

Contribution ID: 685 Contribution code: **contribution ID 685**Type: **Poster**

Evaluation and Implementation of Various Persistent Storage Options for CMSWEB Services in Kubernetes Infrastructure at CERN

This talk summarizes the various storage options that we implemented for the CMSWEB cluster in Kubernetes infrastructure. All CMSWEB services require storage for logs, while some services also require storage for data. We also provide a feasibility analysis of various storage options and describe the pros/cons of each technique from the perspective of the CMSWEB cluster and its users. In the end, we also propose recommendations according to the service needs. The first option is the CephFS which can be mounted multiple times across various clusters and VMs and works very well with k8s. We use it both for data and the logs. The second option is the Cinder volume. It is the block storage that runs the filesystem on top of it. It can only be attached to one instance at a time. We use this option only for the data. The third option is S3 storage. It is object storage that offers a scalable storage service that can be used by applications compatible with the Amazon S3 protocol. It is used for the logs. For S3, we explored two mechanisms. For the first scenario, we consider fluentd that runs as a sidecar container in the service pods and sends logs to S3 bucket. For the second scenario, we considered filebeat that runs as a sidecar container in the service pod and scaps those logs to fluentd which runs as a daemonset in each node and sends those logs to S3 in the end. The fourth option is EOS. We configured EOS inside the pods of the CMSWEB services. The fifth option that we explored is to use dedicated VMs that have Ceph volume attached to them. In EOS and VM, the logs from the service pods are sent to EOS/VM using the rsync approach. The last option is to send service logs to Elasticsearch. It has been implemented using fluentd that runs as a daemonset in each node. In parallel to the sending logs to S3 fluentd also sends those logs to the Elasticsearch infrastructure at CERN.

Significance

References

Speaker time zone

No preference

Primary authors: COLLABORATION, CMS; IMRAN, Muhammad (National Centre for Physics (PK))**Presenter:** IMRAN, Muhammad (National Centre for Physics (PK))**Session Classification:** Posters: Raspberry**Track Classification:** Track 1: Computing Technology for Physics Research