|  |
| --- |
| Centralized logging and analysis of multi-regional AWS WAF logs |

**Steps**

In this step it shows how the business problem is solved with utmost efficiency by using various techniques to reach the goal.

1. AWS WAF is a web application firewall that is monitoring the HTTP(S) requests from HostBooks platform.
2. WAF collected logs which will provide troubleshooting and root-cause analysis for any kind of exception. For say,
   1. Blacklist IP
   2. IP repudiation
   3. BOT Requests
   4. 403 error
   5. Firewall blocking
   6. IP repudiation limit cross etc.
3. The collected logs would be streamed through Amazon Kinesis Data Firehose.
4. The streamed logs will be stored in S3 which can be later referred for historical analysis.
5. Simultaneously, these logs are sent to Amazon Elasticsearch service for security related analysis.
6. The logs from Elasticsearch service are sent to Kibana for further analysis giving security related insights. It helps to aggregate important findings from multiple regions in a single dashboard.

**Challenges Faced and resolved**

There were following issues with the architecture :

1. Log analysis
2. Centralised logging

2. Storage of real-time logs

3. Optimized searching and analysis

4. Aggregation of non-uniform logs

The issues were solved by using AWS WAF.

**Test cases**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Test Case No.** | **Description** | **Test Data** | **Expected Result** | **Actual Result** | **Status (pass/fail)** |
| #1 | Connection check | Logs | Connection successful | Connection successful | Pass |
| #2 | Log streaming | WAF logs | Seamless flow of logs | Seamless flow of logs | Pass |
| #3 | Centalized logging | WAF logs | Logs centralised at one place | Logs centralised at one place | Pass |

Solution Summary - Training & Handover Notes

**About Challenge(s)**

HostBooks has an online portal, used by accountants and commerce professionals. They use the portal to use services like GST filing and E-way bill payments.

Though they had an on-premise firewall in place but the customer was facing multiple challenges in analyzing the logs generated by the firewalls deployed at multiple locations. AWS cloud service WAF can consolidate firewall data from multiple regions at one place.

The biggest challenge here is the collection of WAF firewall logs from client website for analysis and then combining these insights to be displayed together in a single dashboard as a one-stop solution.

The next challenge is performing log analysis which is critical for understanding the effectiveness of any solution offered, it is valuable for day-to-day troubleshooting and also for long-term understanding of how the application is performing.

Also, while troubleshooting logs from multiple regions, how to perform root cause analysis for exceptional cases like IP repudiation, BOT Requests, 403 error, Firewall blocking is a difficult task.

**Proposed Solution**

MIND discussed the problem with the customer and after analyzing the business problem, it was determined that Amazon Kinesis Data Firehose would fit the business problem. Solution flow proposed consisted of the following steps.

* With the access to full AWS WAF logs, we currently have the ability to analyze all the logs generated by AWS WAF while it’s protects the web applications. In addition, Amazon Kinesis Data Firehose is used to forward these logs to Amazon Simple Storage Service (Amazon S3) for the purpose of archival, and to Amazon Elasticsearch Service for further analysis which is then represented in Kibana as dashboard.
* This allows us to find out in near-real time that which AWS WAF rules are getting triggered, the reason why are they being triggered, and by which request.
* Long-term analysis is also done by creating a historical view of previous logs.
* The Centralized Logging solution offering enables organizations to collect, analyze, and display Amazon WAF logs in a single dashboard.
* These collected logs provide troubleshooting and root-cause analysis for any kind of exception for say Blacklist IP, IP repudiation, BOT Requests, 403 error, Firewall blocking, IP repudiation limit cross etc.
* The offering contains a suite of infrastructure services that deploy a centralized logging solution.
* It uses Amazon Elasticsearch Service (Amazon ES) and Kibana, an analytics and visualization platform that is integrated with Amazon ES, which together results in a unified view of all the log events.
* Amazon Kinesis Data Firehose streams the data coming from WAF to Amazon Elasticsearch Service and concurrently stores the data to S3.
* Then these errors are visualized in Kibana which will use streamed data to perform real-time root cause analysis for exceptions in a customizable, user-friendly dashboard.

**AWS Services used**

* Amazon Kinesis Data Firehose
* Amazon WAF
* AWS ElasticSearch
* Amazon S3
* Kibana

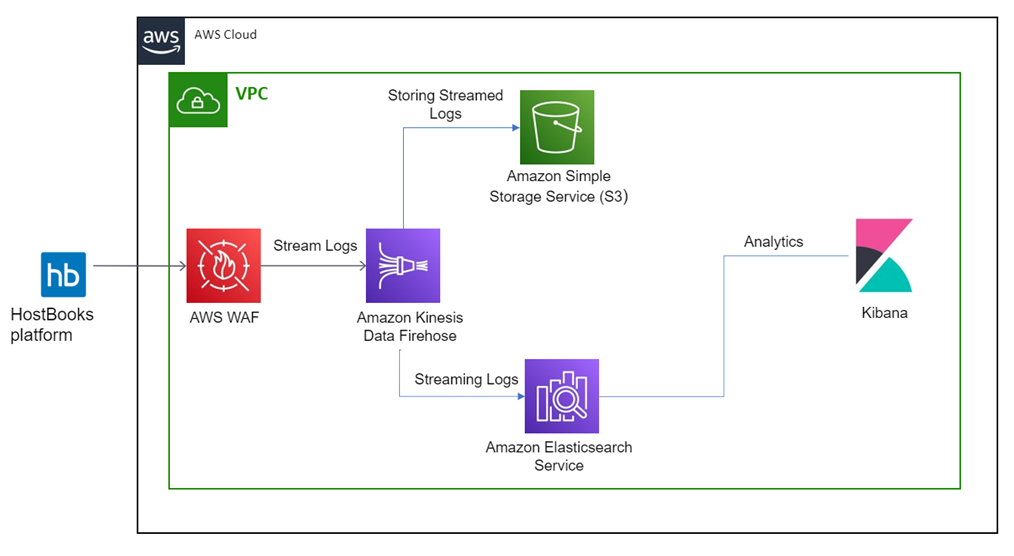
**Solution Outcome**

All in all, we successfully created a solution with the following outcomes.

* Full access to WAF logs in 1 unified location.
* More secure business operations with timely threat detection
* Increased reliability of infrastructure at lower cost & significant improvement in performance to cost ratio

There is increased customer satisfaction which possibly will be translated to more orders from same customer.

**Architecture Diagram**

**How AWS services helped in building the model for sales Forecasting**

**Amazon Kinesis Data Firehose**

Amazon Kinesis Data Firehose is the easiest way to load streaming data into data stores and analytics tools. It can capture, transform, and load streaming data into Amazon S3, Amazon Redshift, Amazon Elasticsearch Service, and Splunk, enabling near real-time analytics with existing business intelligence tools.

**Amazon WAF**

AWS WAF is a web application firewall that lets you monitor the HTTP(S) requests that are forwarded to an Amazon CloudFront distribution, Amazon API Gateway REST API, Application Load Balancer, or AWS AppSync GraphQL API.

**AWS ElasticSearch**

Amazon Elasticsearch Service (Amazon ES) is a managed service that makes it easy to deploy, operate, and scale Elasticsearch clusters in the AWS Cloud. Elasticsearch is a popular open-source search and analytics engine for use cases such as log analytics, real-time application monitoring, and clickstream analysis.

**Amazon S3 to store meta-data**

It is an object storage service that offers industry-leading scalability, data availability, security, and performance. In this solution it.

**Kibana**

Kibana is a popular open source visualization tool designed to work with Elasticsearch. Amazon ES provides an installation of Kibana with every Amazon ES domain.t is an object storage service that offers industry-leading scalability, data availability, security, and performance. In this solution it helped to store raw invoice documents, preprocessed invoice documents and for storing final JSON output files.