End-to-End Log Collection for Ephemeral Kubernetes Jobs using Filebeat Sidecar

# 1. Overview

This document provides an end-to-end solution for collecting logs from ephemeral Kubernetes job pods using Filebeat sidecar containers. Since logs are not stored in hostPath volumes and are lost once the pod exits, the sidecar pattern is the most effective method to collect logs in real time before pod termination.

# 2. Sidecar Pattern Explanation

Each job pod includes two containers:  
1. The main container runs the application or job.  
2. The Filebeat sidecar container tails log files generated during job execution and forwards them to Logstash.

# 3. Kubernetes Job with Filebeat Sidecar

Example YAML for a Kubernetes job with Filebeat sidecar:

apiVersion: batch/v1  
kind: Job  
metadata:  
 name: example-job  
spec:  
 template:  
 metadata:  
 labels:  
 app: example  
 spec:  
 containers:  
 - name: job-runner  
 image: your-job-image  
 command: ["./run-job.sh"]  
 volumeMounts:  
 - name: logs  
 mountPath: /appvol/application/logs  
 - name: filebeat  
 image: <your-nexus>/filebeat:7.18.2  
 args: ["-c", "/etc/filebeat/filebeat.yml", "-e"]  
 volumeMounts:  
 - name: filebeat-config  
 mountPath: /etc/filebeat  
 - name: logs  
 mountPath: /appvol/application/logs  
 restartPolicy: Never  
 volumes:  
 - name: logs  
 emptyDir: {}  
 - name: filebeat-config  
 configMap:  
 name: filebeat-config

# 4. Filebeat ConfigMap

Create the following ConfigMap to configure Filebeat:

apiVersion: v1  
kind: ConfigMap  
metadata:  
 name: filebeat-config  
data:  
 filebeat.yml: |  
 filebeat.inputs:  
 - type: log  
 enabled: true  
 paths:  
 - /appvol/application/logs/\*  
 output.logstash:  
 hosts: ["logstashserv1:5044", "logstashserv2:5044"]  
 loadbalance: true  
 ssl.enabled: false  
 logging.level: info

# 5. Deployment Steps

1. Apply the ConfigMap:  
 kubectl apply -f filebeat-config.yaml  
  
2. Create the Job YAML and apply it:  
 kubectl apply -f example-job.yaml  
  
3. Monitor logs from the Filebeat sidecar or verify in Logstash output.

# 6. Summary

Using a sidecar container ensures that logs are captured and sent to Logstash even if the main job pod exits quickly. This setup is ideal for non-persistent workloads where logs are stored only in the ephemeral container filesystem.