Google Kubernetes Engine (GKE) Monitoring and Logging Solution

Overview

This document outlines the setup and configuration of monitoring and logging solutions for Google Kubernetes Engine (GKE) using managed Prometheus for metrics monitoring and Elastic Stack (EFK) for application logs. Additionally, it covers the creation of dashboards, alerts, and notifications on Google Workspace (formerly G Suite) Teams channel for monitoring critical metrics and logs.

Components

Managed Prometheus: Used for collecting and monitoring metrics from GKE clusters.

Elastic Stack (EFK):

- Elasticsearch: Stores and indexes application logs.
- Fluentd: Collects logs from Kubernetes pods and sends them to Elasticsearch.
- Kibana: Visualizes and explores logs stored in Elasticsearch.

Google Workspace Teams Channel: Used for team communication and receiving alerts and notifications.

Setup and Configuration

1. Managed Prometheus Setup

- Enable Managed Prometheus on your GKE cluster.
- Configure Prometheus to scrape metrics from your applications and Kubernetes components.
- Create custom Prometheus recording and alerting rules for metrics such as CPU, memory, disk space, latency, node uptime, and node downtime.

2. EFK Logging Solution Setup

- Deploy Elasticsearch, Fluentd, and Kibana on your GKE cluster.
- Configure Fluentd to collect logs from all pods running in the cluster and forward them to Elasticsearch.

- Set up index patterns and visualization dashboards in Kibana for log analysis and monitoring.

3. Dashboard Creation

- Utilize and create Google custom dashboards to visualize metrics collected by Prometheus.
- Include graphs and charts for CPU, memory, disk space, latency, node uptime, and node downtime.
- Integrate Kibana visualizations for log analysis directly into Elastic CLoud dashboards for holistic monitoring.

4. Alerting Setup

- Configure Google alert to send alerts based on predefined rules.
- Define alerting rules for critical metrics exceeding thresholds (e.g., CPU utilization > 90%, memory usage > 85%, latency > 500ms).
- Set up alerts for node uptime and downtime.
- Integrate alert notifications with Google Workspace Teams channel using appropriate webhooks or APIs.

5. Team Communication Channel

- Create a dedicated Google Workspace Teams channel for monitoring and incident management.
- Invite relevant team members to the channel for collaboration.
- Configure alert notifications to be sent to the Teams channel for real-time incident response.

Implementation Steps

Enable Managed Prometheus:

- Follow Google Cloud documentation to enable Managed Prometheus for your GKE cluster.

Deploy EFK Logging Solution:

- Use Helm charts or manifests to deploy Elasticsearch, Fluentd, and Kibana onto the GKE cluster.

- Configure Fluentd to collect logs from Kubernetes pods and forward them to Elasticsearch.

Create Grafana Dashboards:

- Log in to Grafana and create custom dashboards for monitoring GKE metrics.
- Integrate Kibana visualizations into Grafana dashboards for unified monitoring.

Set Up Alerts:

- Define alerting rules in Prometheus Alertmanager for critical metrics.
- Configure alert notification receivers to send alerts to Google Workspace Teams channel.

Create Google Workspace Teams Channel:

- Create a dedicated channel for monitoring and incident response.
- Configure incoming webhooks or API integrations to receive alert notifications.

Test and Iterate:

- Test the monitoring and logging setup thoroughly to ensure alerts are triggered accurately.
- Iterate on dashboards and alerts based on feedback and evolving requirements.

Conclusion

Implementing a comprehensive monitoring and logging solution for GKE clusters using Managed Prometheus and EFK provides visibility into system performance and application behavior. By creating custom dashboards, setting up alerts, and integrating with team communication channels like Google Workspace Teams, teams can effectively monitor, troubleshoot, and respond to incidents in real-time, ensuring the reliability and stability of their GKE environments.