Kubernetes Activities at UTA

Armen Vartapetian
University of Texas, Arlington

US ATLAS Computing Facility Meeting, July 22, 2020

Introduction

- Kubernetes (k8s) is container orchestration system for automating computer application deployment, scaling, and management.
- It is a flexible and resilient platform for automation and orchestration of containerized applications on practically any computing infrastructure, for running services and delivering resources.
- Kubernetes got quite popular in recent years.
- Orchestration of containers in a Kubernetes cluster can also be used for the purpose of job scheduling and running.
- Due to that there is a growing interest in using Kubernetes cluster as a computing site.

Kubernetes as a batch system

- Or in other words replacing batch system with Kubernetes.
- Since ATLAS has already migrated to a fully container-based workload, it is simpler and more natural to provide container-native infrastructure at sites.
- That reduces a number of layers in the software and service stack of a traditional grid computing site.
- A job is basically replaced by the concept of a pod, and Kubernetes ensures that a pod terminates successfully.
- That relieves Harvester from having to track the status of individual pods/jobs, and to resubmit ones that fail due to node issue.
- Development of a special instance of Harvester, interfacing ATLAS workload submission engine PanDA with Kubernetes, to directly submit and monitor the status of containerized jobs.
- A dedicated K8s Harvester node at CERN, fully operational to run ATLAS jobs at Kubernetes clusters.

Hardware for Kubernetes Cluster

- Currently the UTA Kubernetes cluster is setup on a group of test machines, which are used for evaluating software applications and updates before they are implemented in the main grid cluster.
- We started with fraction of those machines, in total 80 cores of worker nodes right now, and the size will grow up to 250 cores, once we include all of them.
- Those test resources include all the compute models we have in the SWT2 grid cluster, so that represents a good testbed to investigate the K8s behavior trends on various flavor of worker nodes.
- At the same time we are waiting for the UTA_SWT2 cluster to move to UTA campus, which have got delayed due to current COVID-19 situation.
- That cluster has about 2200 cores and significant fraction of that compute we plan to convert to K8s production cluster.

Kubernetes Cluster Installation, Status

- The UTA Kubernetes cluster was deployed on bare metal using the kubeadm method.
- The cluster is showing healthy, and running uninterrupted, since the initial deployment, during the last 40 days.
- As already mentioned, more worker nodes will be added to the cluster shortly.
- And some automation will be implemented, that the cluster can be fully rebuilt from scratch within minutes.
- We will also investigate how to maintain and operate (switching from one to another) both cluster flavors (batch and k8s) with optimal effort on the same test compute resources.

Coming Next

- Setup the UTA Kubernetes cluster to run ATLAS production jobs, through the dedicated K8s Harvester interface.
- Do the homework on test/development cluster and then move to production one.
- Learn from running a variety of ATLAS production workloads and workflows with Kubernetes - understand inefficiencies, bottlenecks, work on optimization.
- Carry out maximum tuning to optimize the cluster performance, and see whether the results indicate that we can benefit from the Kubernetes model.