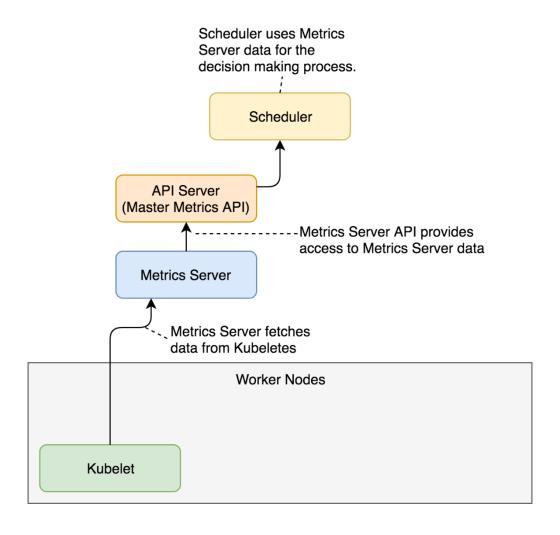
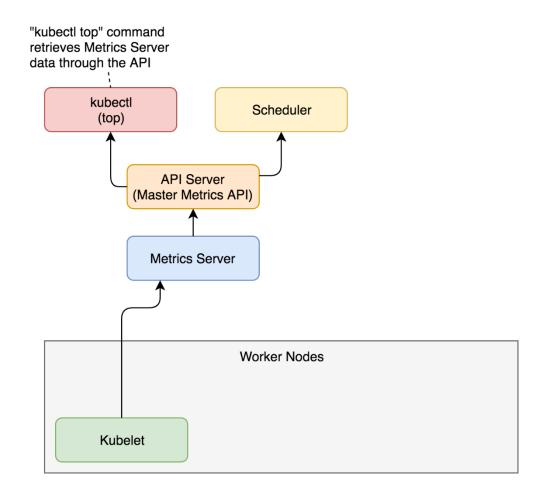
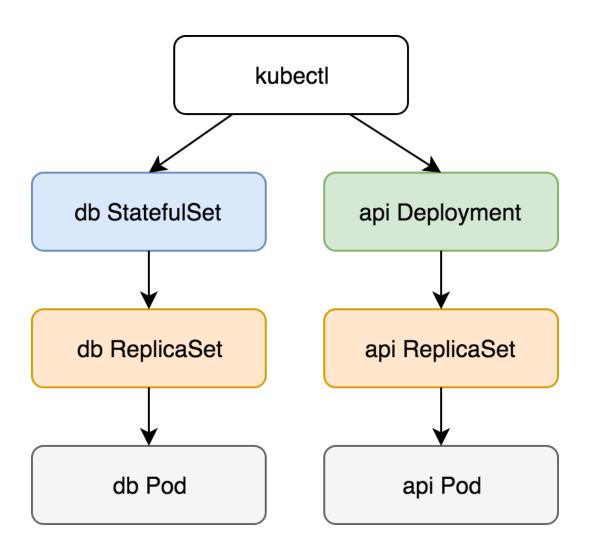
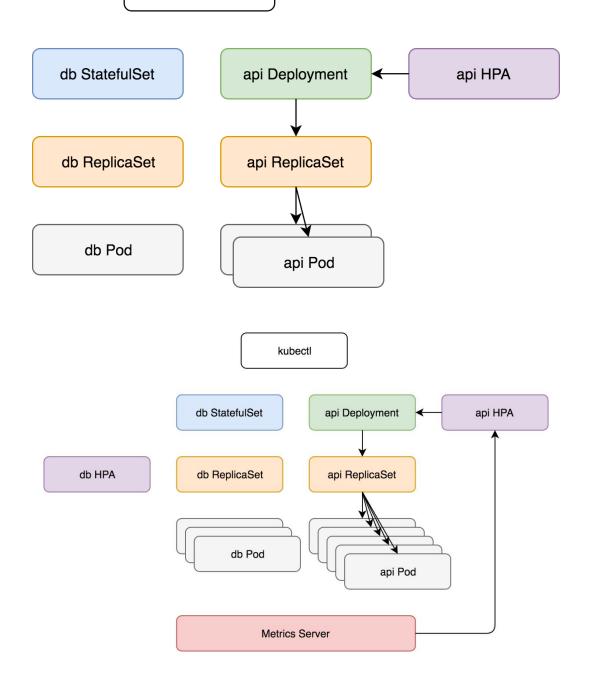
Chapter 1: Autoscaling Deployments and StatefulSets Based on Resource Usage



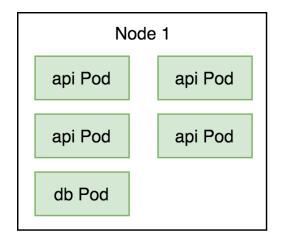


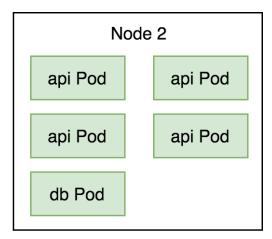


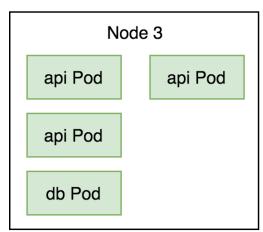
kubectl



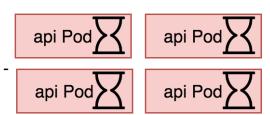
Chapter 2: Auto-scaling Nodes of a Kubernetes Cluster

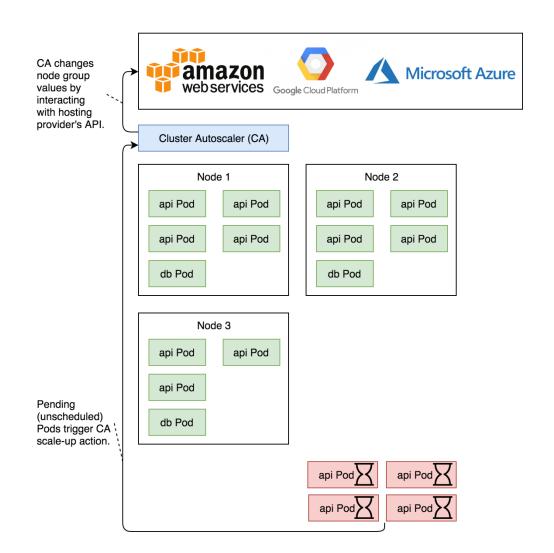


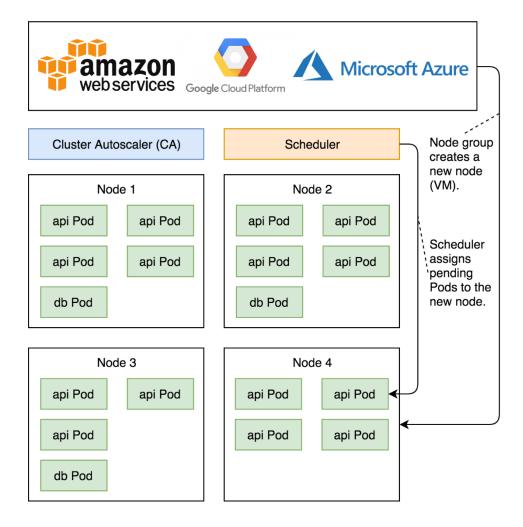


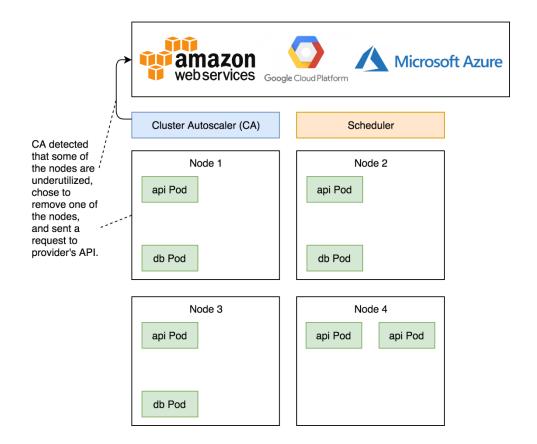


Pods that could not be scheduled due to insufficient capacity of the cluster are in the pending state.

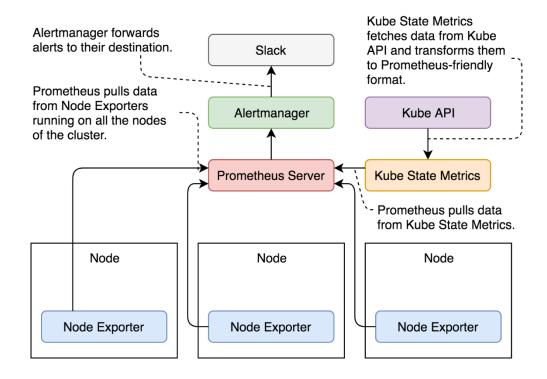




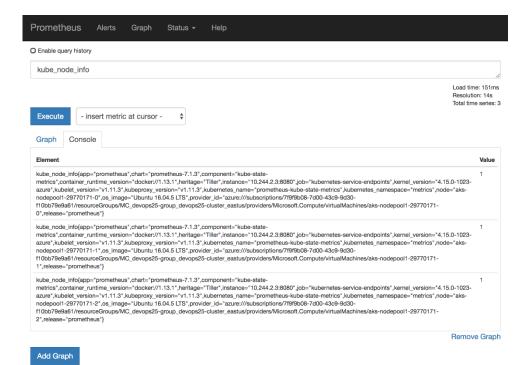




Chapter 3: Collecting and Querying Metrics and Sending Alerts



Targets All Unhealthy kubernetes-apiservers (0/1 up) show less Last Scrape https://172.31.2.21:443/metrics DOWN Instance="172.31.2.21:443" Get https://172.31.2.21:443/metrics: dial tcp 172.3 1.2.21:443: I/o timeout 59.345s ago kubernetes-nodes (3/3 up) show less Endpoint Labels Last Scrape Error Bestpool-"nodepool" beta kubernetes Jo. and 42.233s ago "amotiv" beta kubernetes Jo. and 42.233s ago "amotiv" beta kubernetes Jo. Instanco, type-"8to noder Joža" beta kubernetes Jo. on-"flord" heli noder Joža" beta kubernetes Jo. on-"flord" heli noder Joža vojeknetes Jo. on-smitus" heli noder Joža vojeknetes Joža von-smitus" heli noder Joža vojeknetes Joža von-smitus" instance on-"aks-nodepool "2077/017-10" kubernetes joznes von-duster-"Nd-Gevopooli"-gova devoposit"-busta estatus" kubernetes Joža vojeknetes Joža vojeknete Joža vojeknetes Joža vojeknete https://kubernetes.default.svc:443/api/v1/n odes/aks-nodepool1-29770171-0/proxy/m etrics agamtpool-"nodepool" | beta Jubamentee Jo arch | 16:142s ago "am654" | beta Jubamentee Jo Instance, type-"8te noder (Bai" | beta Jubamentee Jo Instance, type-"8te noder (Bai" | beta Jubamentee Jo open"ersters" | Bail | red, domain Justa Jubamentee Jo open "enters" | final large domain, beta Jubamentee Jo zone-"1" | instance | linear | https://kubernetes.default.svc:443/api/v1/n odes/aks-nodepool1-29770171-1/proxy/m etrics ce="aks-nodepool1-29770171-1" kubernetes_zuze
com_duster="MC_davops25-group_davops25-clusto
r_eastus" kubernetes_lo_hostname="aks-nodepool
1-29770171-1" kubernetes_lo_role="agent" store "nodepool1" beta_kubernetes_io_arch 16.879s ago odes/aks-nodepool1-29770171-2/proxy/m



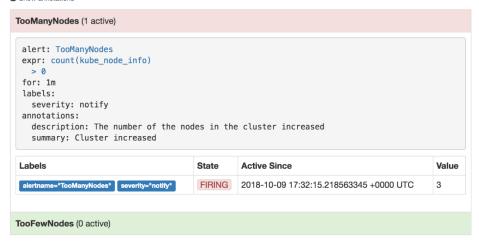
Alerts

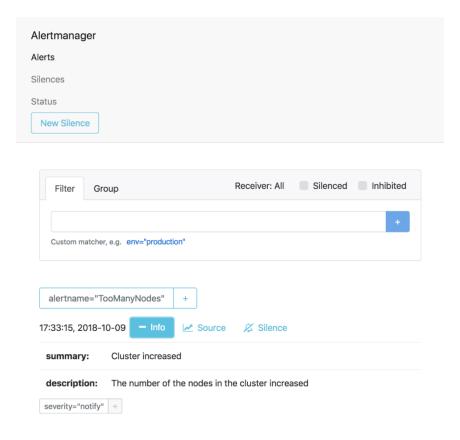
```
TooFewNodes (0 active)

alert: TooFewNodes
expr: count(kube_node_info)
< 1
for: 15m
labels:
severity: notify
annotations:
description: The number of the nodes in the cluster decreased
summary: Cluster decreased

TooManyNodes (0 active)
```

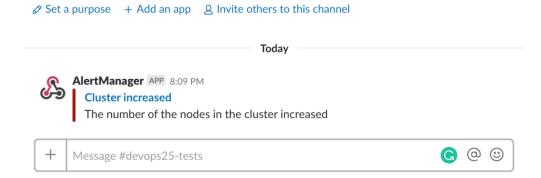
Alerts

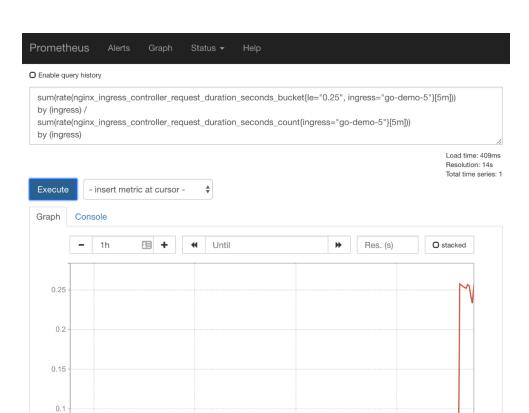




devops25-tests

You created this channel on October 7th. This is the very beginning of the #devops25-tests channel.





0.05

16:15

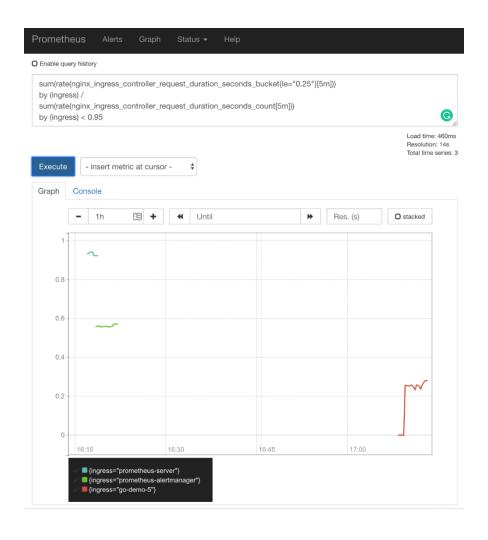
[{ingress="go-demo-5"}

16:30

16:45

Remove Graph

17:00



Alerts

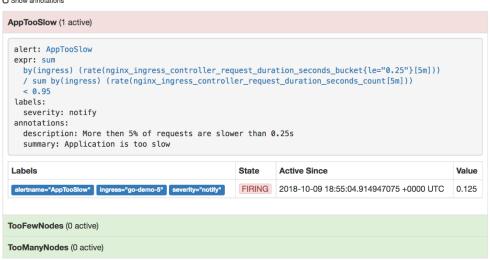
O Show annotations

```
alert: AppTooSlow
expr: sum
by(ingress) (rate(nginx_ingress_controller_request_duration_seconds_bucket{le="0.25"}[5m]))
/ sum by(ingress) (rate(nginx_ingress_controller_request_duration_seconds_count[5m]))
< 0.95
labels:
severity: notify
annotations:
description: More then 5% of requests are slower than 0.25s
summary: Application is too slow

TooFewNodes (0 active)
```

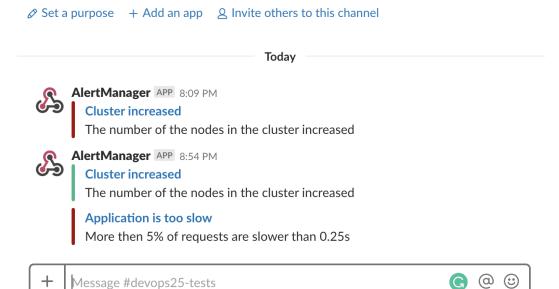
Prometheus Alerts Graph Status - Hel

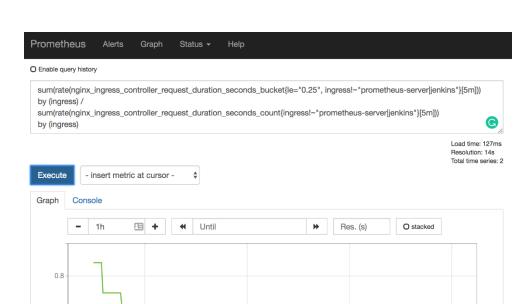
Alerts



devops25-tests

You created this channel on October 7th. This is the very beginning of the #devops25-tests channel.





18:30

0.6

0.4

0.2

0

18:15

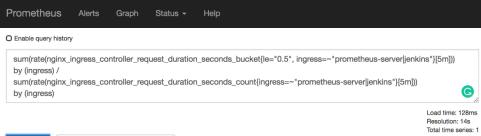
[{ingress="prometheus-alertmanager"}
[{ingress="go-demo-5"}

Remove Graph

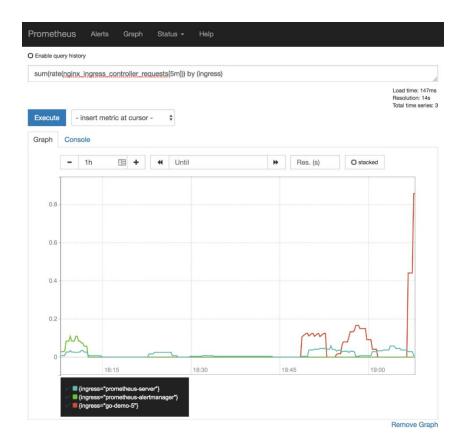
19:00

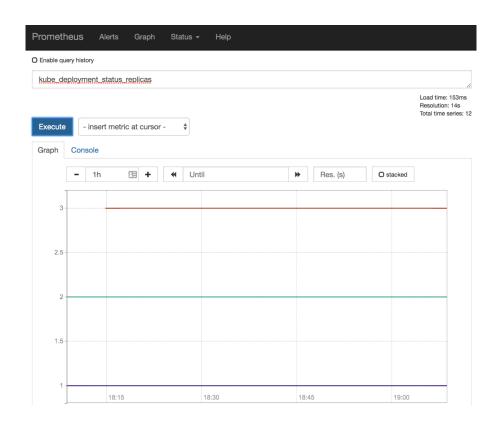
W

18:45



Remove Graph





O Enable query history

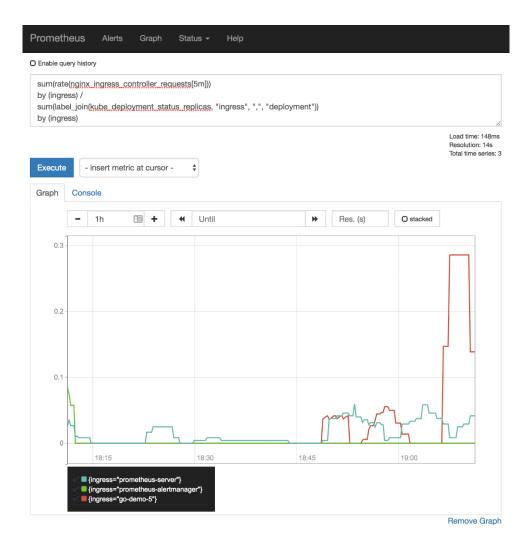
label_join(kube_deployment_status_replicas, "ingress", ",", "deployment")

Load time: 172ms Resolution: 14s Total time series: 12

Execute - insert metric at cursor -

Graph Console

Element	Valu
kube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="go-demo-5",heritage="Tiller",ingress="go-demo-5",instance="10.244.2.3:8080",job="kubernetes-service-endpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="go-demo-5",release="prometheus"}	3
kube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state- metrics",deployment="heapster",heritage="Tiller",ingress="heapster",instance="10.244.2.3:8080",job="kubernetes-service- andpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="kube-system",release="prometheus"}	1
kube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="kube-dns- /20",heritage="Tiller",ingress="kube-dns-v20",instance="10.244.2.3:8080",job="kubernetes-service-endpoints",kubernetes_name="prometheus-kube- state-metrics",kubernetes_namespace="metrics",namespace="kube-system",release="prometheus"}	2
kube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="kubernetes-dashboard",heritage="Tiller",ingress="kubernetes-dashboard",instance="10.244.2.3:8080",job="kubernetes-service-endpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="kube-system",release="prometheus"}	1
kube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="metrics-server",heritage="Tiller",ingress="metrics-server",instance="10.244.2.3:8080",job="kubernetes-service-endpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="kube-system",release="prometheus"}	1
tube_deployment_status_replicas(app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="nginx-ingress-controller",instance="10.244.2.3:8080",job="kubernetes-service-undpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="ingress-nginx",release="prometheus"}	1
tube_deployment_status_replicas{app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="prometheus- ilertmanager",heritage="Tiller",ingress="prometheus-alertmanager",instance="10.244.2.3:8080",job="kubernetes-service- indpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="metrics",release="prometheus"}	1
tube_deployment_status_replicas(app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="prometheus-kube-state-metrics",heritage="Tiller",ingress="prometheus-kube-state-metrics",instance="10.244.2.3:8080",job="kubernetes-service-sundpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="metrics",release="prometheus")	1
kube_deployment_status_replicas(app="prometheus",chart="prometheus-7.1.3",component="kube-state-metrics",deployment="prometheus- bushgateway",heritage="Tiller",ingress="prometheus-pushgateway",instance="10.244.2.3:8080",job="kubernetes-service- andpoints",kubernetes_name="prometheus-kube-state-metrics",kubernetes_namespace="metrics",namespace="metrics",release="prometheus"}	1



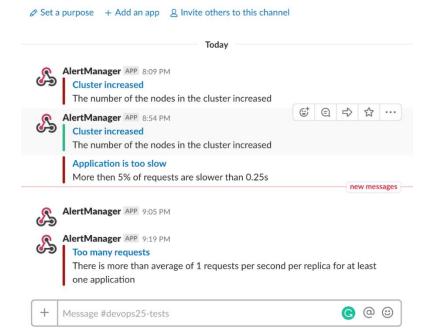
Alerts

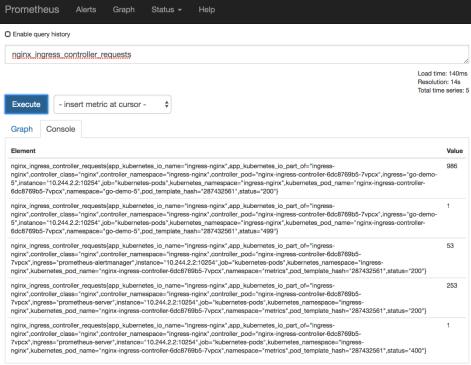
O Show annotations



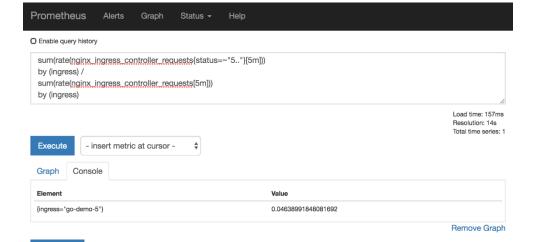
devops25-tests

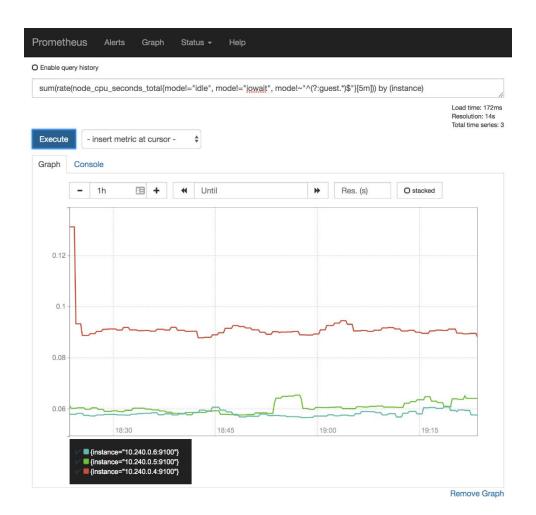
You created this channel on October 7th. This is the very beginning of the #devops25-tests channel.

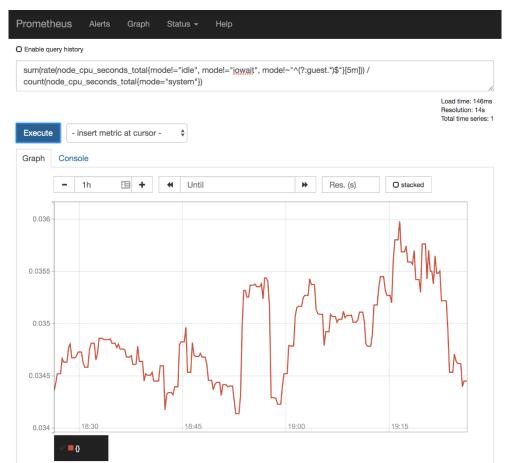




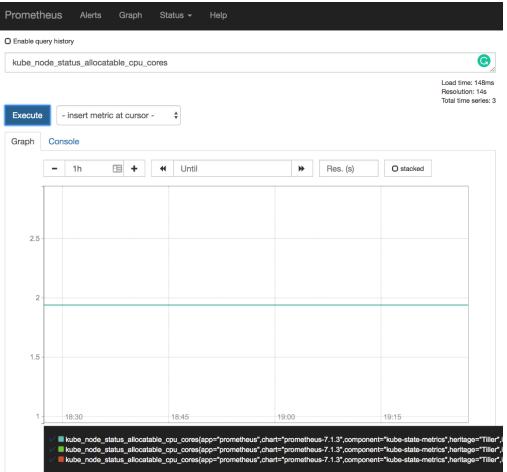
Remove Graph



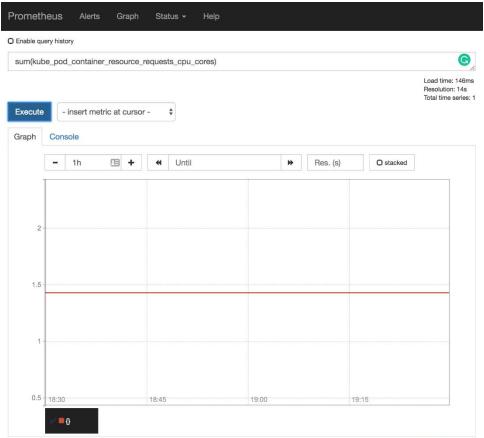




Remove Graph



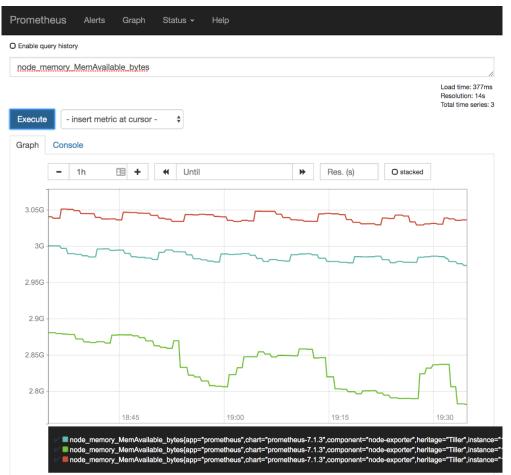
Remove Graph



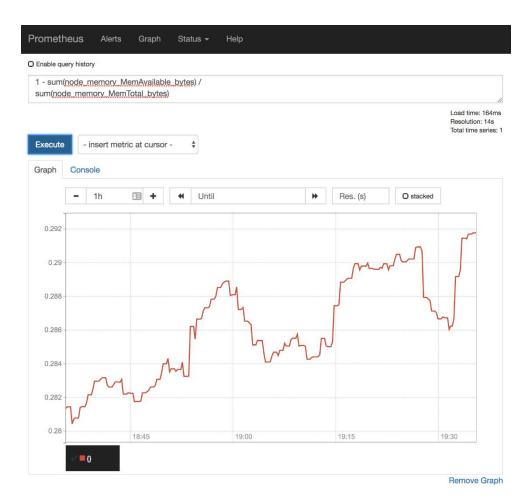
Remove Graph

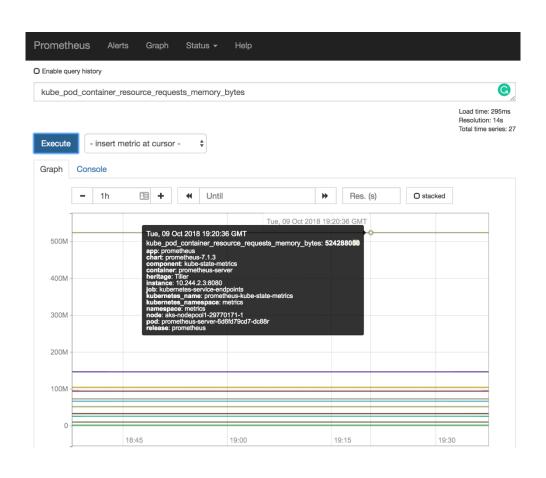
Alerts

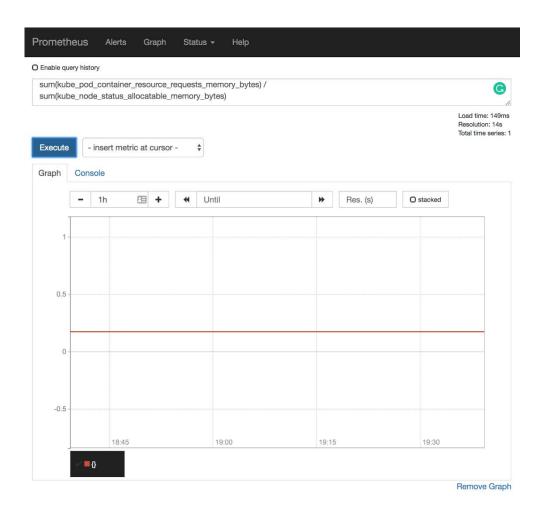
```
AppTooSlow (0 active)
NotEnoughCPU (0 active)
 alert: NotEnoughCPU
 expr: sum(rate(node_cpu_seconds_total{mode!="idle",mode!="iowait",mode!~"^(?:guest.*)$"}[5m]))
/ count(node_cpu_seconds_total{mode="system"}) > 0.9
 for: 30m
 labels:
  severity: notify
 annotations:
   description: CPU usage of the cluster is above 90% summary: There's not enough CPU
TooFewNodes (0 active)
TooManyNodes (0 active)
TooManyRequests (0 active)
TooMuchCPURequested (0 active)
 alert: TooMuchCPURequested
 expr: sum(kube_pod_container_resource_requests_cpu_cores)
   / sum(kube_node_status_allocatable_cpu_cores) > 0.9
  for: 30m
 labels:
   severity: notify
 annotations:
   description: More than 90% of allocatable CPU is requested
   summary: There's not enough allocatable CPU
```



Remove Graph







Alerts

```
AppTooSlow (0 active)
NotEnoughCPU (0 active)
NotEnoughMemory (0 active)
 alert: NotEnoughMemory
 expr: 1
   - sum(node_memory_MemAvailable_bytes) / sum(node_memory_MemTotal_bytes) > 0.9
 for: 30m
 labels:
  severity: notify
 \verb"annotations:"\\
   description: Memory usage of the cluster is above 90%
   summary: There's not enough memory
TooFewNodes (0 active)
TooManyNodes (0 active)
TooManyRequests (0 active)
TooMuchCPURequested (0 active)
TooMuchMemoryRequested (0 active)
 alert: TooMuchMemoryRequested
 expr: sum(kube_pod_container_resource_requests_memory_bytes)
  / sum(kube_node_status_allocatable_memory_bytes) > 0.9
 for: 30m
 labels:
   severity: notify
 annotations:
   description: More than 90% of allocatable memory is requested
   summary: There's not enough allocatable memory
```

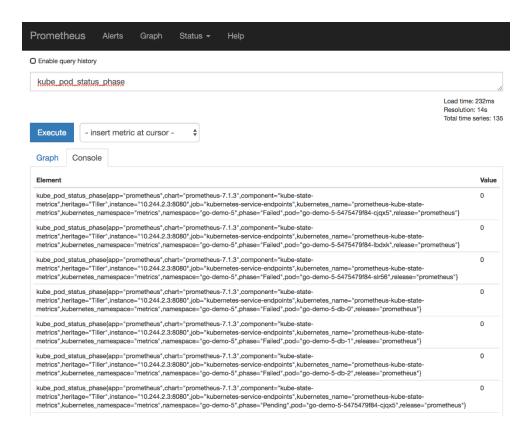
Alerts

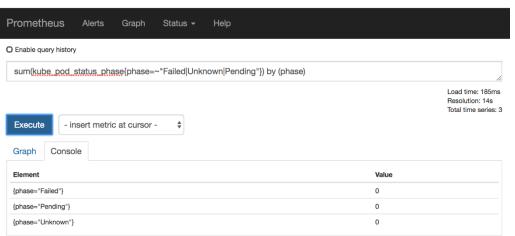
O Show annotations

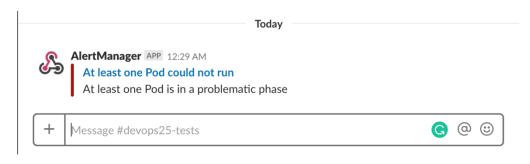
TooMuchCPUAndMemory (3 active)

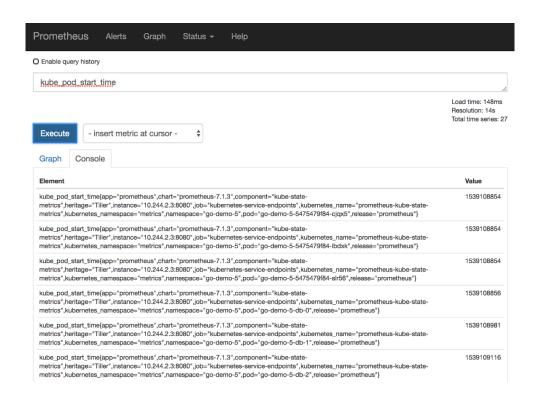
```
alert: TooMuchCPUAndMemory
expr: (sum
  by(instance) (rate(node_cpu_seconds_total{mode!="idle",mode!="iowait",mode!~"^(?:guest.*)$"}[5m]))
  / count by(instance) (node_cpu_seconds_total{mode="system"})) < 0.5 and
  (1 - sum by(instance) (node_memory_MemAvailable_bytes) / sum by(instance) (node_memory_MemTotal_by
tes))
  < 0.5
for: 30m
labels:
  severity: notify
annotations:
  description: Less than 50% of CPU and 50% of memory is used on at least one node
  summary: Too much unused CPU and memory</pre>
```

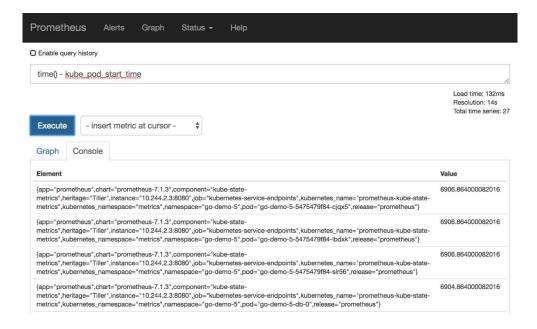
Labels	State	Active Since	Value
alertname="TooMuchCPUAndMemory" Instance="10.240.0.6:9100" severity="notify"	PENDING	2018-10-09 19:44:15.218563345 +0000 UTC	0.027553615303143517
alertname="TooMuchCPUAndMemory" Instance="10.240.0.4:9100" severity="notify"	PENDING	2018-10-09 19:44:15.218563345 +0000 UTC	0.04563008130081312
alertname="TooMuchCPUAndMemory" Instance="10.240.0.5:9100" severity="notify"	PENDING	2018-10-09 19:44:15.218563345 +0000 UTC	0.03303275222793209













AlertManager APP 9:19 PM

Too many requests

There is more than average of 1 requests per second per replica for at least one application



AlertManager APP 9:29 PM

Too many requests

There is more than average of 1 requests per second per replica for at least one application



AlertManager APP 9:48 PM

Application is too slow

More then 5% of requests are slower than 0.25s



AlertManager APP 10:09 PM

Old Pods

At least one Pod has not be updated to more than 90 days

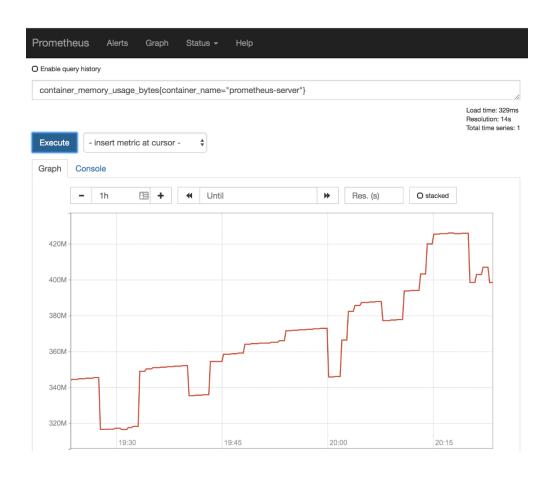


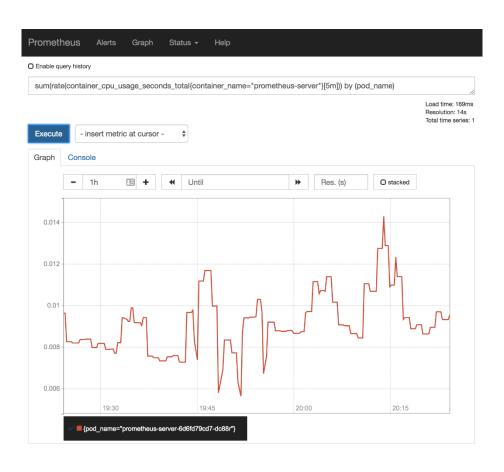
Message #devops25-tests

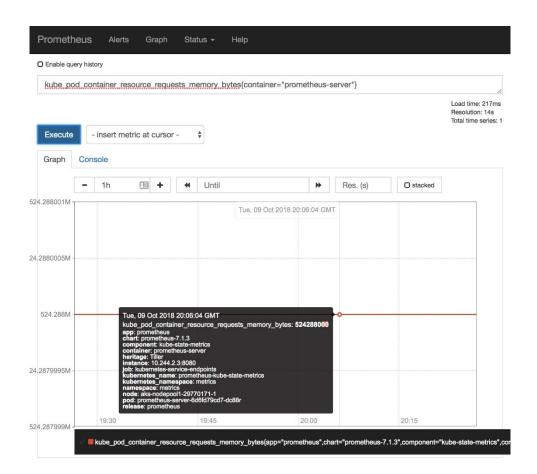


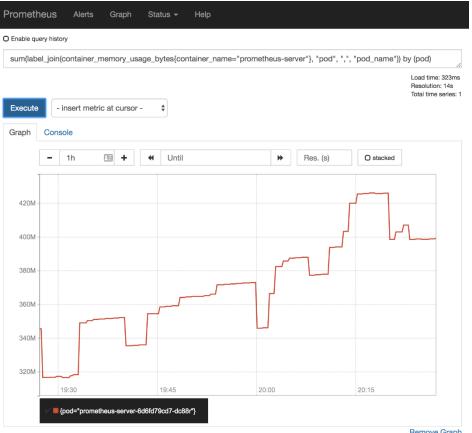




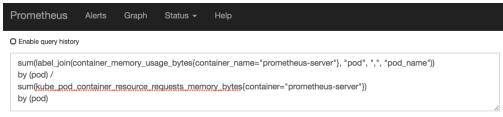




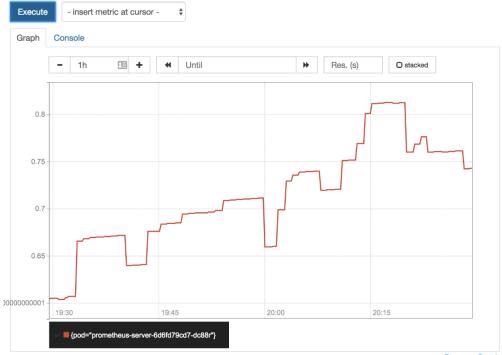




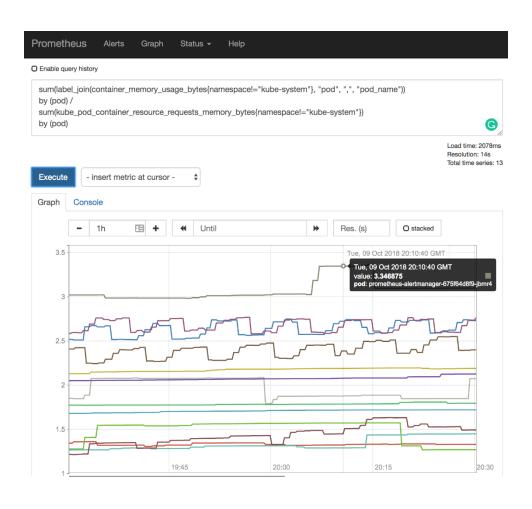
Remove Graph

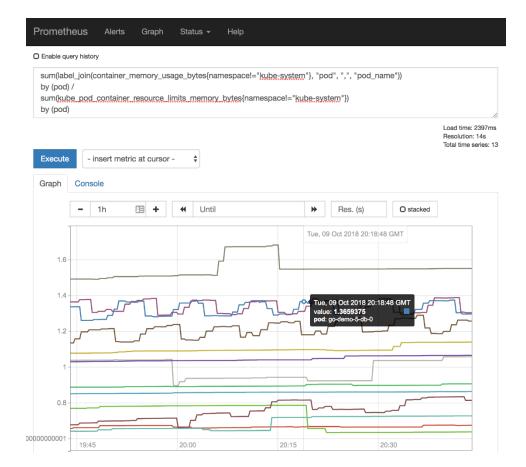


Load time: 204ms Resolution: 14s Total time series: 1

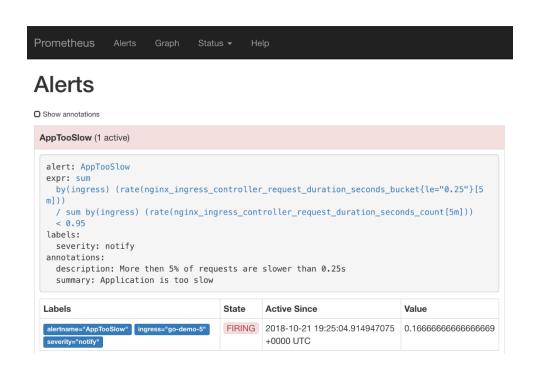


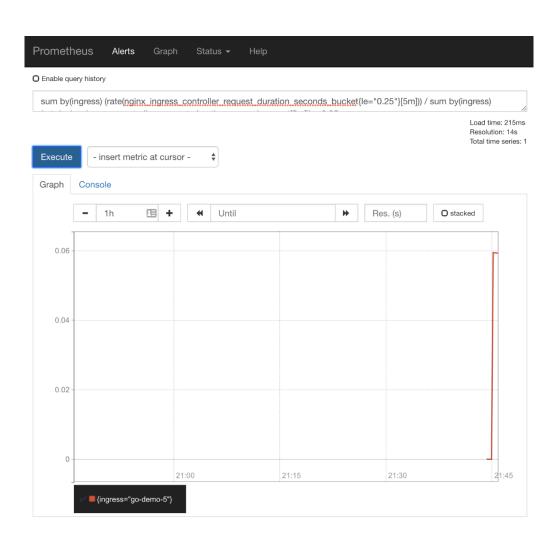
Remove Graph





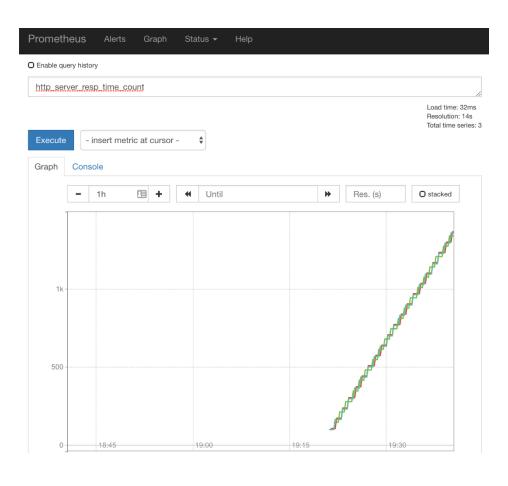
Chapter 4: Debugging Issues Discovered Through Metrics and Alerts

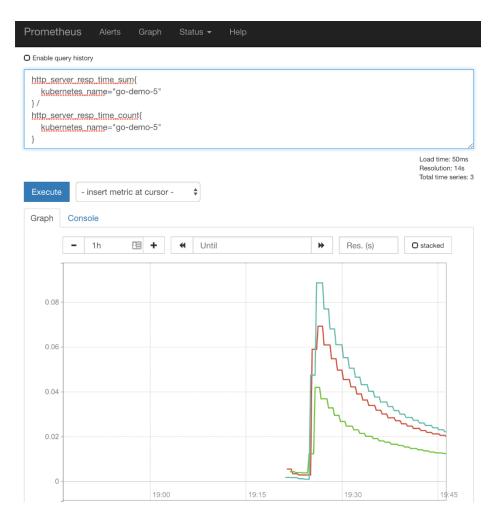




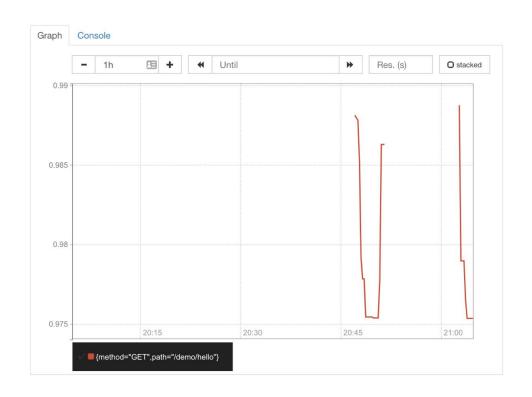
kubernetes-service-endpoints (8/11 up) show less

Endpoint	State	Labels	Last Scrape	Error
http://10.0.2.15:9100/metrics	UP	app="prometheus" chart="prometheus-7.1.3" component="no de-exporter" heritage="Tiller" instance="10.0.2.15:9100" kub ernetes_name="prometheus-node-exporter" kubernetes_namespace="metrics" release="prometheus"	26.001s a go	
http://172.17.0.12:8080/me trics	UP	app="go-demo-5" chart="go-demo-5-0.0.1" heritage="Tiller" instance="172.17.0.12:8080" kubernetes_name="go-demo-5" kubernetes_namespace="go-demo-5" release="go-demo-5"	21.523s a go	
http://172.17.0.13:8080/me trics	UP	app="go-demo-5" chart="go-demo-5-0.0.1" heritage="Tiller" instance="172.17.0.13:8080" k ubernetes_name="go-demo-5" kubernetes_namespace="go-demo-5" release="go-demo-5"	49.894s a go	
http://172.17.0.14:8080/me trics	UP	app="go-demo-5" chart="go-demo-5-0.0.1" heritage="Tiller" instance="172.17.0.14:8080" kubernetes_name="go-demo-5" kubernetes_namespace="go-de	29.902s a go	

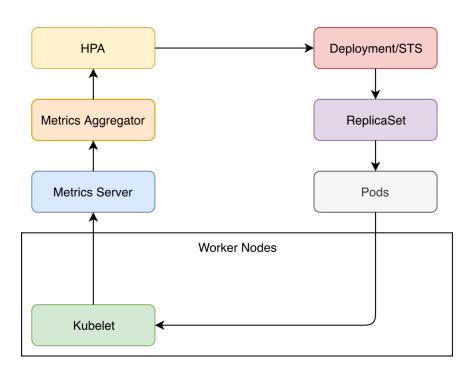


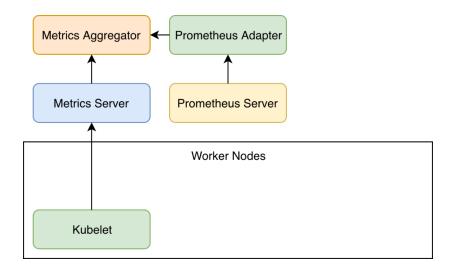


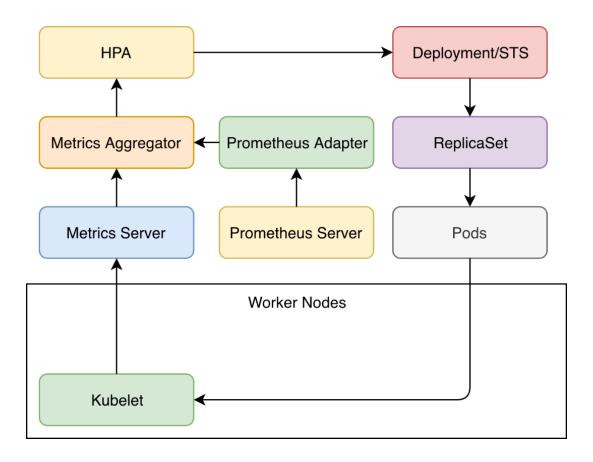




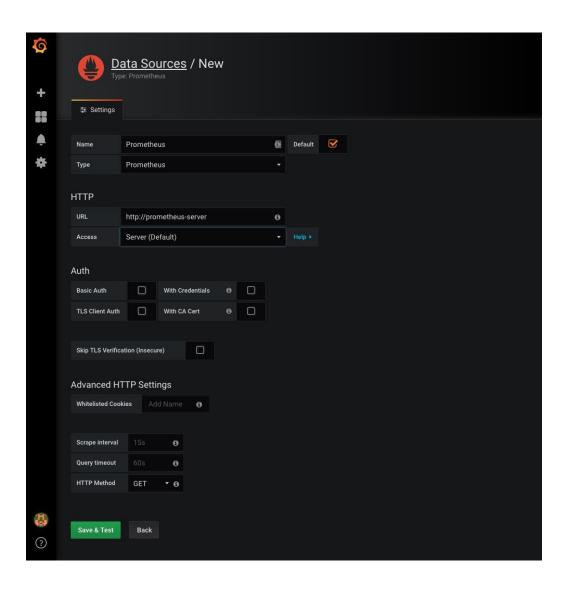
Chapter 5: Extending HorizontalPodAutoscaler with Custom Metrics

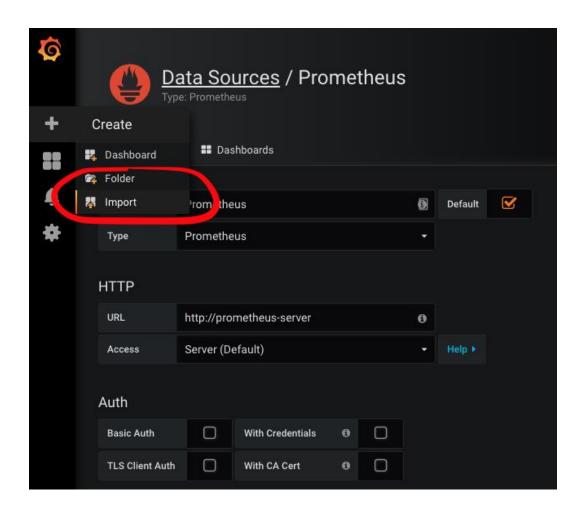


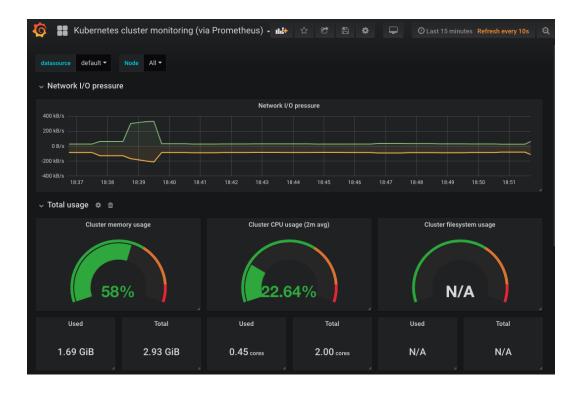


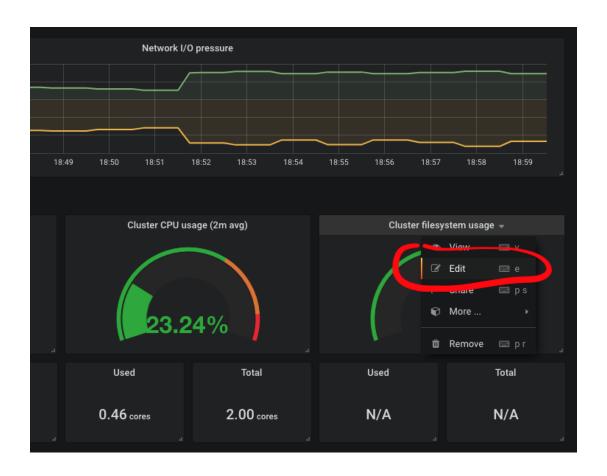


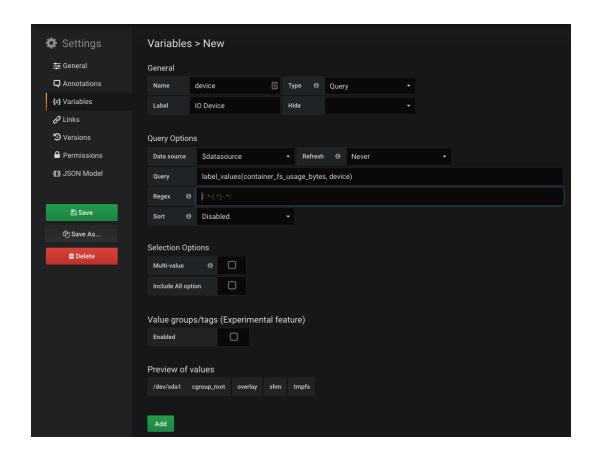
Chapter 6: Visualizing Metrics and Alerts

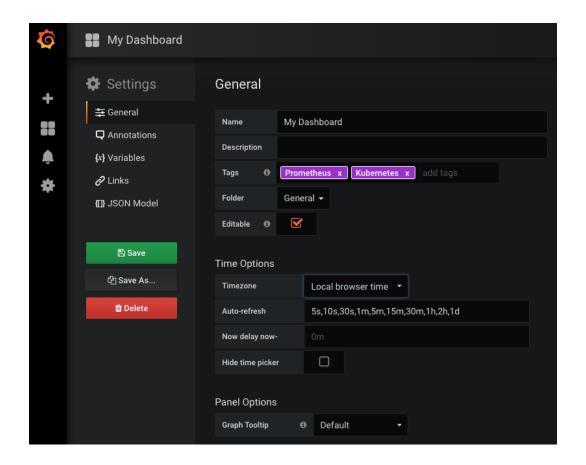


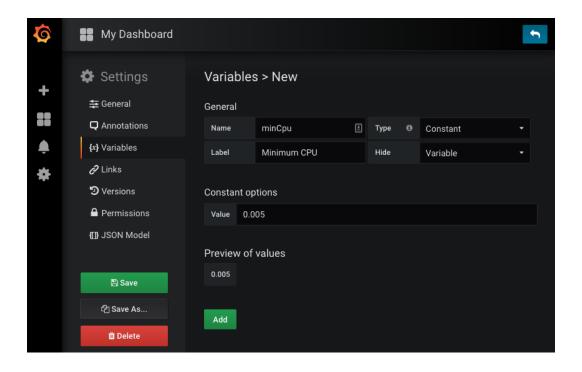


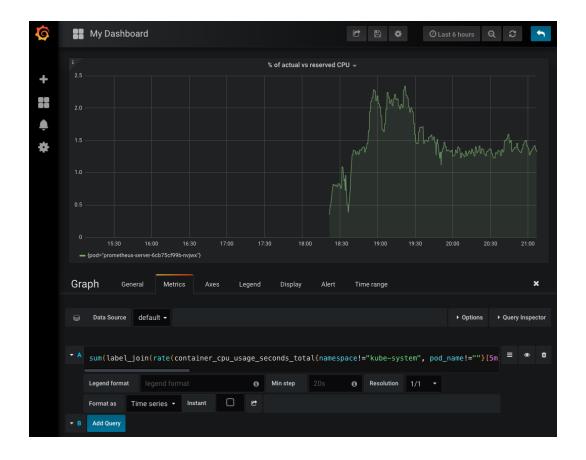


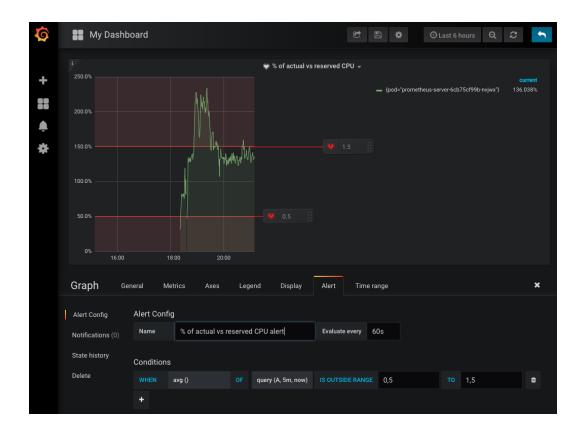


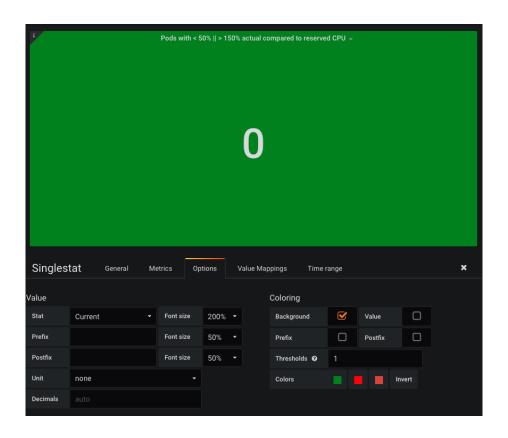




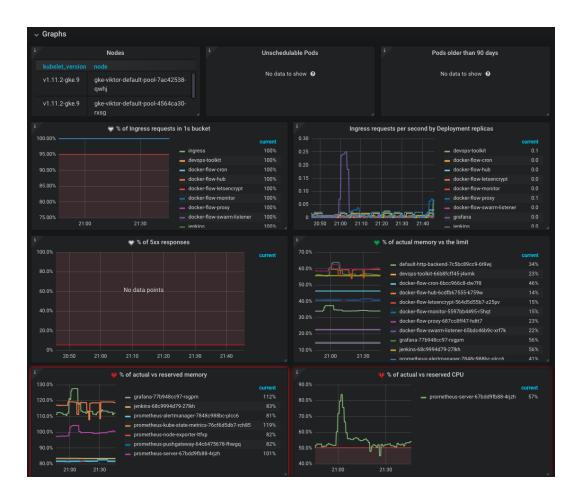




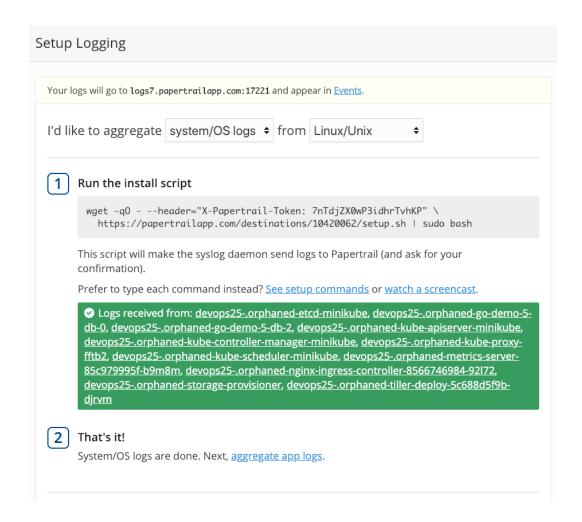


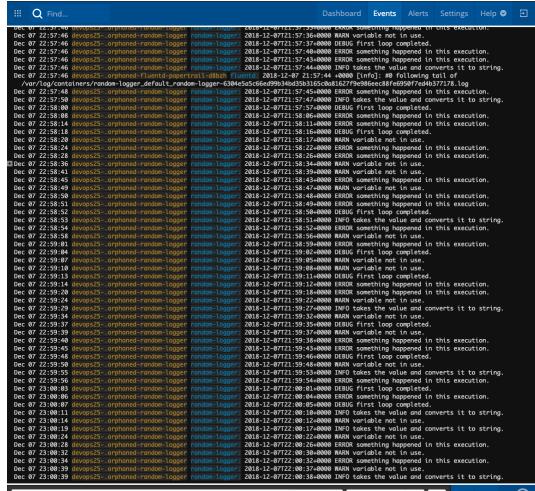






Chapter 7: Collecting and Querying Logs





🔾 random-logger











