IMPROVE HORIZONTAL POD AUTOSCALING IN CONTAINER ORCHESTRATION TO ADOPT FREQUENT OSCILLATION IN SERVICE REQUESTS

Divya Premanantha

A dissertation submitted in partial fulfilment of the requirement for
Master of Science degree in Computer Science

Department of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2021

Abstract

Containers are lightweight stand-alone self-contained units that package their dependencies together for fine-grained resource sharing. It is a platform as a service that uses OS-level virtualization. Due to this managing container lifecycle by automating scaling, scheduling and managing the health of the container has become critical tasks of a container as a service platform such as Kubernetes container orchestrater.

A flexible infrastructure is required for applications with dynamic workloads to leverage performance measures and minimize resource costs. Request with various fluctuation patterns has an impact on autoscaling the deployment to cater for the resource requirements. This project finds a solution to adopt frequent service oscillation during service requests into consideration when performing horizontal pod autoscaling in Kubernetes.

Keywords: HPA, Kubernetes, Fluctuation, Autoscaling