:05:40.239Z 2026/01/15 19:05:40 [notice] 1#1: nginx/1.28.1
- 2026-01-15T19:05:40.239Z 2026/01/15 19:05:40 [notice] 1#1: built by gcc 14.2.0 (Debian 14.2.0-19)
- 2026-01-15T19:05:40.239Z 2026/01/15 19:05:40 [notice] 1#1: 05: Linux 5.10.45+245.983.amzn2.x86_64
- 2026-01-15T19:05:40.239Z 2026/01/15 19:05:40 [notice] 1#1: getrlimit(RLIMIT_NOFILE): 65535:65535
- 2026-01-15T19:05:40.240Z 2026/01/15 19:05:40 [notice] 1#1: start worker processes
- 2026-01-15T19:05:40.240Z 2026/01/15 19:05:40 [notice] 1#1: start worker process 28
- 2026-01-15T19:05:40.240Z 2026/01/15 19:05:40 [notice] 1#1: start worker process 28

10. Durability and Backup Strategy

All logs delivered through Firehose are backed up to S3.

Lifecycle policies archive data to Glacier for long-term retention.

This provides durability, disaster recovery, and compliance support.

11. Governance and Documentation

All logging configurations, access policies, and retention standards were documented.

This supported:

- Audit reviews
- Compliance reporting
- Incident investigations
- Operational continuity

12. Outcomes and Impact

This implementation delivered:

- Centralized security visibility
- Improved incident response capability
- Compliance-aligned log retention
- Secure SIEM deployment
- Reliable log delivery
- Reduced operational blind spots

13. Conclusion

I designed and implemented a secure, scalable, and compliant centralized logging and observability platform on AWS.

Through automated log collection, secure analytics, and structured retention policies, this solution supports enterprise security monitoring, investigations, and regulatory requirements.