Deploy, scale, and update your website with Google Kubernetes Engine (GKE)

About this codelab

subject Last updated Jun 2, 2022

account_circle Written by Mike Verbanic, Kevin Holbrook

1. Introduction

Running websites and applications is hard.

Things go wrong when they shouldn't, servers crash, increase in demand causes more resources to be utilized, and making changes without downtime is complicated and stressful.

Imagine a tool that could help you do all that and even allow you to automate it! With GKE, all of that is not only possible, it's easy! In this codelab, you assume the role of a developer running an ecommerce website for a fictional company—Fancy Store. Due to problems with scaling and outages, you're tasked with deploying your application to GKE!

The exercises are ordered to reflect a common cloud developer's experience:

1. Create a GKE cluster.

2. Create a Docker container.

3. Deploy the container to GKE.

4. Expose the container via a service.

5. Scale the container to multiple replicas.

6. Modify the website.

7. Roll out a new version with zero downtime.

Architecture diagram

What you'll learn

How to create a GKE cluster

How to create a Docker image

How to deploy Docker images to Kubernetes

How to scale an application on Kubernetes

How to perform a rolling update on Kubernetes

Prerequisites

- A Google Account with administrative access to create projects or a project with a project-owner role

- A basic understanding of Docker and Kubernetes (If you lack a basic understanding, then please review Docker and Kubernetes now.)