Learning promQL

I have a 2 node kubernetes cluster.

We are using the below values

For counter metrics , you will be more interested in rate of change per second

For gauge metrics , we would like to aggregate by sum, avg, min,max

Observe these (for say rates of failures(counter), or for gauges)

Aggregate or Sum the rate of CPU in idle mode per instance for a period of 5 minutes

sum by (instance)(rate(node\_cpu\_seconds\_total{mode="idle"}[5m]))

Percentage of Idle CPU

(sum by (instance)(rate(node\_cpu\_seconds\_total{mode="idle"}[5m])))/ (sum by (instance)(rate(node\_cpu\_seconds\_total[5m])))

Percentage of User CPU

(sum by (instance)(rate(node\_cpu\_seconds\_total{mode="user"}[5m])))/ (sum by (instance)(rate(node\_cpu\_seconds\_total[5m])))

Percentage of system CPU

(sum by (instance)(rate(node\_cpu\_seconds\_total{mode="system"}[5m])))/ (sum by (instance)(rate(node\_cpu\_seconds\_total[5m])))

API server Request Duration

sum by (group) (rate(apiserver\_request\_duration\_seconds\_count[5m]))

Other Queries

sum by (group,namespace) (rate(apiserver\_request\_duration\_seconds\_sum[1d])/rate(apiserver\_request\_duration\_seconds\_count[1d]))

Learning about K8s API Server

sum by (namespace,instance) (rate(apiserver\_request\_duration\_seconds\_count[5m]))

Total Number of requests received by API Server per instance

sum by (instance) (apiserver\_request\_total)

Apiserver requests where status code is not 200 or 201

(apiserver\_request\_total{code!~"200|201"})

Apiserver requests where status code is 500 or something like 500,502,503,504 etc

(apiserver\_request\_total{code=~"50\*"})

The same query as above, but aggregated by iinstance

sum by (instance)(apiserver\_request\_total{code=~"50\*"})

Aggregate by namespace and code

sum by (namespace,code) (rate(apiserver\_request\_total[5m]))

Aggregate configmap by namespaces and pod

sum by (namespace, pod) (kube\_configmap\_info)

All pods by images

sum by (image) (kube\_pod\_container\_info)

All containers in a waiting state

kube\_pod\_container\_status\_waiting!=0

The rate of container restarts in pod aggregated by namespace

sum by (namespace)(rate(kube\_pod\_container\_status\_restarts\_total[5m]))

sum by (instance) (rate(kube\_pod\_container\_status\_restarts\_total[5m]))

**More about the performance of a web application**

Averages from Summaries

Percentiles from Histogram