Document of Understanding – Requirement

**AWS-Console logs collector**

Prepared by

**Anurag Gundappa**

**Background:**

Currently there is official plugin of graylog which is able to collect AWS logs – service logs, VPC flow logs. The limitation of this plugin is that in order to proper working of this plugin AWS Kinesis needs to be activate. Kinesis is responsible to convert all the CloudWatch logs into continuous stream of data.

# **Amazon CloudWatch**

Amazon CloudWatch is a monitoring and management service that provides data and actionable insights for AWS, hybrid, and on-premises applications and infrastructure resources. With CloudWatch, you can collect and access all your performance and operational data in form of logs and metrics from a single platform. This allows you to overcome the challenge of monitoring individual systems and applications in silos (server, network, database, etc.). CloudWatch enables you to monitor your complete stack (applications, infrastructure, and services) and leverage alarms, logs, and events data to take automated actions and reduce Mean Time to Resolution (MTTR). This frees up important resources and allows you to focus on building applications and business value.

You can use CloudWatch Container Insights to monitor, troubleshoot, and alarm on your containerized applications and microservices. CloudWatch collects, aggregates, and summarizes compute utilization information like CPU, memory, disk, and network data, as well as diagnostic information like container restart failures, to help DevOps engineers isolate issues and resolve them quickly. Container Insights gives you insights from container management services such as Amazon ECS for Kubernetes (EKS), Amazon’s Elastic Container Service (ECS), AWS Fargate, and standalone Kubernetes (k8s).

**Purpose of the document:**

This document explains the need and implementation of AWS logs collector module.

**Functional Requirement:**

Here we are trying to eliminate the usage of AWS Kinesis. There are others ways to bring logs to logger from cloud. AWS cloudwatch is the place from which we need to fetch the logs of all the services and VPC.

1. There can be graylog plugin which will fetch the logs from AWS without using AWS Kinesis stream generator.

2. There can be python module which will carry heavy lifting of log data from AWS. Python module will fetch logs in continuous and can write log file simulteneously. These logs can be deposited into **.log** file and then this ever growing file can to mapped to logger as data input.

The second approach would need to extract information from the logs using logger extractors.

This will form the first part of this development. In the second part there is VSOCbox plugin which will generate events in SIEM from these RAW logs.

This way we can maintain the AWS logs in both SIEM and in CASB.

**Impact Analysis:**

Impacted Area of CASB -

Inputs - > AWS

Impacted Area of SIEM is-

Events→Security Events

**Benefits (Value Add with Change):**

1. 1. SIEM can now handle AWS logs and generate events from those logs with this set of plugins

**Condition/Constraint:**

1. 1. There will be dependency on some of the AWS services. Currently 30 AWS services can send logs to Amazon CloudWatch.
2. There is a need for setting up one cloudtrail Trail, one S3 bucket and one Cloudwatch log group
3. Trail need to be configured to send events to S3 bucket and Cloudwatch log group as simulteneously.

**Reference:**

1. <https://github.com/sherzberg/graylog-plugin-s3>
2. <https://github.com/Graylog2/graylog-plugin-aws>
3. <https://dzone.com/articles/log-collection-graylog-aws>
4. <https://techblog.bozho.net/log-collection-with-graylog-on-aws/>
5. <http://www.neeleshgurjar.co.in/manage-aws-api-logs-with-cloudtrail-and-graylogs/>
6. <https://operatingops.org/2019/09/21/cloudwatch-logs-structured-as-json-with-python-lambda-functions/>