

LIST OF FIGURES

Fig 2.1 Main page of the Kubernetes Engine in Google Cloud Platform.....	7
Fig 2.2 Diagram of a CA process scaling up [10].....	8
Fig 2.3 Diagram of how Kubernetes VPA works in an example [11].....	9
Fig 2.4 Design for the Kubernetes HPA example [12].....	10
Fig 2.5 Example of a Slack in a Kubernetes Node [13].....	11
Fig 2.6 Resource report for applications taken from Zalando [13].....	11
Fig 2.7 NGINX web page deployed onto the cluster with a custom template.....	13
Fig 3.1 Proposed block diagram illustrating the system used for this thesis	14
Fig 3.2 Flowchart of how the research is processed	15
Fig 4. 1 k6 script code for CA Test 1	23
Fig 4.2 k6 script code for CA Test 2.....	23
Fig 4.3 GKE comes with a built-in Vertical Pod Autoscaler editor, that allows users to configure accordingly	24
Fig 4.4 The Edit Policy menu allows for admins to edit scaling mode and what resources to control along with request limits.....	25
Fig 4.5 Compute resources are scaled with their requests and limits	25
Fig 4.6 Uses the same code as the CA Test, where the parameters are similar for a fair test.....	26
Fig 4.7 Similar to the previous figure, in order to conduct a fair trial, the same code is used as CA Test 2	26
Fig 4.8 The blue line shows when the CA Test 1 was conducted.....	27
Fig 4.9 Memory requests shown during trial test of both k6 test.....	28
Fig 4.10 K6 results for CA Test 1	28
Fig 4.11 The green line below the graph shows when the CA Test 2 was conducted	29
Fig 4.12 K6 test results for CA Test 2 showing more drops and interrupted iterations due to a failed autoscale	29
Fig 4.13 Shows the memory request in the nodes for the Kubernetes engine	30
Fig 4.14 Cluster status showing an error notification, being unable to scale up due to limitations	30
Fig 4.15 Results showing varying errors during VPA Test 1	31
Fig 4.16 Results from VPA Test 1	31
Fig 4.17 Shows the CPU and Memory Usage for VPA Test 1	32
Fig 4.18 Shows pods terminating after new VPA policies are enabled	32
Fig 4.19 Google Cloud Platform Pod Autoscaling Events	32
Fig 4.20 While upscaling, pods needed some time to redeploy so errors occurred	33

Fig 4.21 Shows the results of VPA Test 2 with errors shown, possibly due to pod termination and limits 33

Fig 4.22 The red line shown in the figure represents the time when VPA Test 2 was conducted 34