Monitoring of Monitoring

in open source project

VICTORIA
METRICS

Roman Khavronenko | github.com/hagen1778

Roman Khavronenko

- I am a software engineer
- I like Go, playing video games and listening podcasts
- I lived in Ukraine, UK, Austria and now live in Poland
- Created ClickHouse datasource for Grafana and load balancer proxy (chproxy) for ClickHouse
- For last few years I'm working on VictoriaMetrics

https://github.com/hagen1778

https://twitter.com/hagen1778

VICTORIA METRICS

The High Performance
Open Source Time Series Database & Monitoring Solution

13k stars
200+ contributors

Apache 2.0 license

Grammarly, CERN,
Roblox, Adidas, Wix

What this talk is about

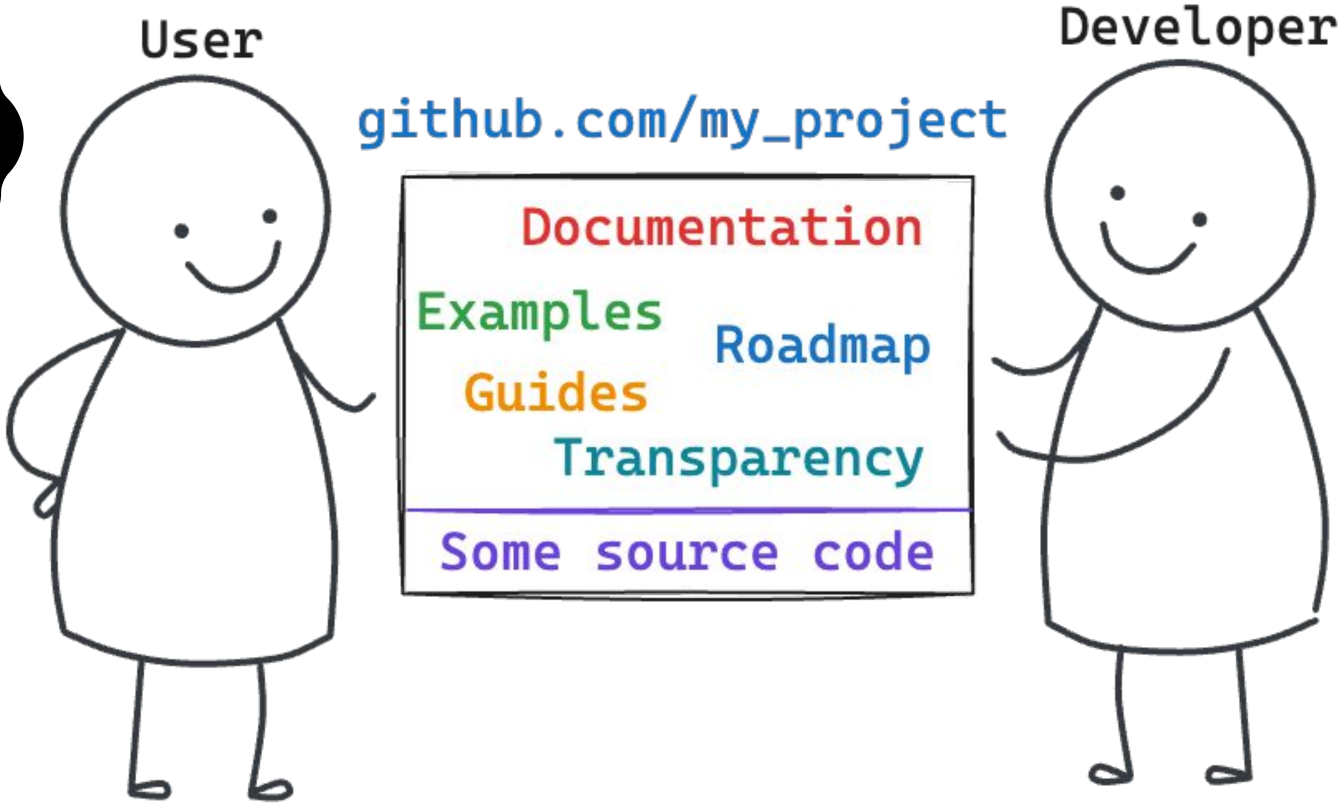
1. Maintaining **open source projects**

2. **Monitoring** and **distributed systems**

3. For people familiar with **Grafana**, **Prometheus** or **VictoriaMetrics**

How open source project should look like

100%
HUMAN
CREATED
CONTENT
NO AI



User

Developer

github.com/my_project

Documentation

Examples

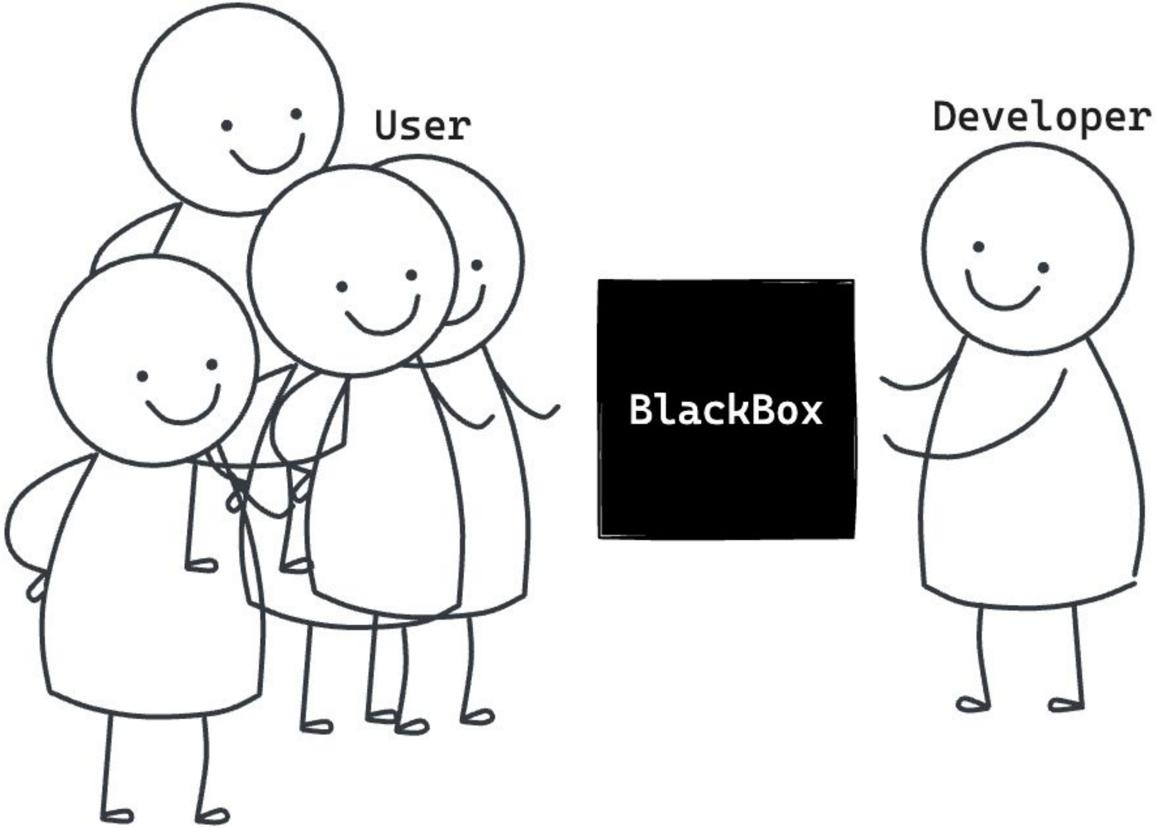
Roadmap

Guides

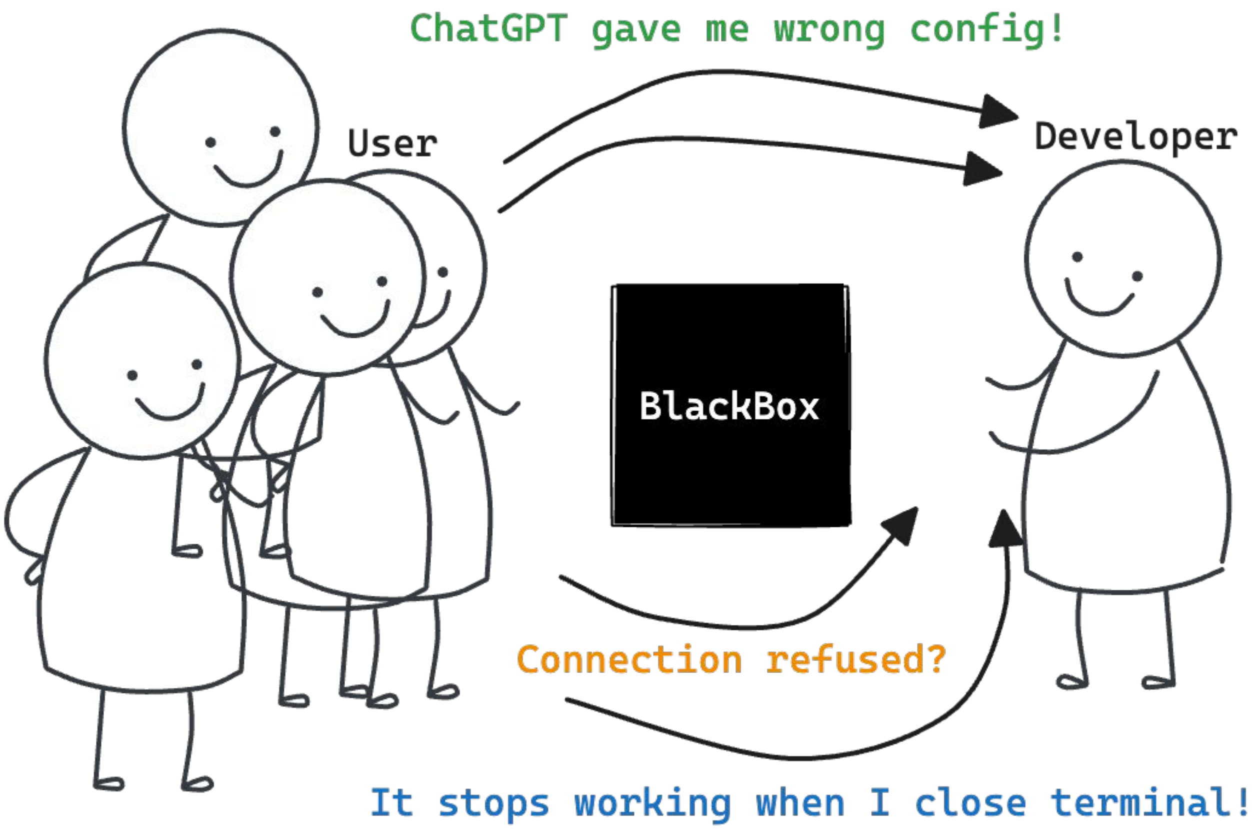
Transparency

Some source code

How it usually looks like

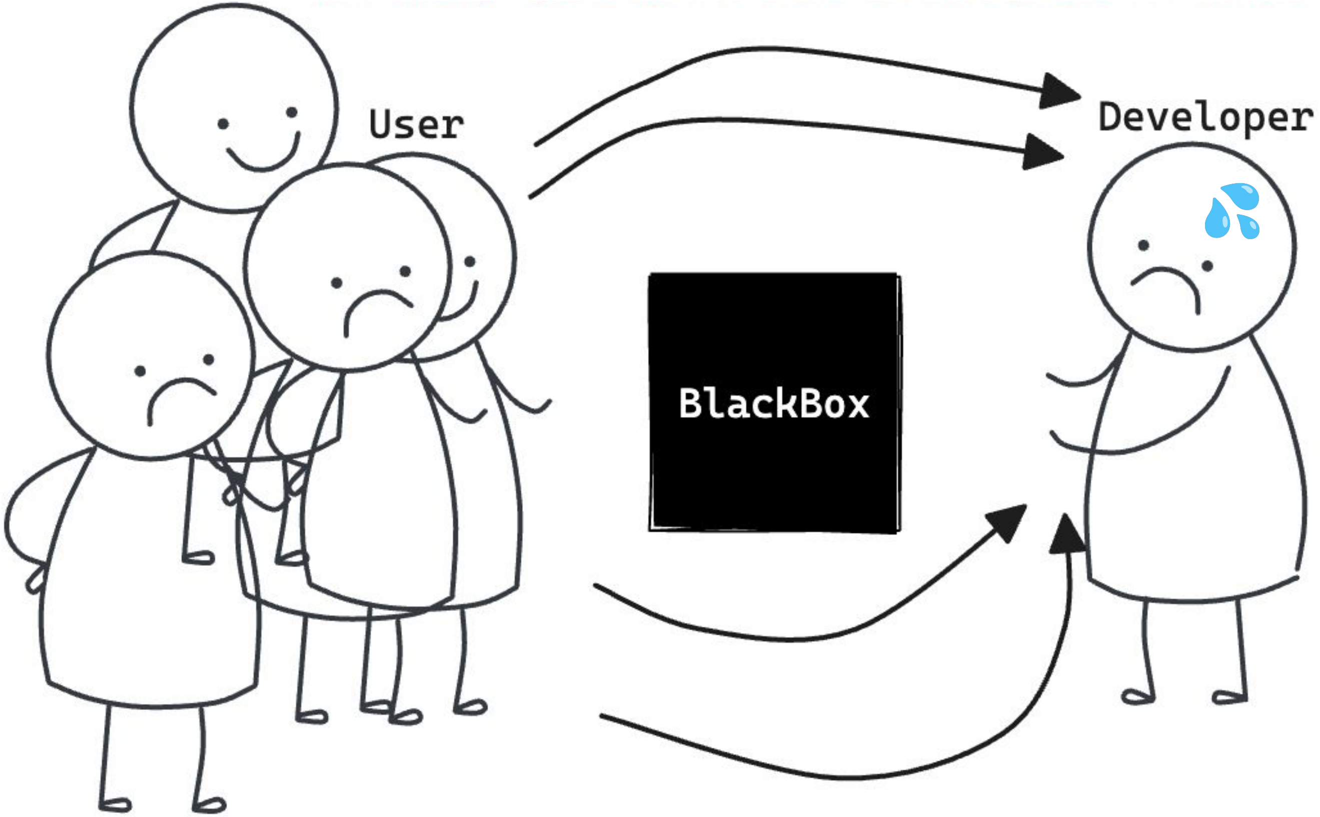


When project is a blackbox - maintainer is a documentation

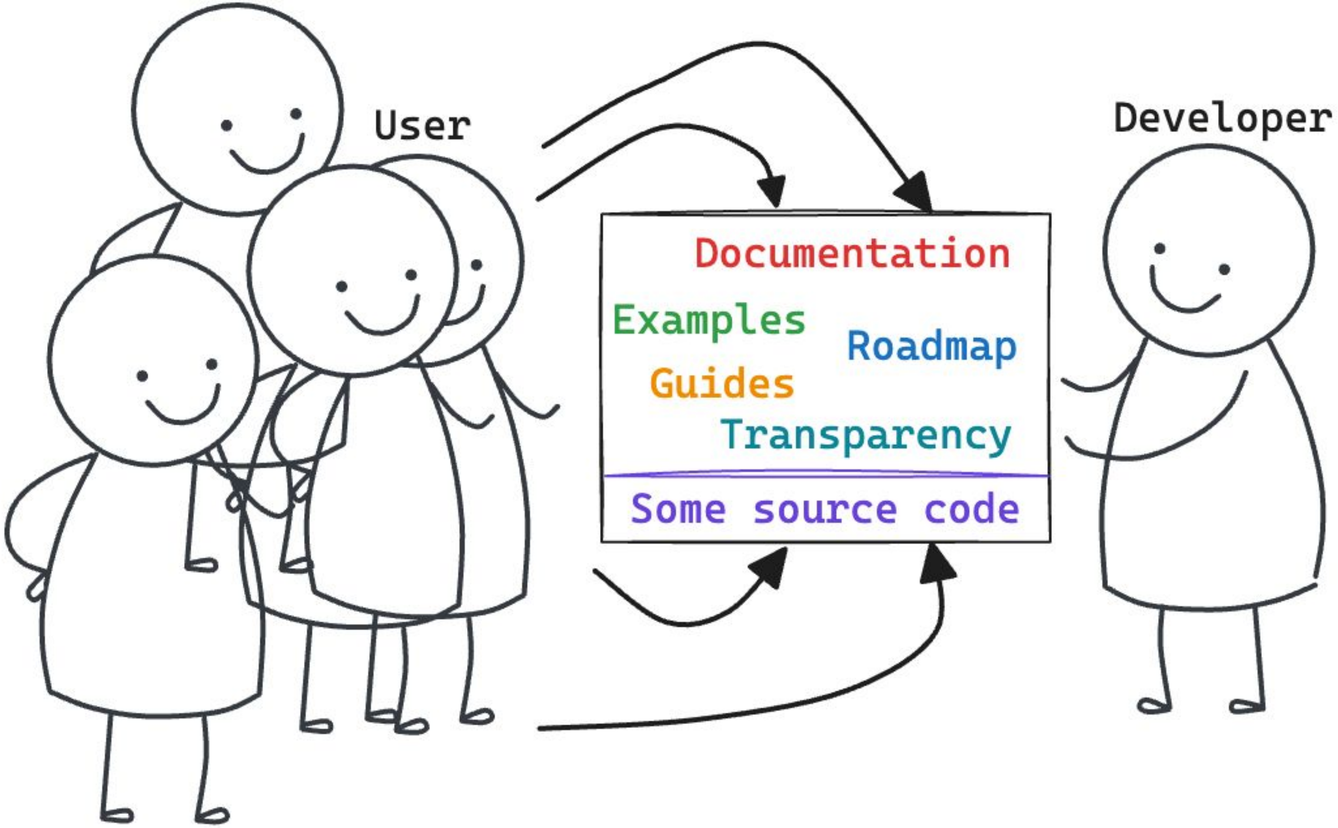


ChatGPT gave me wrong config!
User
Developer
BlackBox
Connection refused?
It stops working when I close terminal!

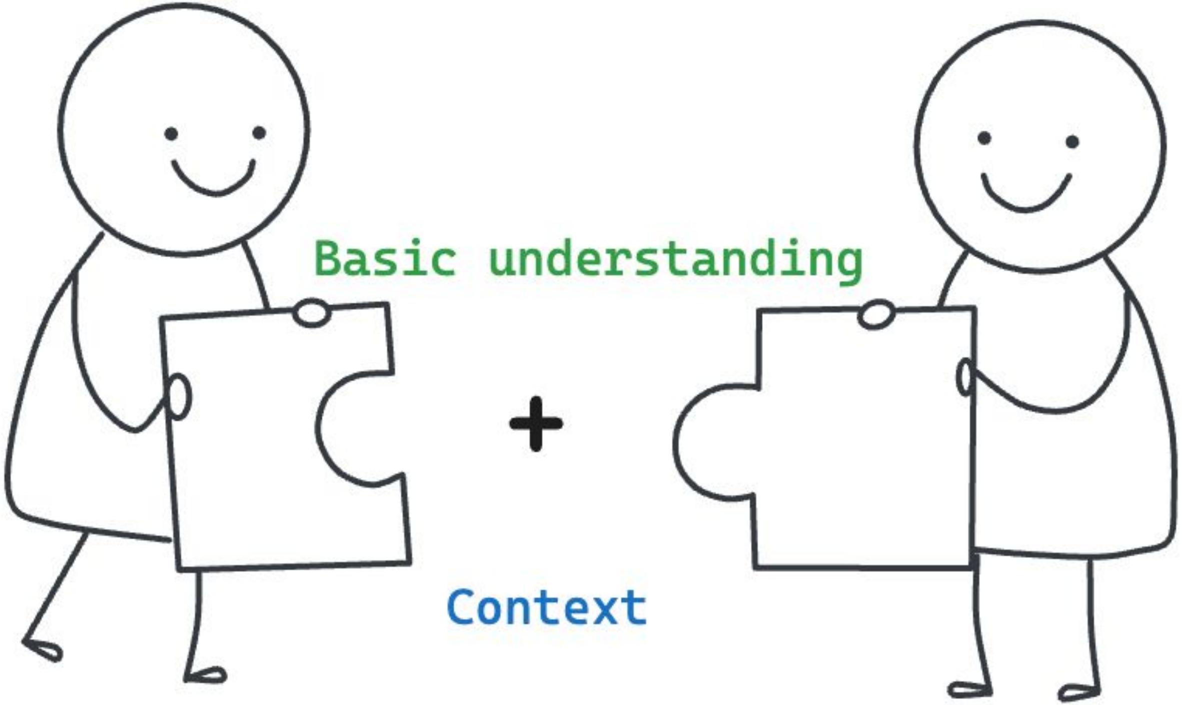
Too many questions for maintainer to handle...



Make it transparent to reduce number of questions



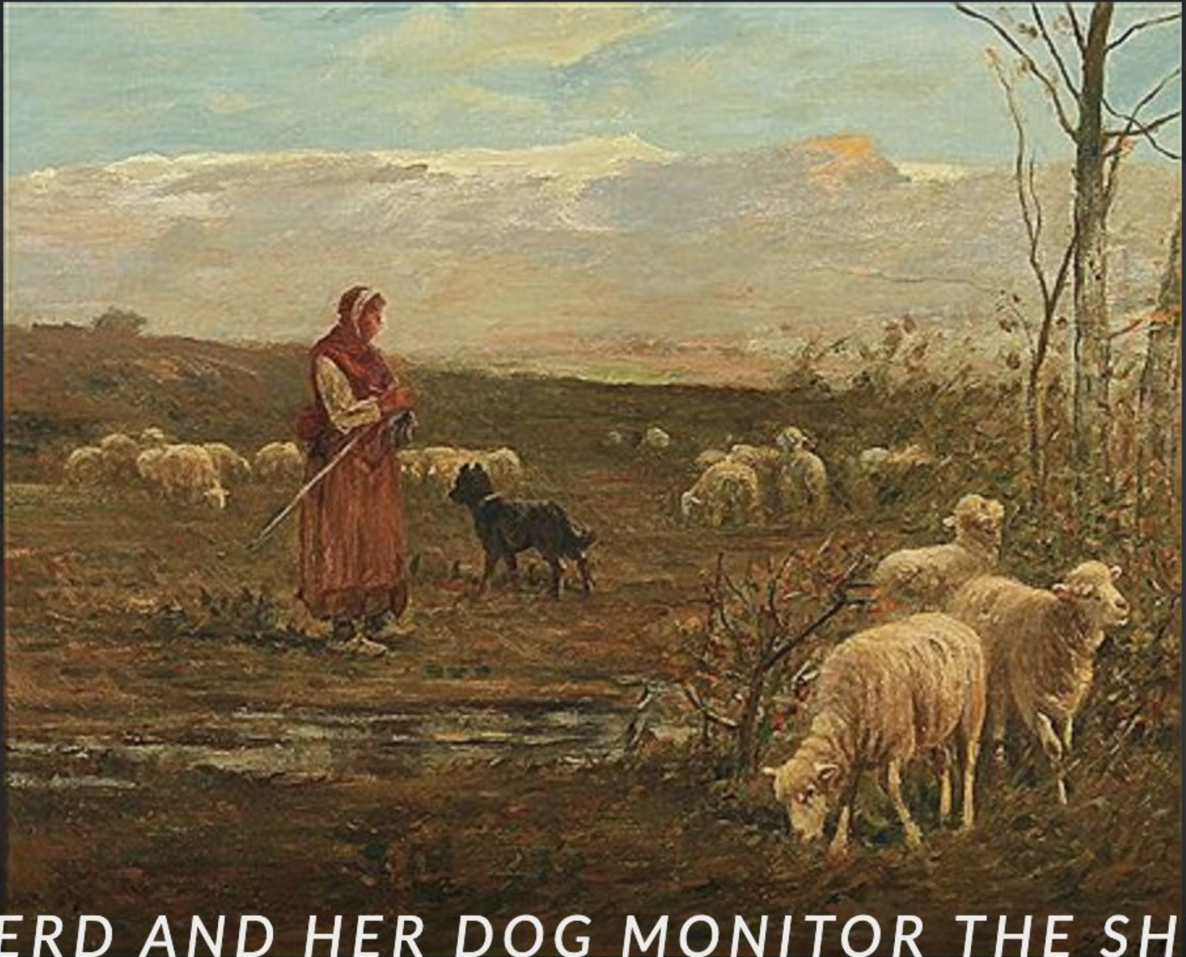
Help users to understand the project better



Basic understanding

+

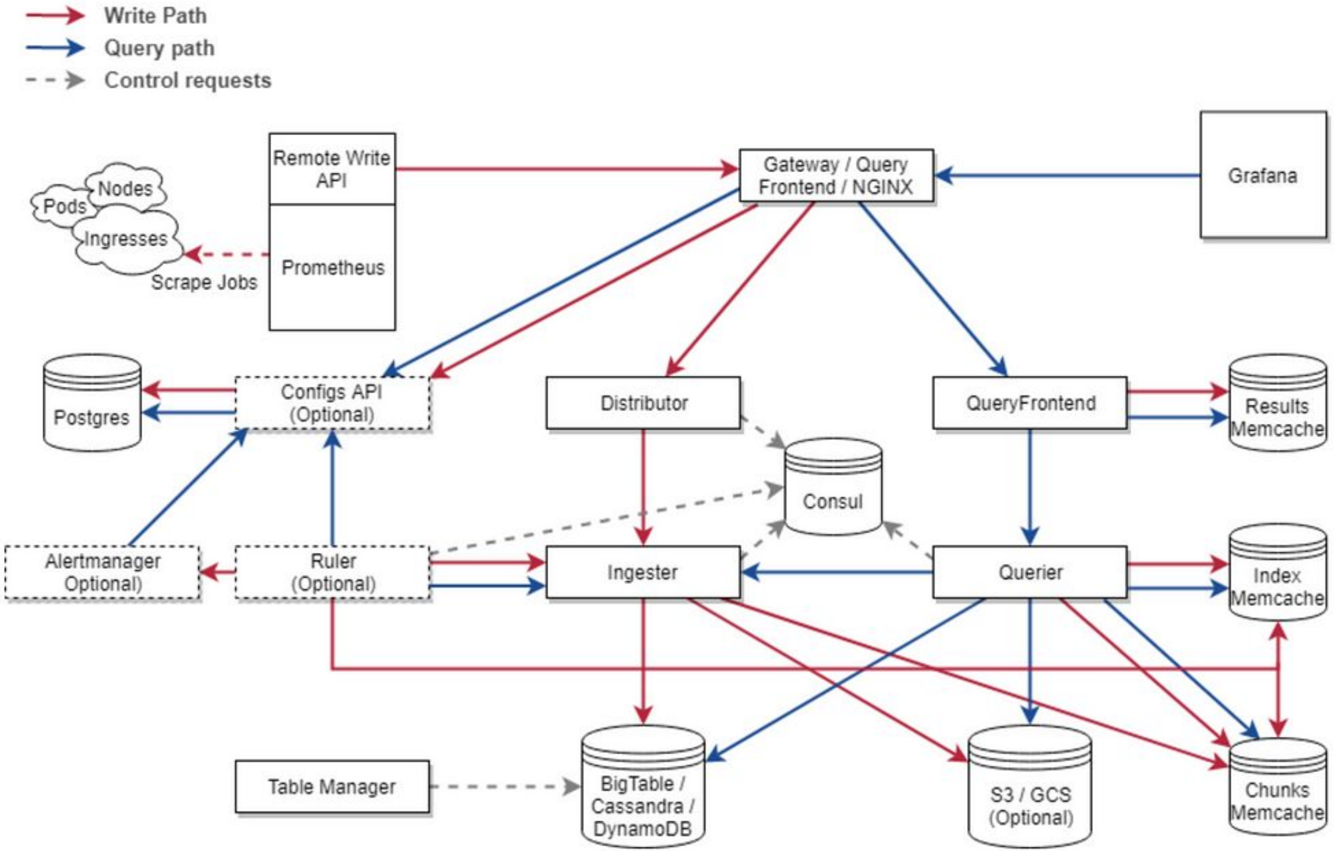
Context



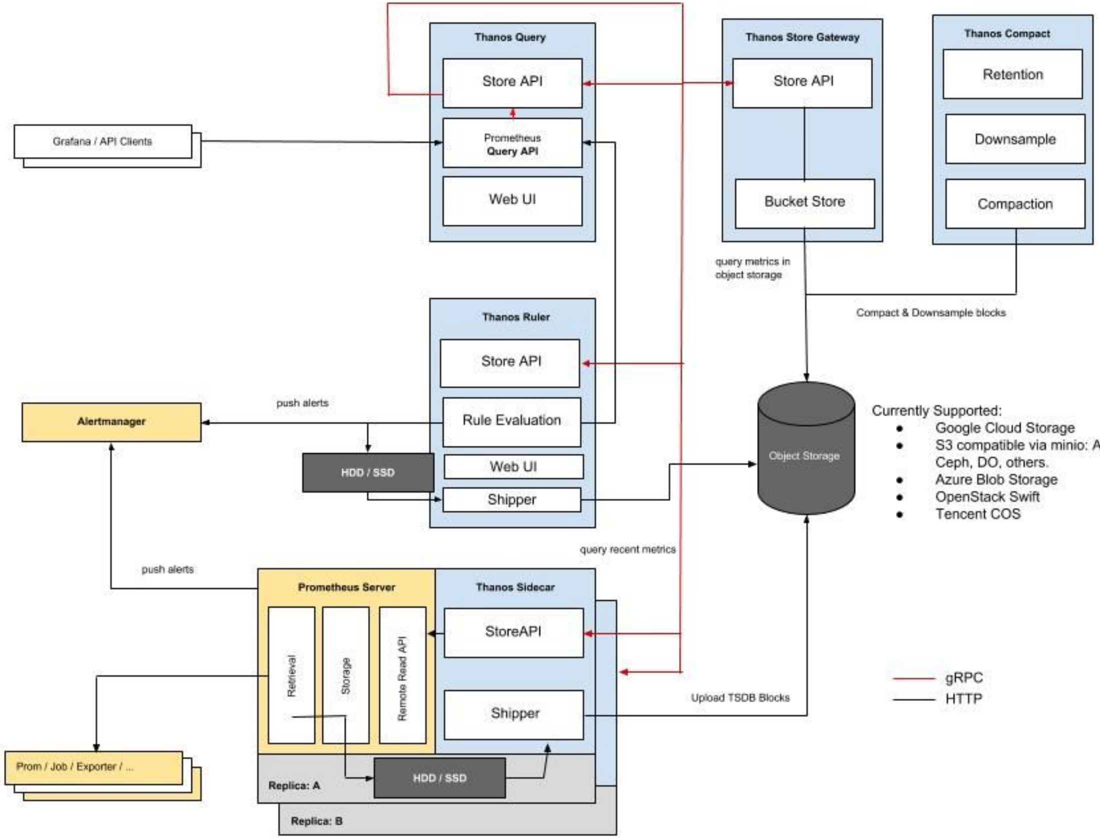
A SHEPHERD AND HER DOG MONITOR THE SHEEP.

Albert Charpin

Cortex architecture



Thanos architecture



Currently Supported:

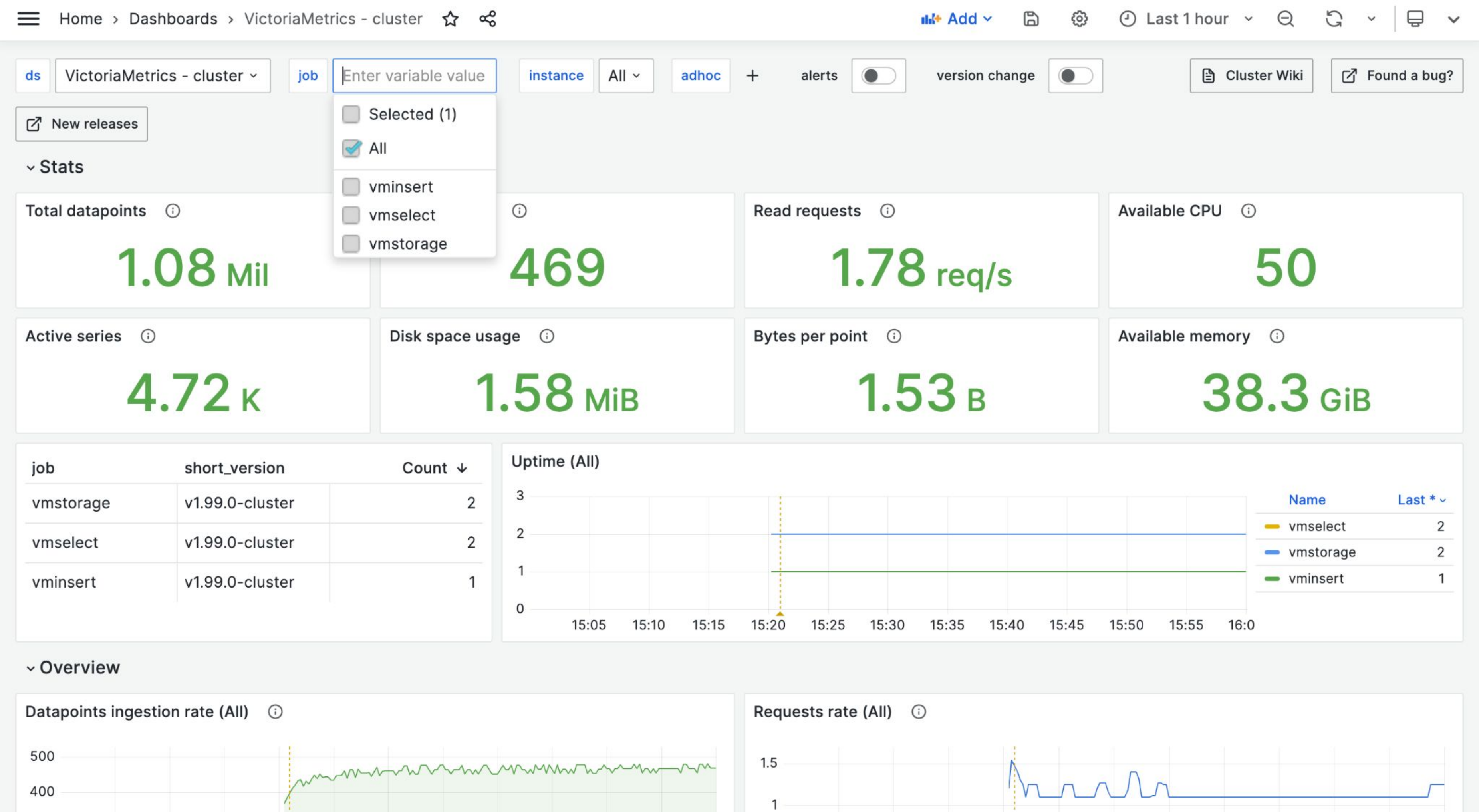
- Google Cloud Storage
- S3 compatible via minio: AWS S3, Ceph, DO, others.
- Azure Blob Storage
- OpenStack Swift
- Tencent COS

How to make it easier to understand?

- Write good Documentation
- Instrument it with helpful Logs and meaningful Metrics
- Provide Dashboards (Grafana)
- And Alerting rules

Grafana

VictoriaMetrics - cluster



Home › Dashboards › VictoriaMetrics - cluster

Add · Last 1 hour

ds: VictoriaMetrics - cluster · job: Enter variable value · instance: All · adhoc · + · alerts · version change · Cluster Wiki · Found a bug?

New releases

□ Selected (1)
☑ All
□ vminsert
□ vselect
□ vmstorage

Stats

Total datapoints		Read requests	Available CPU
1.08 Mil	469	1.78 req/s	50

Active series	Disk space usage	Bytes per point	Available memory
4.72 K	1.58 MiB	1.53 B	38.3 GiB

job	short_version	Count ↓
vmstorage	v1.99.0-cluster	2
vmselect	v1.99.0-cluster	2
vminsert	v1.99.0-cluster	1

Uptime (All)

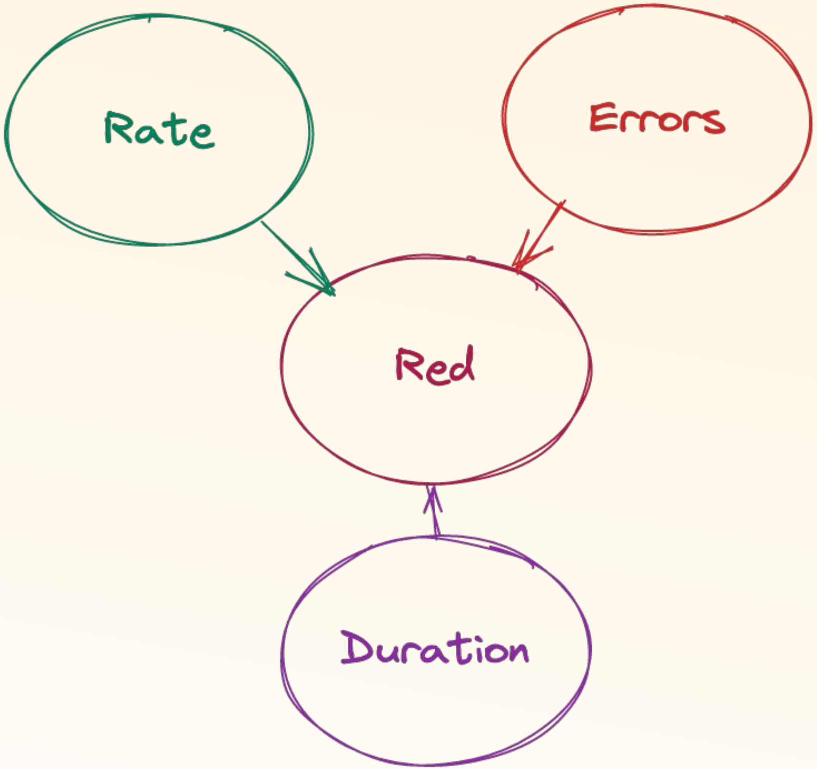
Name	Last *
vmselect	2
vmstorage	2
vminsert	1

Overview

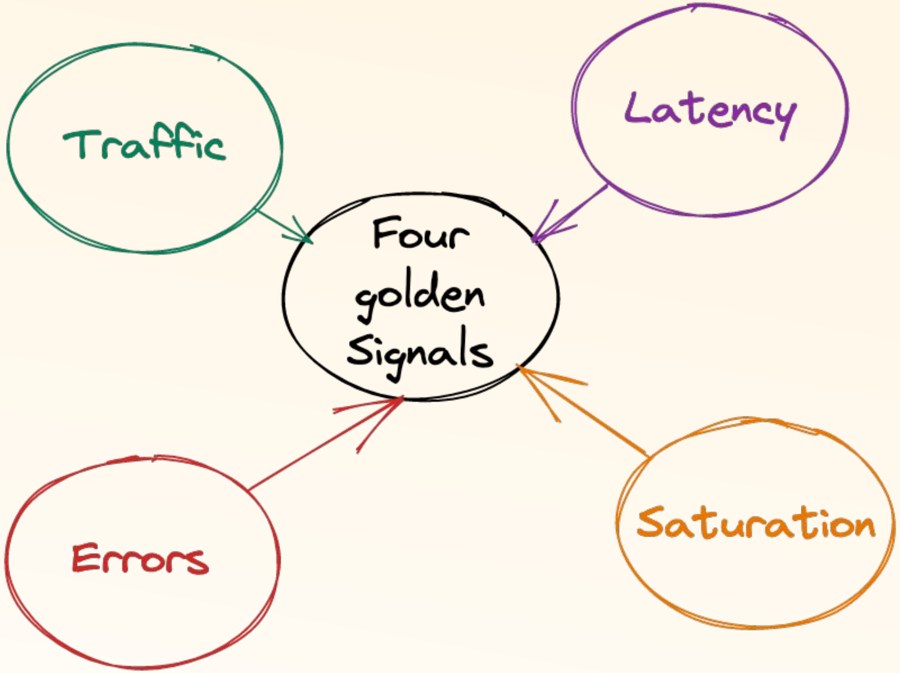
Datapoints ingestion rate (All)

Requests rate (All)

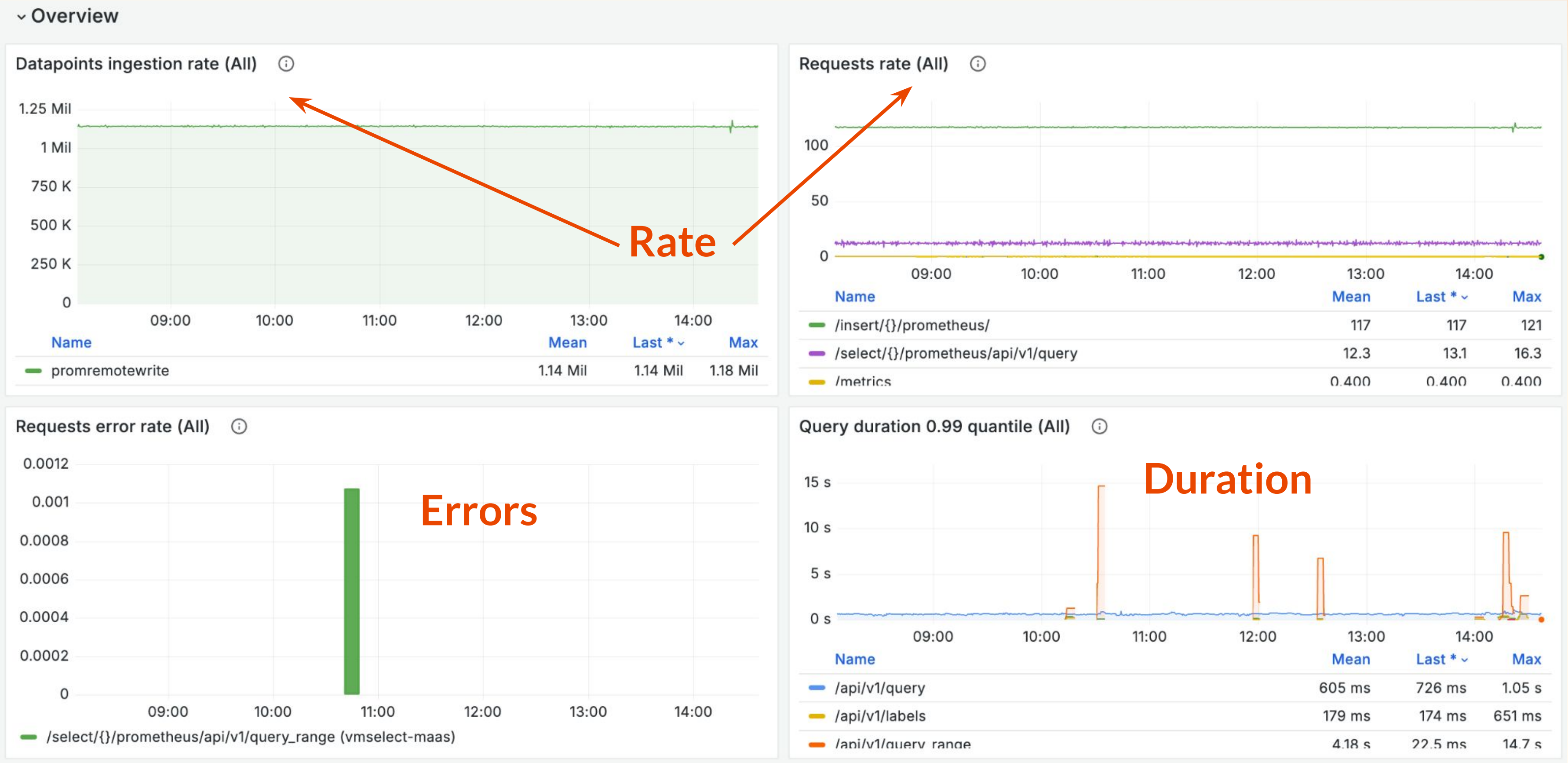
RED method by Weave.works



The Four Golden Signals method by Google



RED signals on the dashboard



Overview

Rate

Errors

Duration

Datapoints ingestion rate (All)

Name	Mean	Last *	Max
🟩 promremotewrite	1.14 Mil	1.14 Mil	1.18 Mil

Requests rate (All)

Name	Mean	Last *	Max
🟩 /insert/{}/prometheus/	117	117	121
🟪 /select/{}/prometheus/api/v1/query	12.3	13.1	16.3
🟨 /metrics	0.400	0.400	0.400

Requests error rate (All)

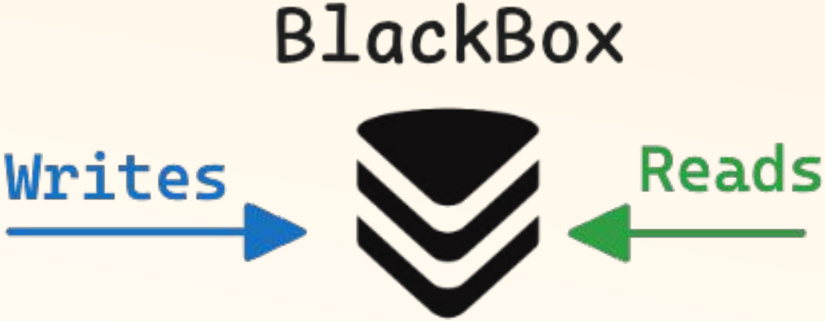
🟩 /select/{}/prometheus/api/v1/query_range (vmselect-maas)

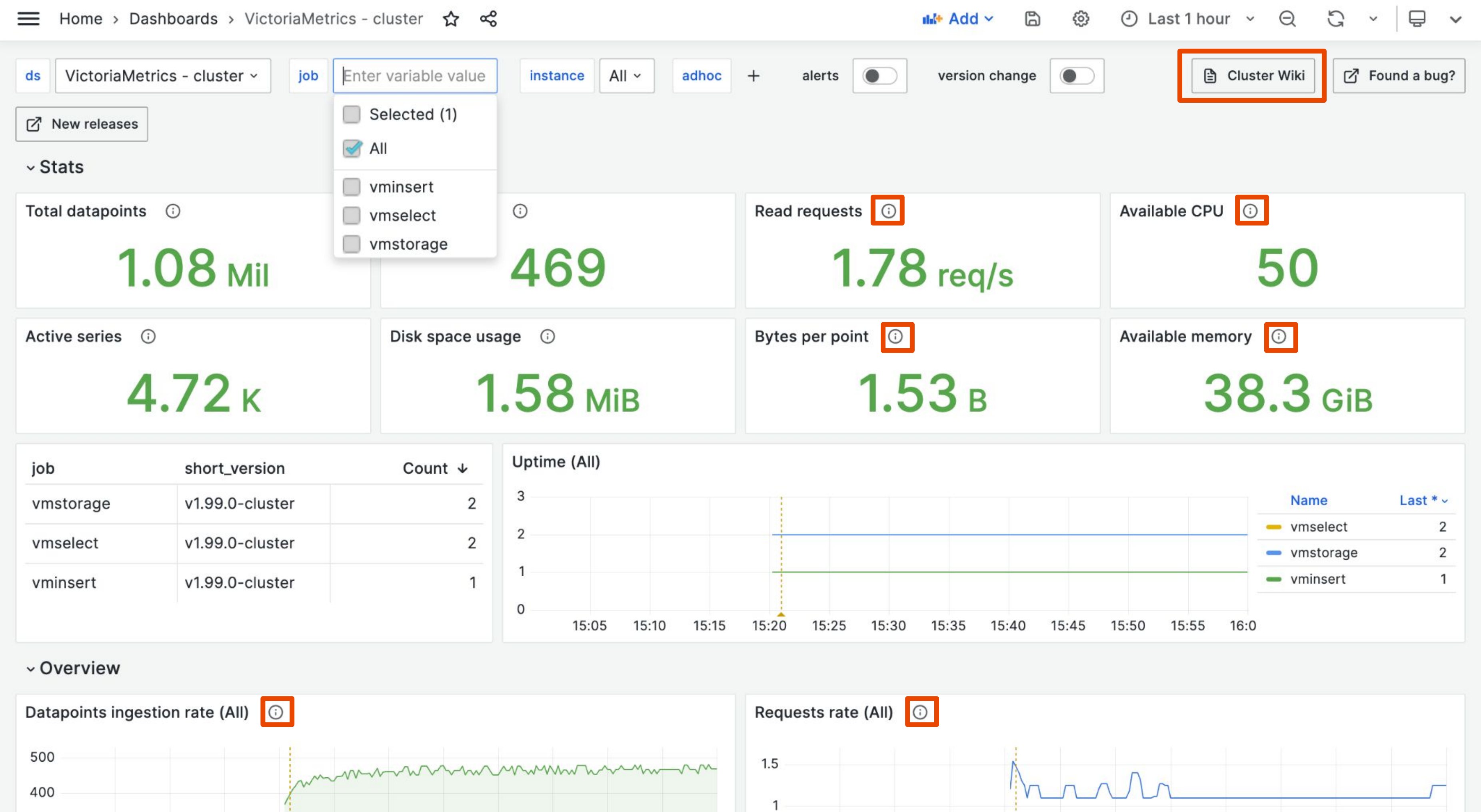
Query duration 0.99 quantile (All)

Name	Mean	Last *	Max
🟦 /api/v1/query	605 ms	726 ms	1.05 s
🟨 /api/v1/labels	179 ms	174 ms	651 ms
🟧 /api/v1/query_range	4.18 s	22.5 ms	14.7 s

Neither RED nor 4 Golden signals answers on **WHY**

- **Why** errors started to appear?
- **Why** latency is high?
- **Why** resources were saturated?





Home › Dashboards › VictoriaMetrics - cluster

Add | Last 1 hour

ds: VictoriaMetrics - cluster | job: Enter variable value | instance: All | adhoc | + | alerts | version change

Cluster Wiki | Found a bug?

New releases

□ Selected (1)
☑ All
□ vminsert
□ vmselect
□ vmstorage

Stats

Total datapoints	469	Read requests	Available CPU
1.08 Mil		1.78 req/s	50

Active series	Disk space usage	Bytes per point	Available memory
4.72 K	1.58 MiB	1.53 B	38.3 GiB

job	short_version	Count ↓
vmstorage	v1.99.0-cluster	2
vmselect	v1.99.0-cluster	2
vminsert	v1.99.0-cluster	1

Uptime (All)

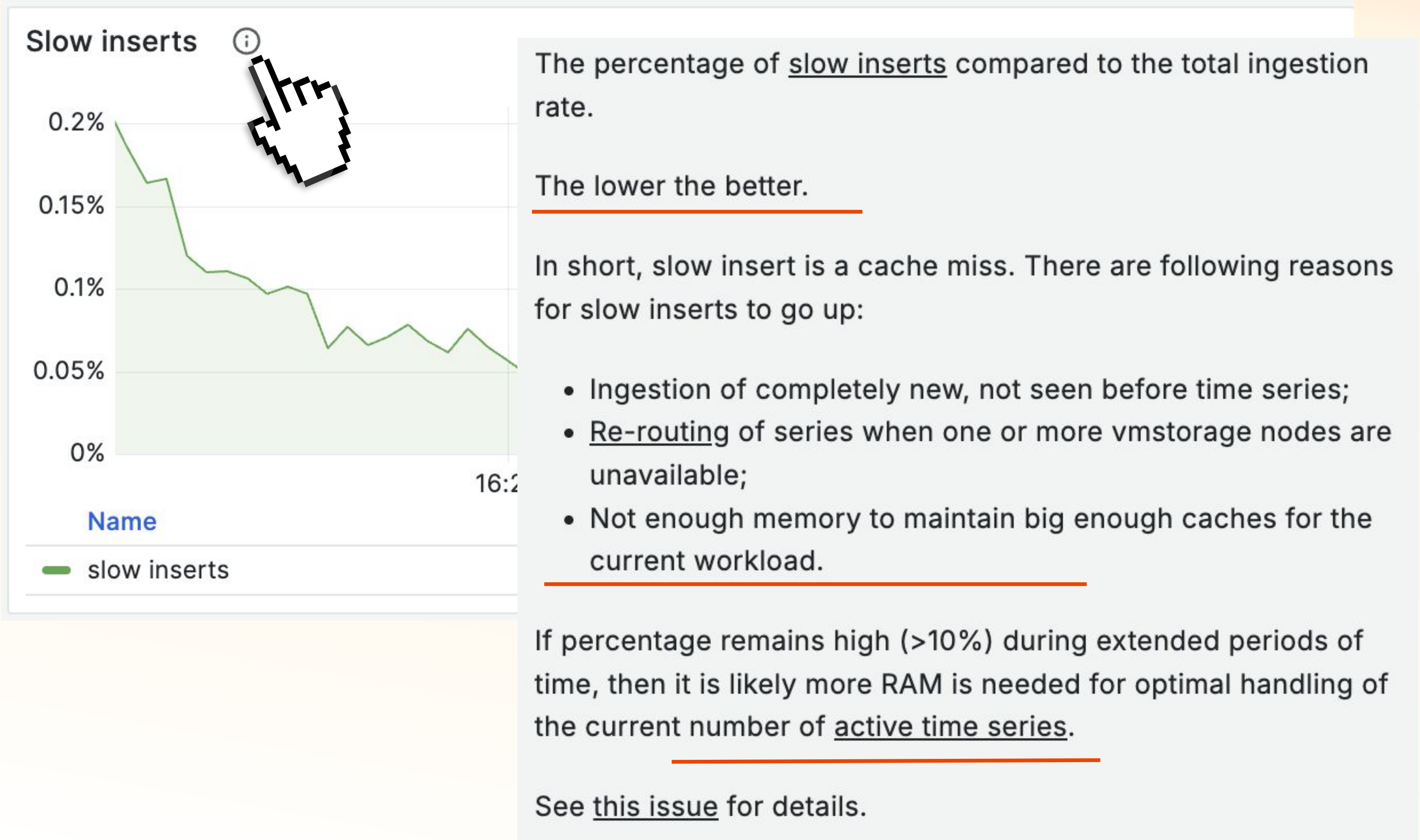
Name	Last *
vmselect	2
vmstorage	2
vminsert	1

Overview

Datapoints ingestion rate (All)

Requests rate (All)

Example of the panel description



The percentage of slow inserts compared to the total ingestion rate.

The lower the better.

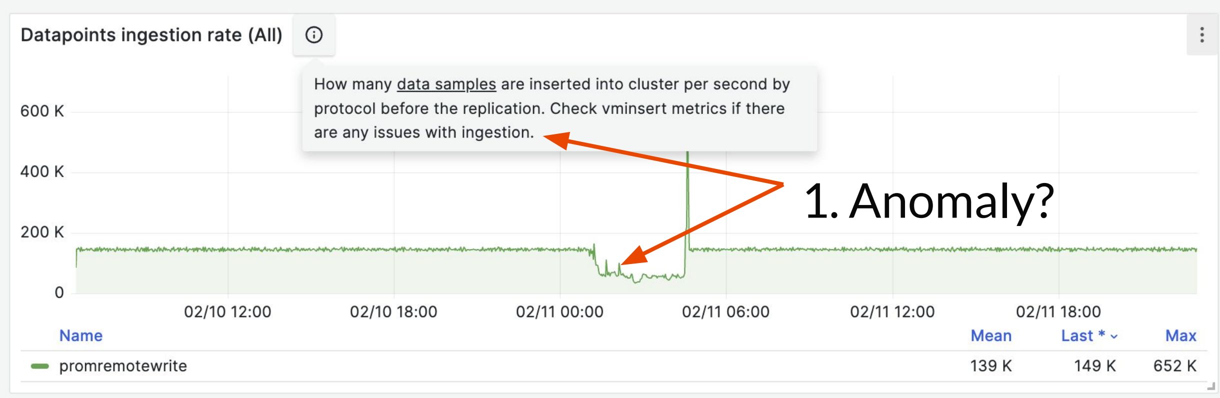
In short, slow insert is a cache miss. There are following reasons for slow inserts to go up:

- Ingestion of completely new, not seen before time series;
- Re-routing of series when one or more vmstorage nodes are unavailable;
- Not enough memory to maintain big enough caches for the current workload.

If percentage remains high (>10%) during extended periods of time, then it is likely more RAM is needed for optimal handling of the current number of active time series.

See this issue for details.

Troubleshooting flow in VM dashboards

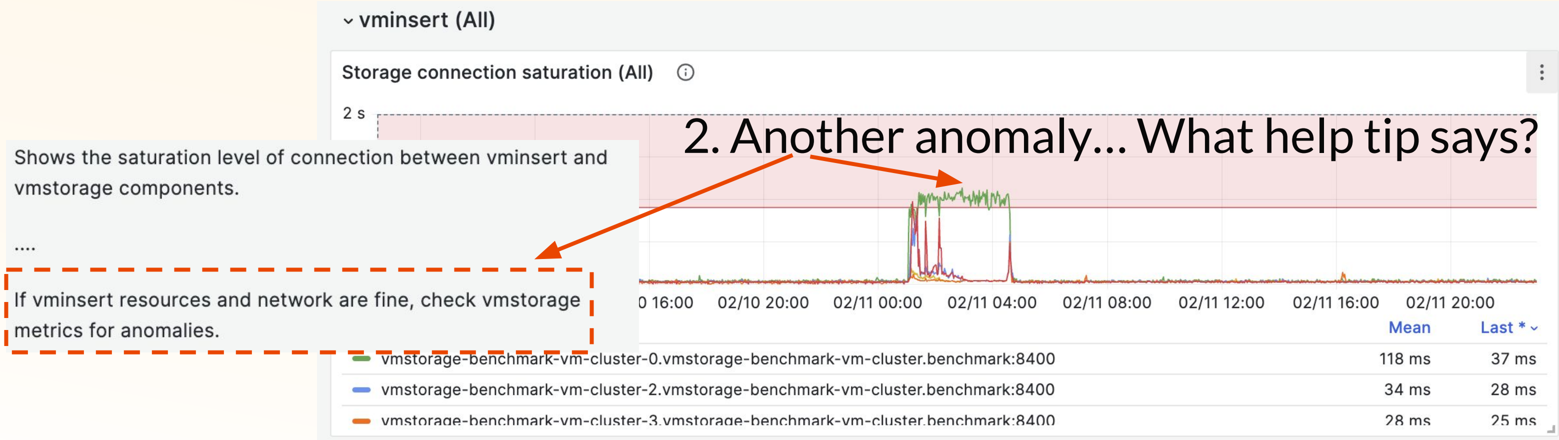


Datapoints ingestion rate (All)

How many data samples are inserted into cluster per second by protocol before the replication. Check vminsert metrics if there are any issues with ingestion.

Name	Mean	Last *	Max
promremotewrite	139 K	149 K	652 K

1. Anomaly?



vminsert (All)

Storage connection saturation (All)

Shows the saturation level of connection between vminsert and vmstorage components.

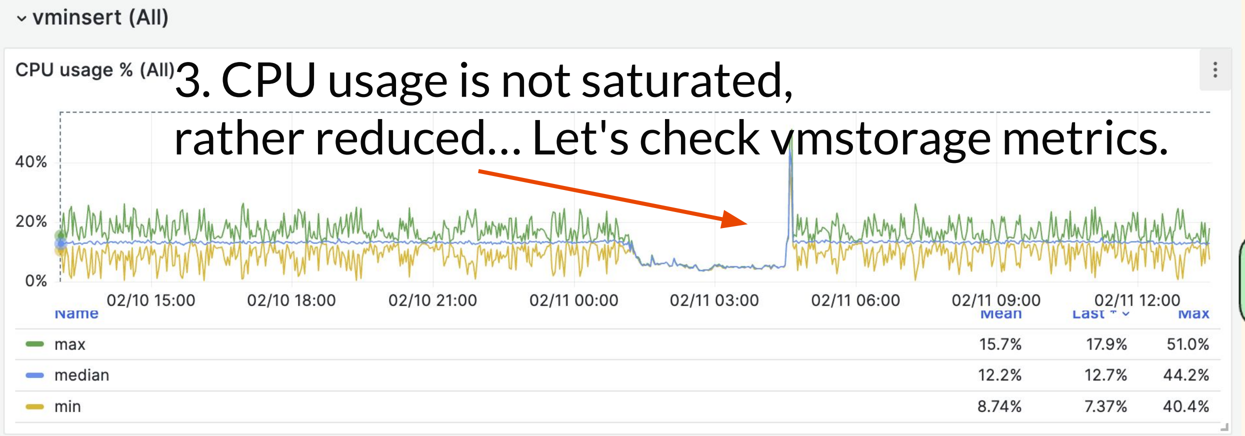
....

If vminsert resources and network are fine, check vmstorage metrics for anomalies.

	Mean	Last *
vmstorage-benchmark-vm-cluster-0.vmstorage-benchmark-vm-cluster.benchmark:8400	118 ms	37 ms
vmstorage-benchmark-vm-cluster-2.vmstorage-benchmark-vm-cluster.benchmark:8400	34 ms	28 ms
vmstorage-benchmark-vm-cluster-3.vmstorage-benchmark-vm-cluster.benchmark:8400	28 ms	25 ms

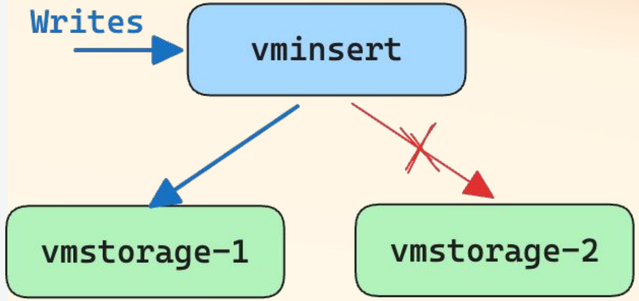
2. Another anomaly... What help tip says?

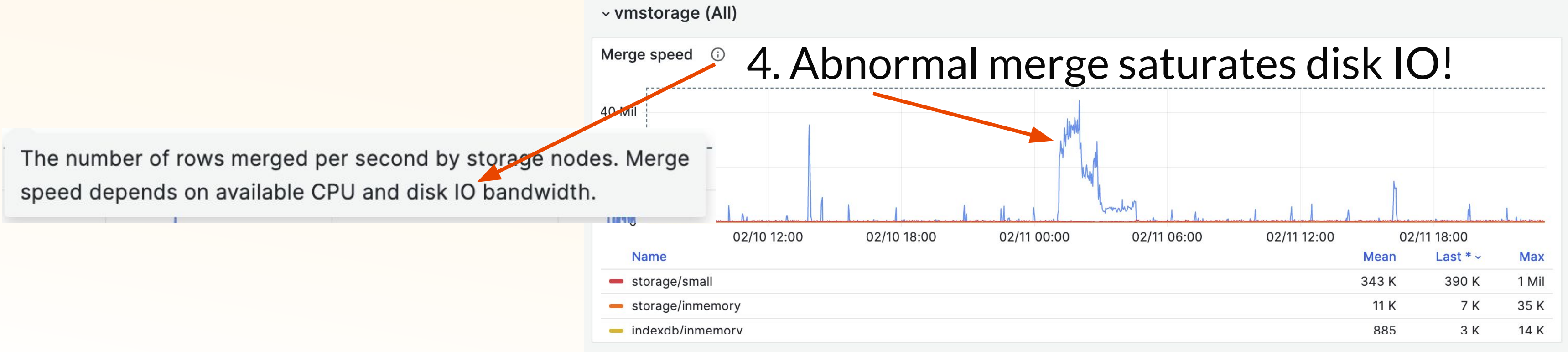
Troubleshooting flow in VM dashboards



3. CPU usage is not saturated, rather reduced... Let's check vmstorage metrics.

Name	Mean	Last *	Max
max	15.7%	17.9%	51.0%
median	12.2%	12.7%	44.2%
min	8.74%	7.37%	40.4%





4. Abnormal merge saturates disk IO!

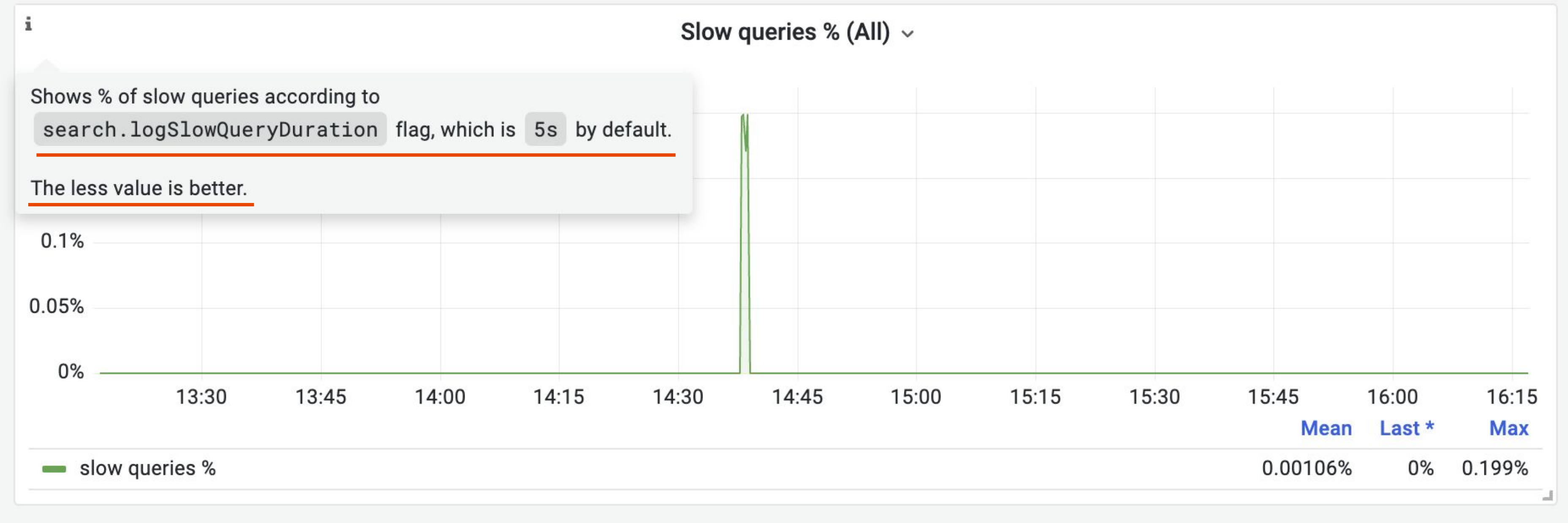
The number of rows merged per second by storage nodes. Merge speed depends on available CPU and disk IO bandwidth.

Name	Mean	Last *	Max
storage/small	343 K	390 K	1 Mil
storage/inmemory	11 K	7 K	35 K
indexdb/inmemory	885	3 K	14 K

Transparency: seeing through code and metrics

Top cool features we use in Grafana dashboards

Tooltips with hints

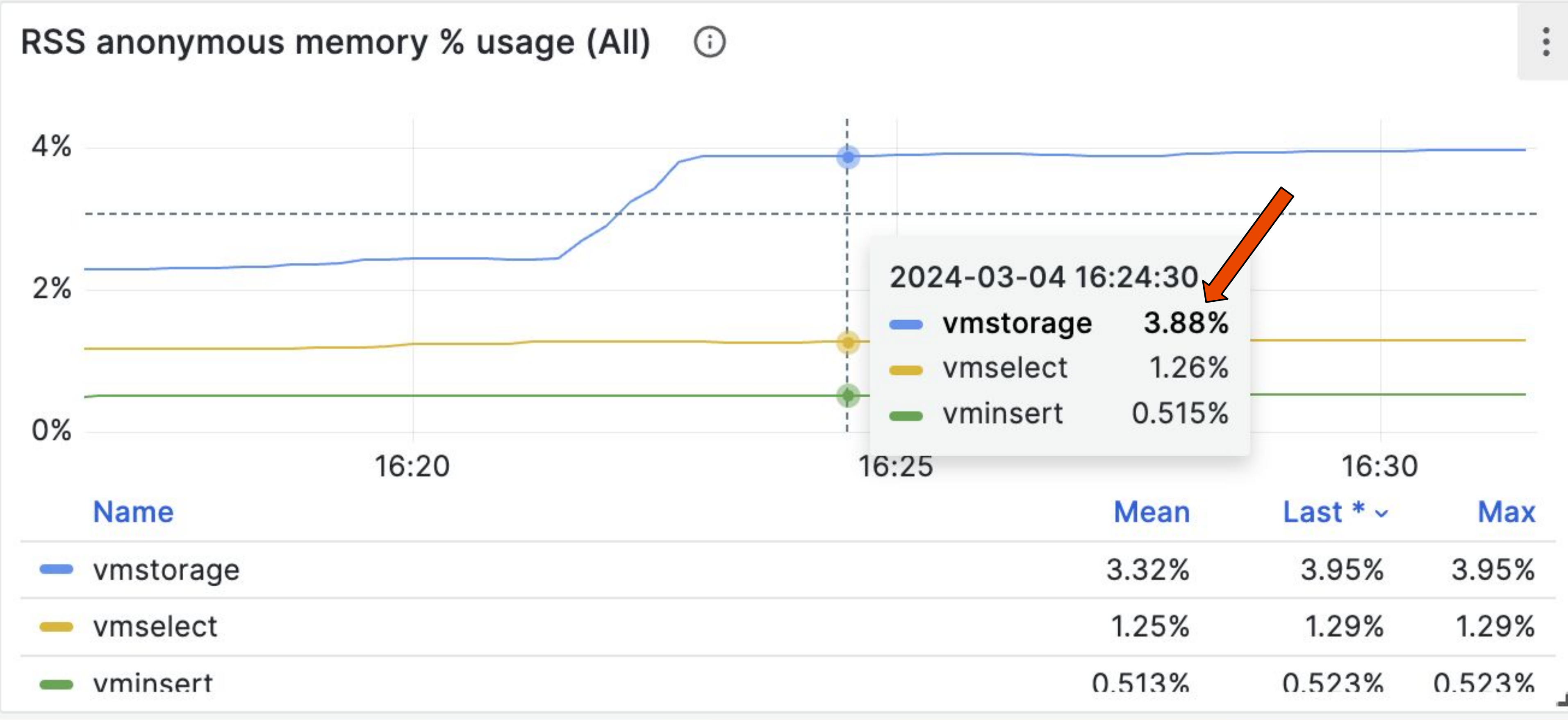


Shows % of slow queries according to `search.logSlowQueryDuration` flag, which is `5s` by default.

The less value is better.

	Mean	Last *	Max
🟩 slow queries %	0.00106%	0%	0.199%

Show only outliers (max) across all components



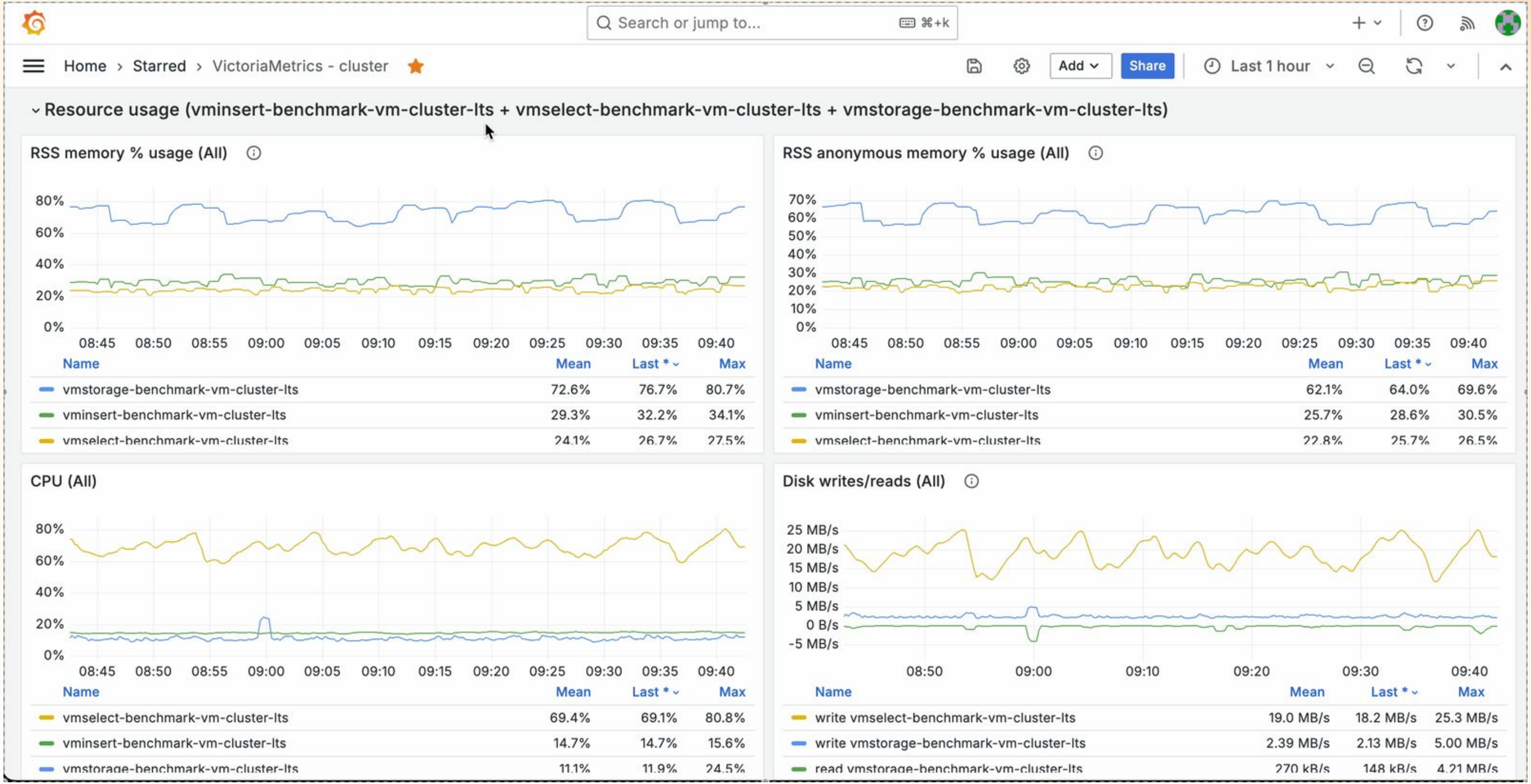
RSS anonymous memory % usage (All)

2024-03-04 16:24:30

vmstorage	**3.88%**
vmselect	1.26%
vminsert	0.515%

Name	Mean	Last * ⌄	Max
vmstorage	3.32%	3.95%	3.95%
vmselect	1.25%	1.29%	1.29%
vminsert	0.513%	0.523%	0.523%

Let user to see detailed view if needed: Drilldown



Resource usage (vminsert-benchmark-vm-cluster-lts + vmselect-benchmark-vm-cluster-lts + vmstorage-benchmark-vm-cluster-lts)

RSS memory % usage (All)

Name	Mean	Last *	Max
vmstorage-benchmark-vm-cluster-lts	72.6%	76.7%	80.7%
vminsert-benchmark-vm-cluster-lts	29.3%	32.2%	34.1%
vmselect-benchmark-vm-cluster-lts	24.1%	26.7%	27.5%

RSS anonymous memory % usage (All)

Name	Mean	Last *	Max
vmstorage-benchmark-vm-cluster-lts	62.1%	64.0%	69.6%
vminsert-benchmark-vm-cluster-lts	25.7%	28.6%	30.5%
vmselect-benchmark-vm-cluster-lts	22.8%	25.7%	26.5%

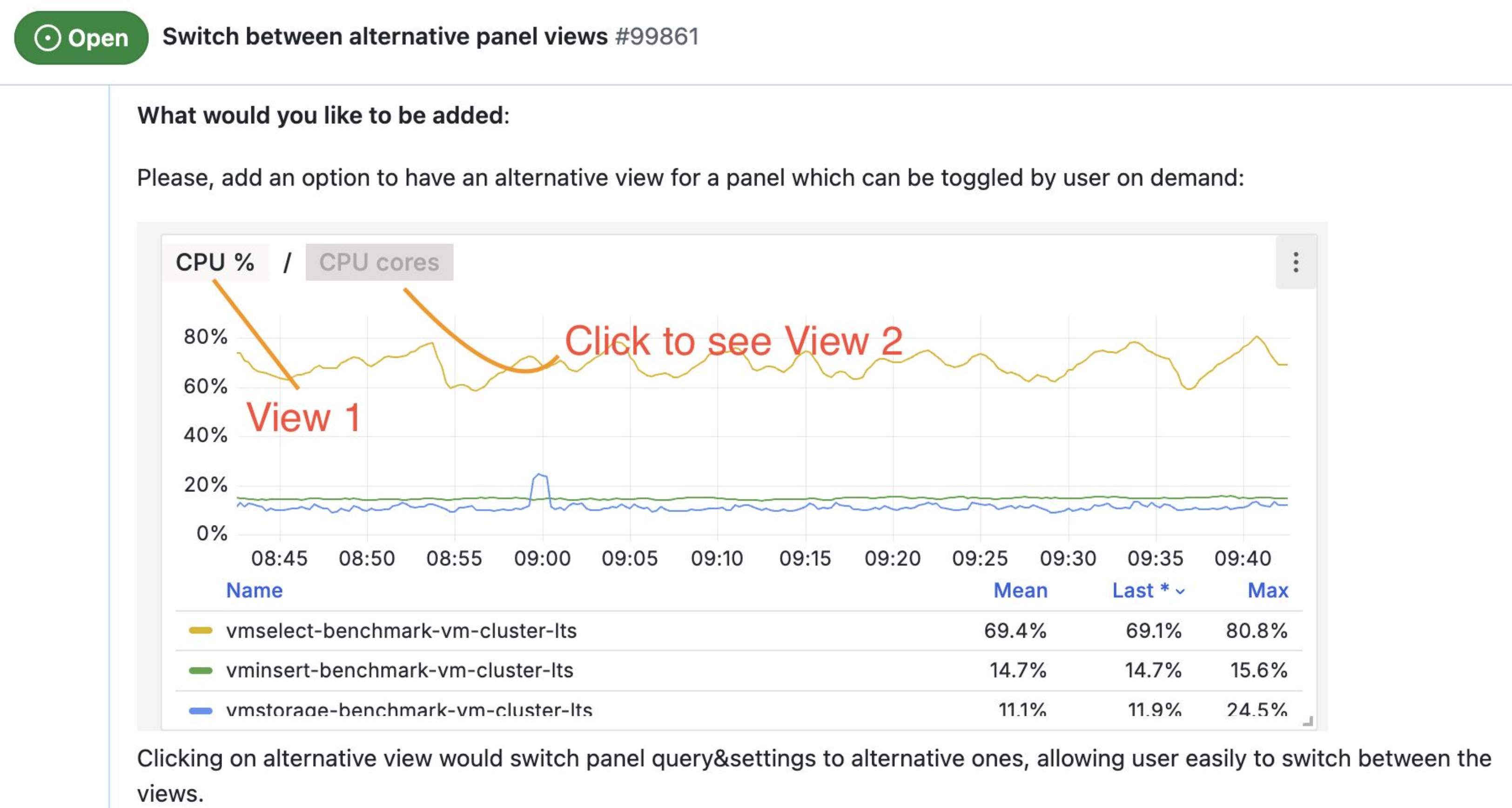
CPU (All)

Name	Mean	Last *	Max
vmselect-benchmark-vm-cluster-lts	69.4%	69.1%	80.8%
vminsert-benchmark-vm-cluster-lts	14.7%	14.7%	15.6%
vmstorage-benchmark-vm-cluster-lts	11.1%	11.9%	24.5%

Disk writes/reads (All)

Name	Mean	Last *	Max
write vmselect-benchmark-vm-cluster-lts	19.0 MB/s	18.2 MB/s	25.3 MB/s
write vmstorage-benchmark-vm-cluster-lts	2.39 MB/s	2.13 MB/s	5.00 MB/s
read vmstorage-benchmark-vm-cluster-lts	270 kB/s	148 kB/s	4.21 MB/s

Support my feature request #99861 for Grafana!



Open **Switch between alternative panel views** #99861

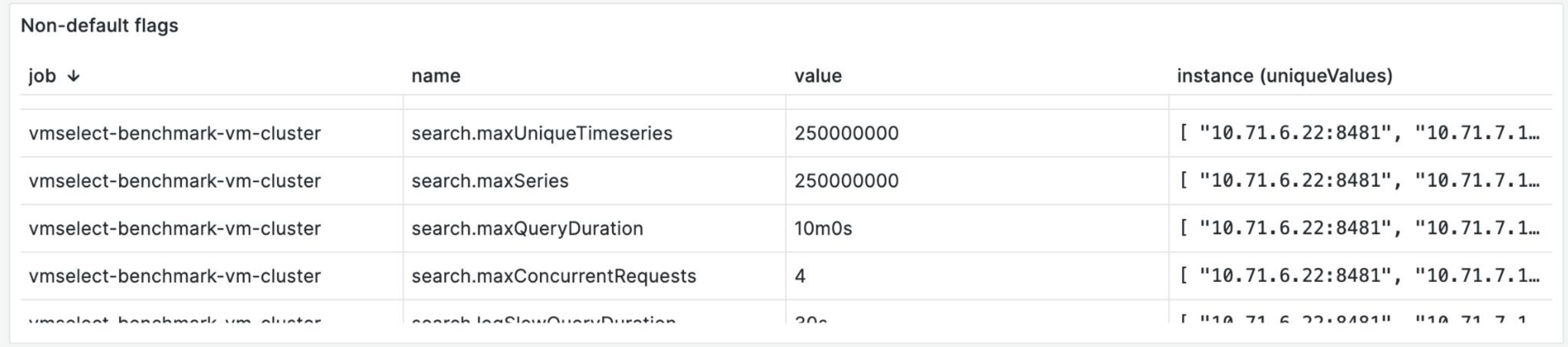
What would you like to be added:

Please, add an option to have an alternative view for a panel which can be toggled by user on demand:

Name	Mean	Last * ⌄	Max
vmselect-benchmark-vm-cluster-lts	69.4%	69.1%	80.8%
vminsert-benchmark-vm-cluster-lts	14.7%	14.7%	15.6%
vmstorage-benchmark-vm-cluster-lts	11.1%	11.9%	24.5%

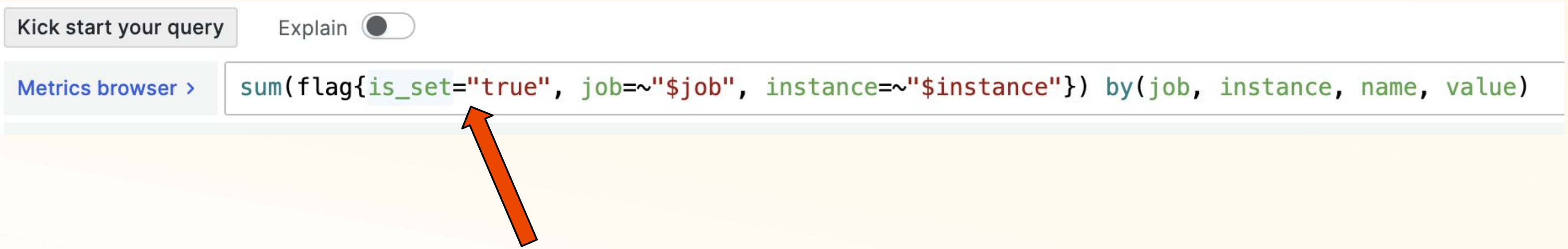
Clicking on alternative view would switch panel query&settings to alternative ones, allowing user easily to switch between the views.

Show flags set by user



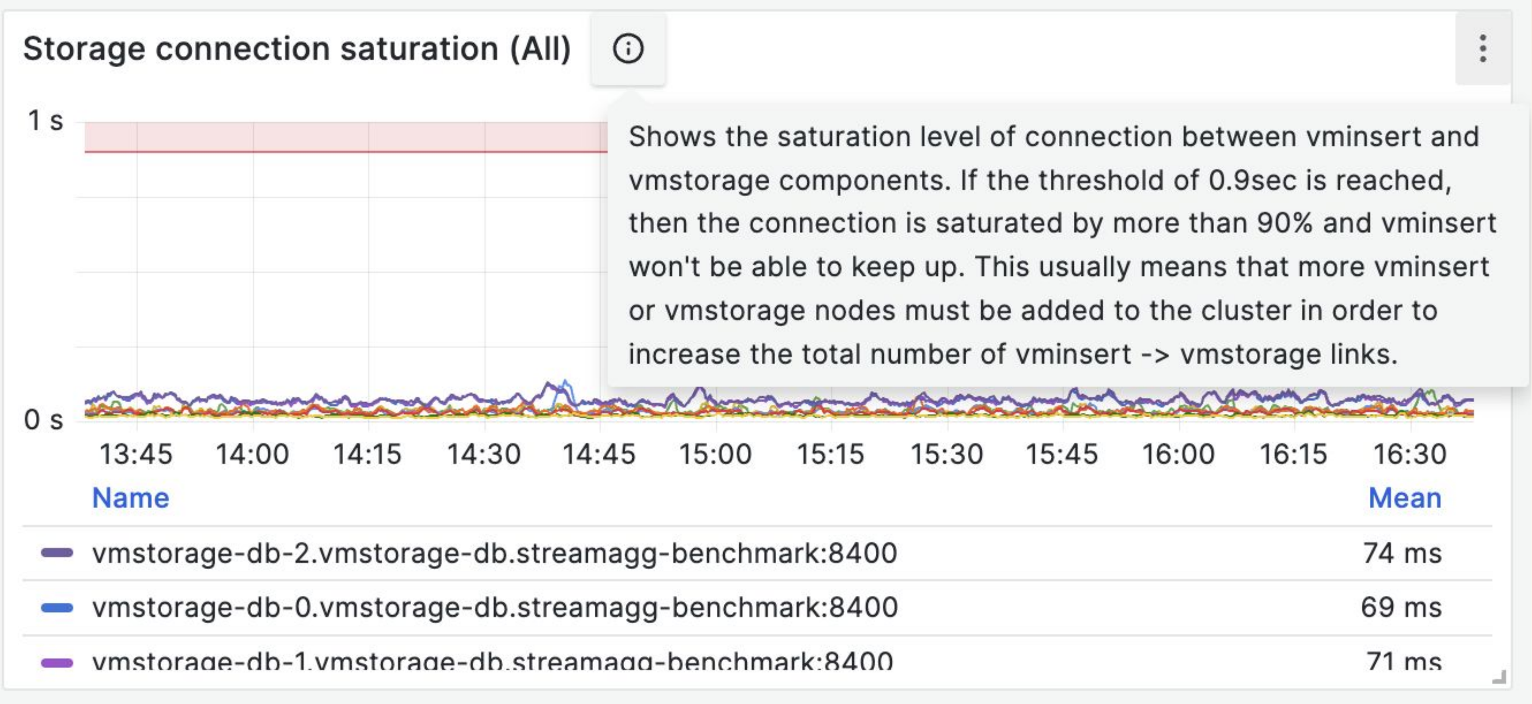
Non-default flags

job ↓	name	value	instance (uniqueValues)
vmselect-benchmark-vm-cluster	search.maxUniqueTimeseries	250000000	["10.71.6.22:8481", "10.71.7.1…
vmselect-benchmark-vm-cluster	search.maxSeries	250000000	["10.71.6.22:8481", "10.71.7.1…
vmselect-benchmark-vm-cluster	search.maxQueryDuration	10m0s	["10.71.6.22:8481", "10.71.7.1…
vmselect-benchmark-vm-cluster	search.maxConcurrentRequests	4	["10.71.6.22:8481", "10.71.7.1…
vmselect-benchmark-vm-cluster	search.logSlowQueryDuration	30s	["10.71.6.22:8481", "10.71.7.1



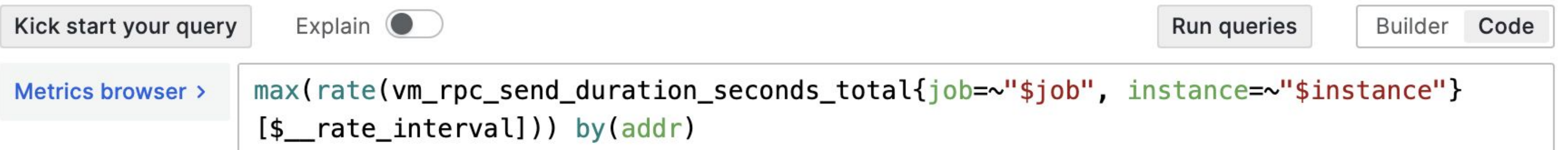
```
sum(flag{is_set="true", job=~"$job", instance=~"$instance"}) by(job, instance, name, value)
```

Show numbers that are easy to understand



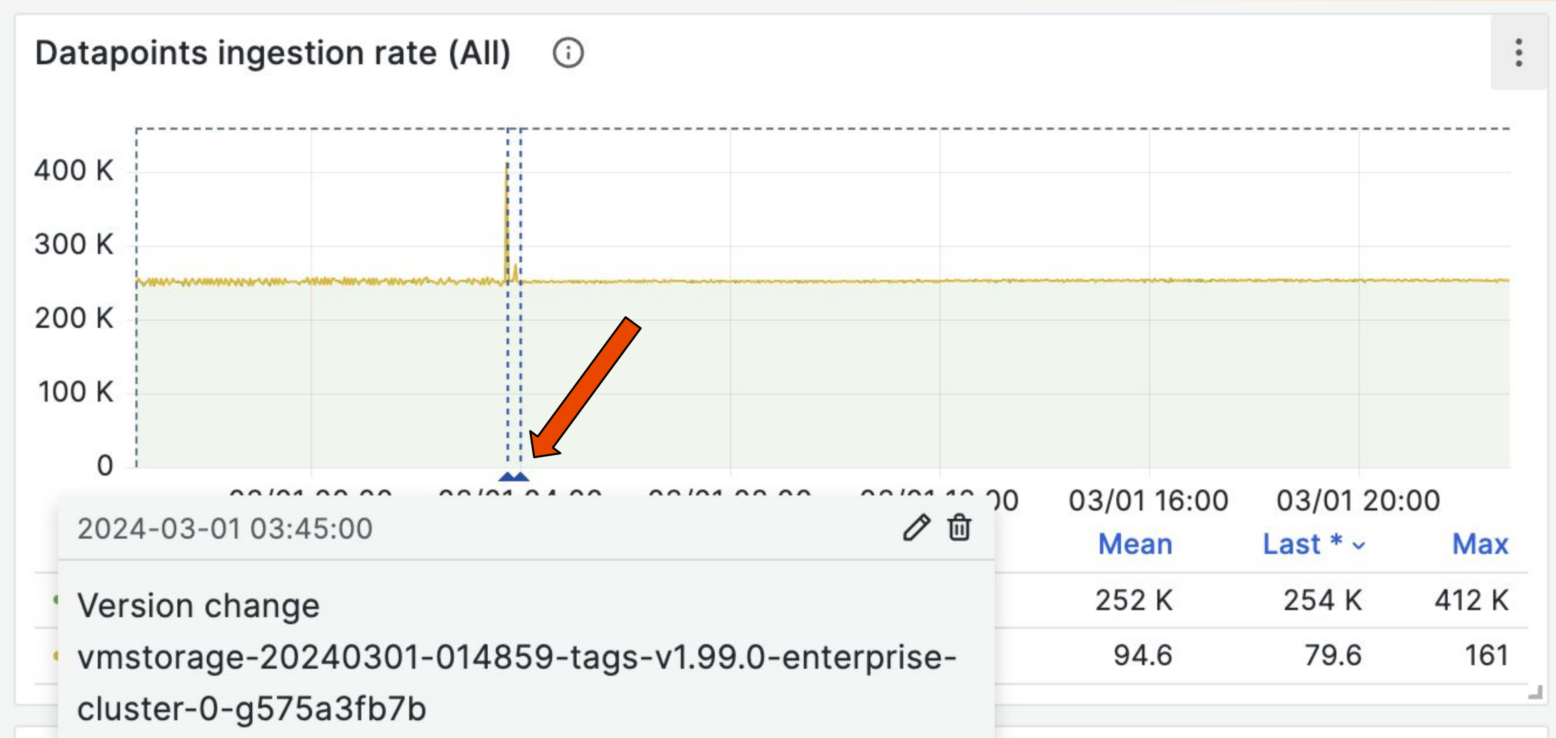
Shows the saturation level of connection between vminsert and vmstorage components. If the threshold of 0.9sec is reached, then the connection is saturated by more than 90% and vminsert won't be able to keep up. This usually means that more vminsert or vmstorage nodes must be added to the cluster in order to increase the total number of vminsert -> vmstorage links.

Name	Mean
vmstorage-db-2.vmstorage-db.streamagg-benchmark:8400	74 ms
vmstorage-db-0.vmstorage-db.streamagg-benchmark:8400	69 ms
vmstorage-db-1.vmstorage-db.streamagg-benchmark:8400	71 ms



```
max(rate(vm_rpc_send_duration_seconds_total{job=~"$job", instance=~"$instance"}
[$__rate_interval])) by(addr)
```

Annotations: show version change



2024-03-01 03:45:00

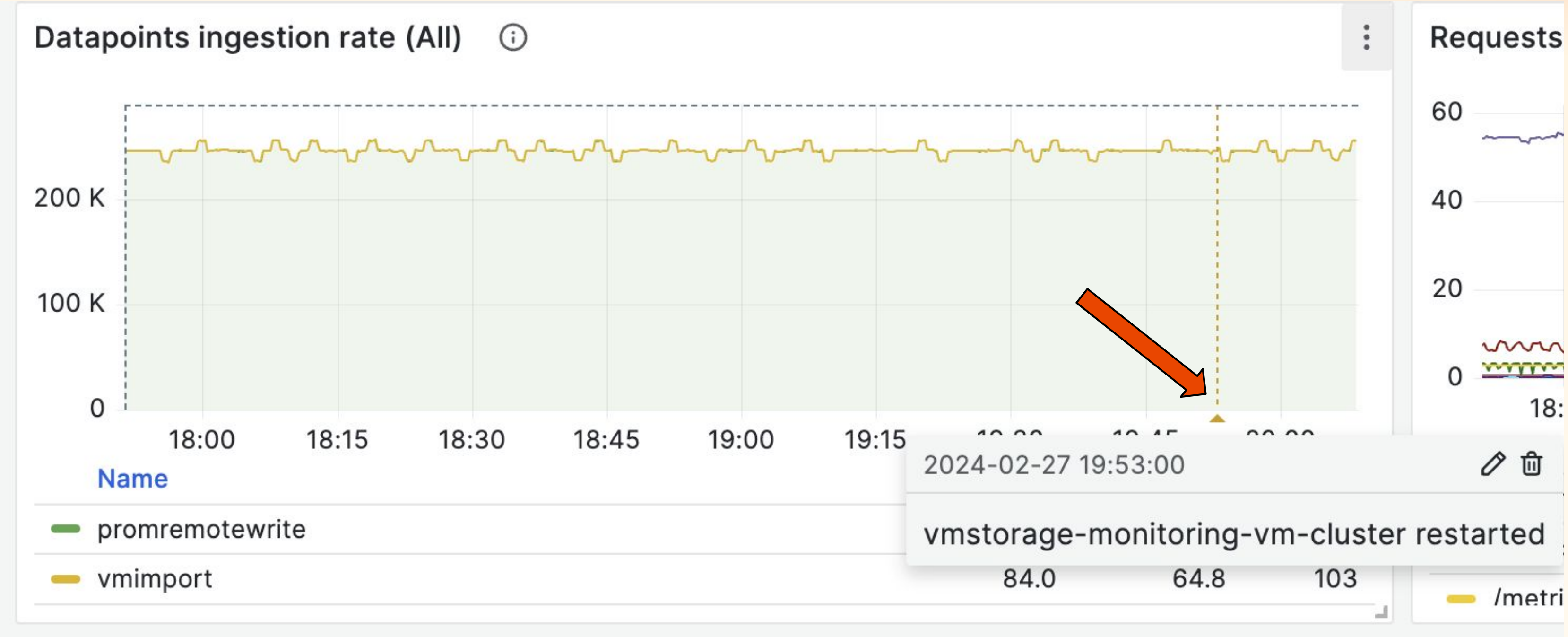
Version change

vmstorage-20240301-014859-tags-v1.99.0-enterprise-cluster-0-g575a3fb7b

	Mean	Last *	Max
	252 K	254 K	412 K
	94.6	79.6	161

```
sum(vm_app_version{job=~"$job", instance=~"$instance"}) by(version) unless (sum(vm_app_version
{job=~"$job", instance=~"$instance"} offset 20m) by(version))
```

Annotations: show process restarts



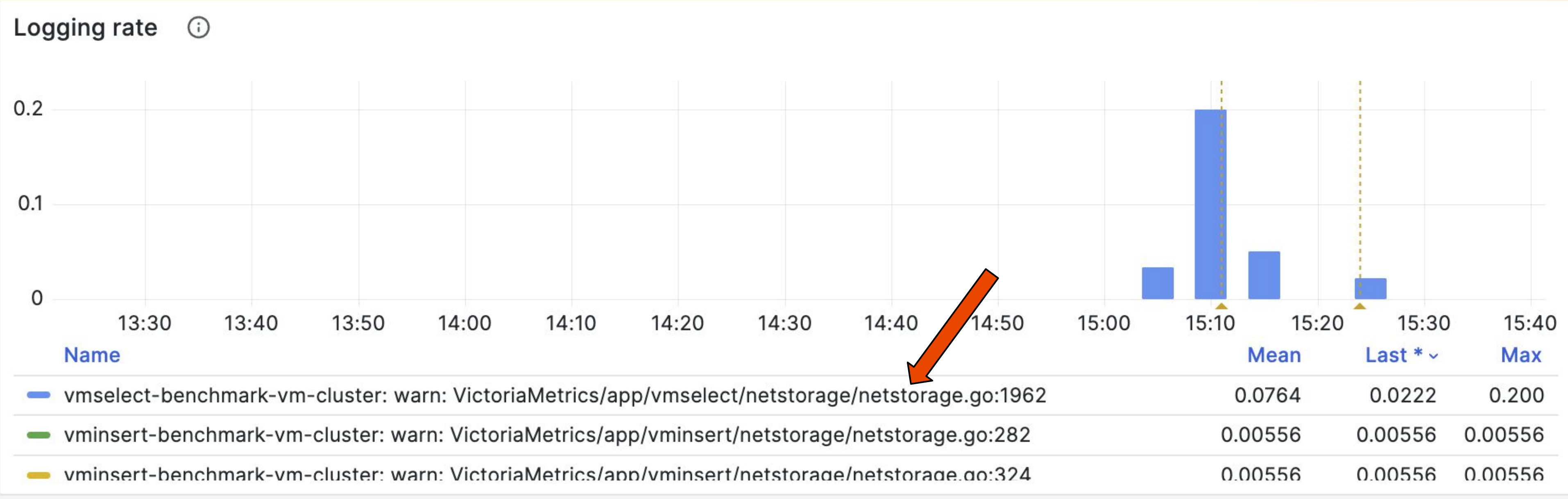
2024-02-27 19:53:00

vmstorage-monitoring-vm-cluster restarted

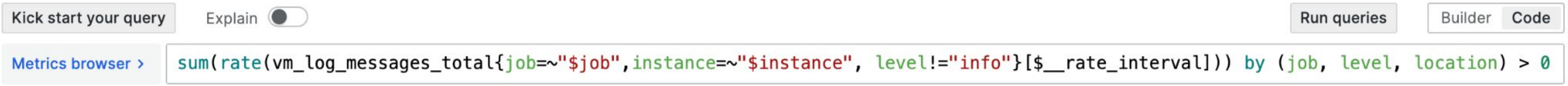
Name			
promremotewrite			
vmimport	84.0	64.8	103

```
sum(changes(vm_app_start_timestamp{job=~"$job", instance=~"$instance"})) by(job)
```

Track the number of logs via metrics



Name	Mean	Last *	Max
vmselect-benchmark-vm-cluster: warn: VictoriaMetrics/app/vmselect/netstorage/netstorage.go:1962	0.0764	0.0222	0.200
vminsert-benchmark-vm-cluster: warn: VictoriaMetrics/app/vminsert/netstorage/netstorage.go:282	0.00556	0.00556	0.00556
vminsert-benchmark-vm-cluster: warn: VictoriaMetrics/app/vminsert/netstorage/netstorage.go:324	0.00556	0.00556	0.00556



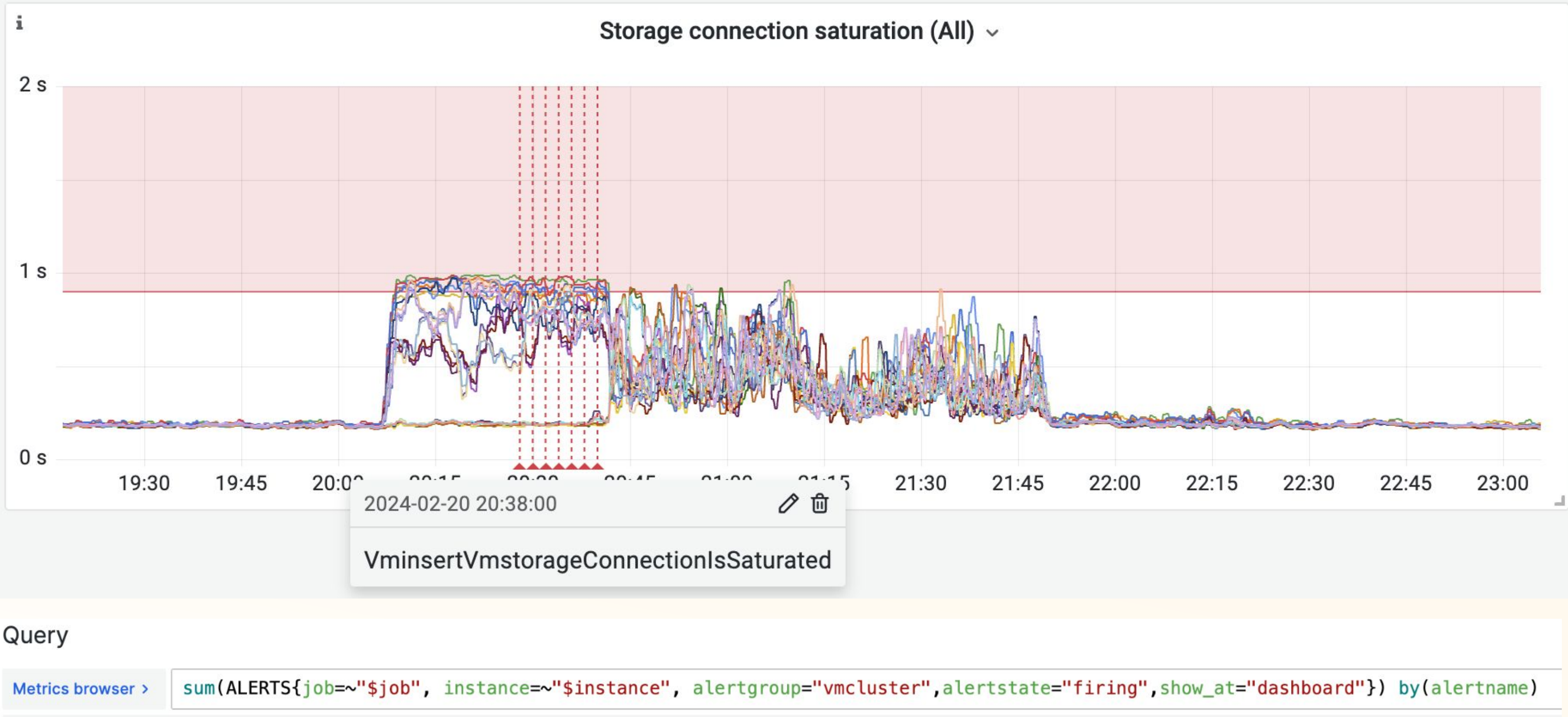
```
sum(rate(vm_log_messages_total{job=~"$job",instance=~"$instance", level!="info"}[$__rate_interval])) by (job, level, location) > 0
```

Software goes with alerting rules out of box

- Track critical metrics
- Contain both: context and explanation
- Are included to helm charts, and are available on github
- Interlinked with dashboards

```
-alert: VminsertVmstorageConnectionIsSaturated
 expr: rate(vm_rpc_send_duration_seconds_total[5m]) > 0.9
 for: 15m
 labels:
   show_at: dashboard
 annotations:
   dashboard:
"http://localhost:3000/d/oS7Bi_0Wz?viewPanel=139&var-instance={{
$labels.instance }}"
   summary: "Connection between vminsert on {{ $labels.instance }} and
vmstorage on {{ $labels.addr }} is saturated"
   description: "The connection between vminsert (instance {{
$labels.instance }}) and vmstorage (instance {{ $labels.addr }})
     is saturated by more than 90% and vminsert won't be able to keep up
     This usually means that more vminsert or vmstorage nodes must be
added to the cluster in order to increase
     the total number of vminsert -> vmstorage links."
```

Annotation displays the active alerts



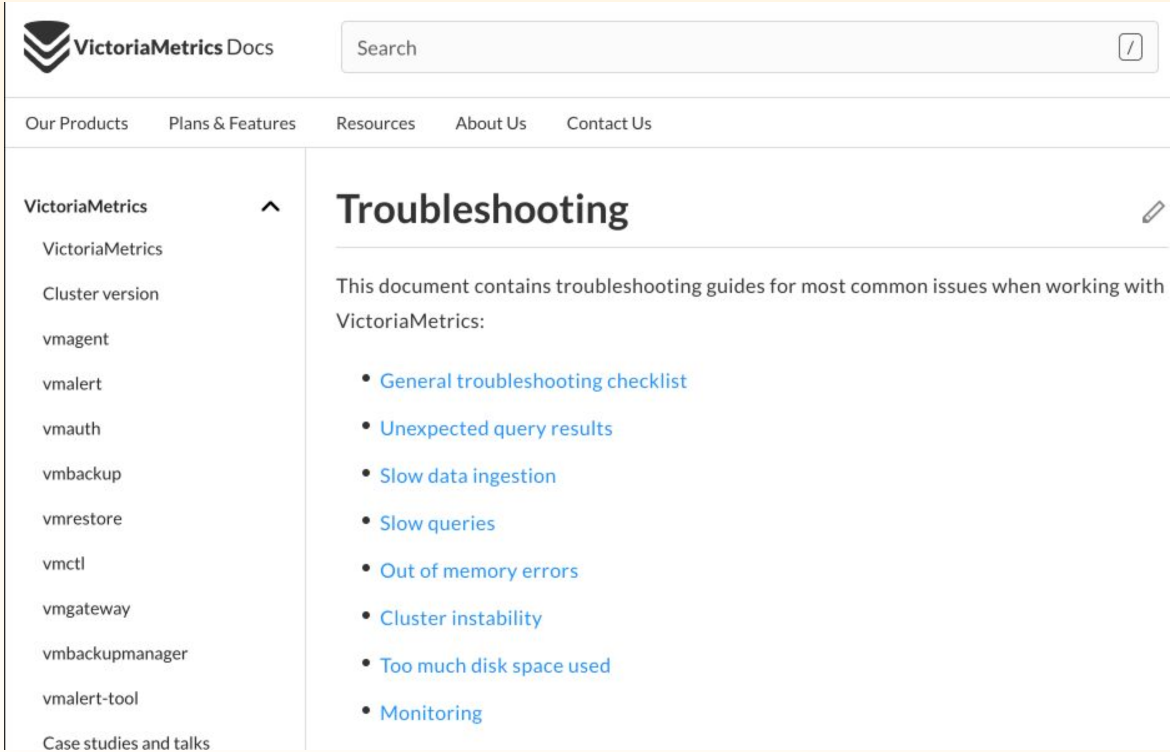
2024-02-20 20:38:00

VminsertVmstorageConnectionIsSaturated

Query

```
sum(ALERTS{job=~"$job", instance=~"$instance", alertgroup="vmcluster",alertstate="firing",show_at="dashboard"}) by(alertname)
```

Documentation



VictoriaMetrics Docs

Search

Our Products · Plans & Features · Resources · About Us · Contact Us

- VictoriaMetrics
  - VictoriaMetrics
  - Cluster version
  - vmagent
  - vmalert
  - vmauth
  - vmbackup
  - vmrestore
  - vmctl
  - vmgateway
  - vmbackupmanager
  - vmalert-tool
  - Case studies and talks

Troubleshooting

This document contains troubleshooting guides for most common issues when working with VictoriaMetrics:

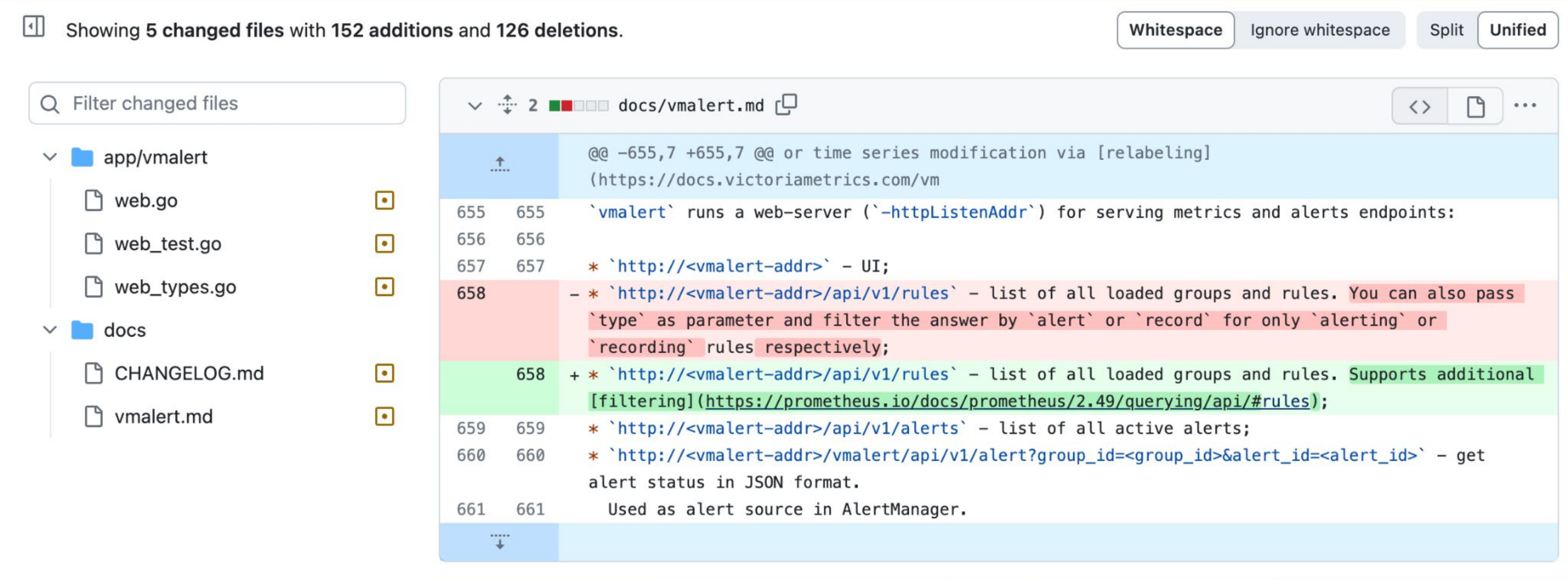
- General troubleshooting checklist
- Unexpected query results
- Slow data ingestion
- Slow queries
- Out of memory errors
- Cluster instability
- Too much disk space used
- Monitoring

```
wc -l docs/**/*.md
...
91053 total
```

Change should always be reflected in the docs

- The actual change
- Tests to test the change

- Update documentation
- Mention in the CHANGELOG



Showing **5 changed files** with **152 additions** and **126 deletions.**

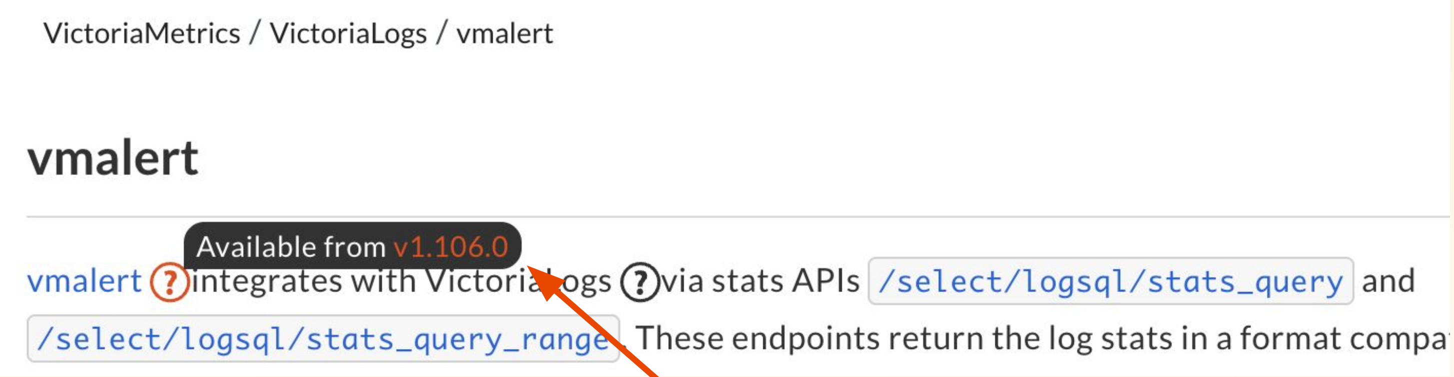
- app/vmalert
  - web.go
  - web_test.go
  - web_types.go
- docs
  - CHANGELOG.md
  - vmalert.md

2 docs/vmalert.md

```
@@ -655,7 +655,7 @@ or time series modification via [relabeling]
(https://docs.victoriametrics.com/vm
`vmalert` runs a web-server (`-httpListenAddr`) for serving metrics and alerts endpoints:

* `http://<vmalert-addr>` - UI;
- * `http://<vmalert-addr>/api/v1/rules` - list of all loaded groups and rules. You can also pass
`type` as parameter and filter the answer by `alert` or `record` for only `alerting` or
`recording` rules respectively;
+ * `http://<vmalert-addr>/api/v1/rules` - list of all loaded groups and rules. Supports additional
[filtering](https://prometheus.io/docs/prometheus/2.49/querying/api/#rules);
* `http://<vmalert-addr>/api/v1/alerts` - list of all active alerts;
* `http://<vmalert-addr>/vmalert/api/v1/alert?group_id=<group_id>&alert_id=<alert_id>` - get
alert status in JSON format.
  Used as alert source in AlertManager.
```

Shortcode for versioning features



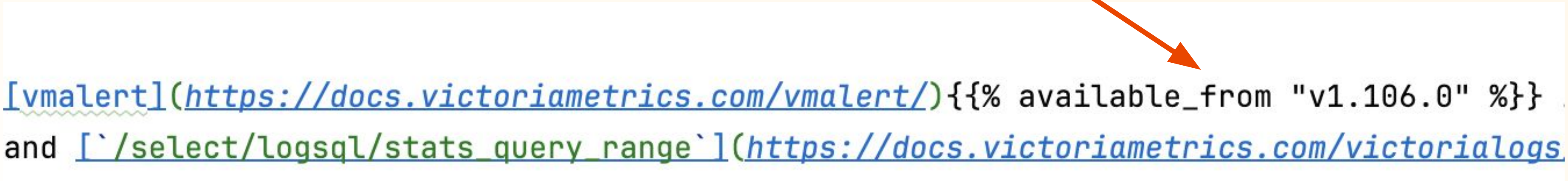
VictoriaMetrics / VictoriaLogs / vmalert

vmalert

Available from v1.106.0

vmalert ? integrates with VictoriaLogs ? via stats APIs `/select/logsql/stats_query` and `/select/logsql/stats_query_range`. These endpoints return the log stats in a format compa

When this feature was released



```
[vmalert](https://docs.victoriametrics.com/vmalert/){{% available_from "v1.106.0" %}} 
and [`/select/logsql/stats_query_range`](https://docs.victoriametrics.com/victorialogs
```

Cross-refer documentation to keep it fresh

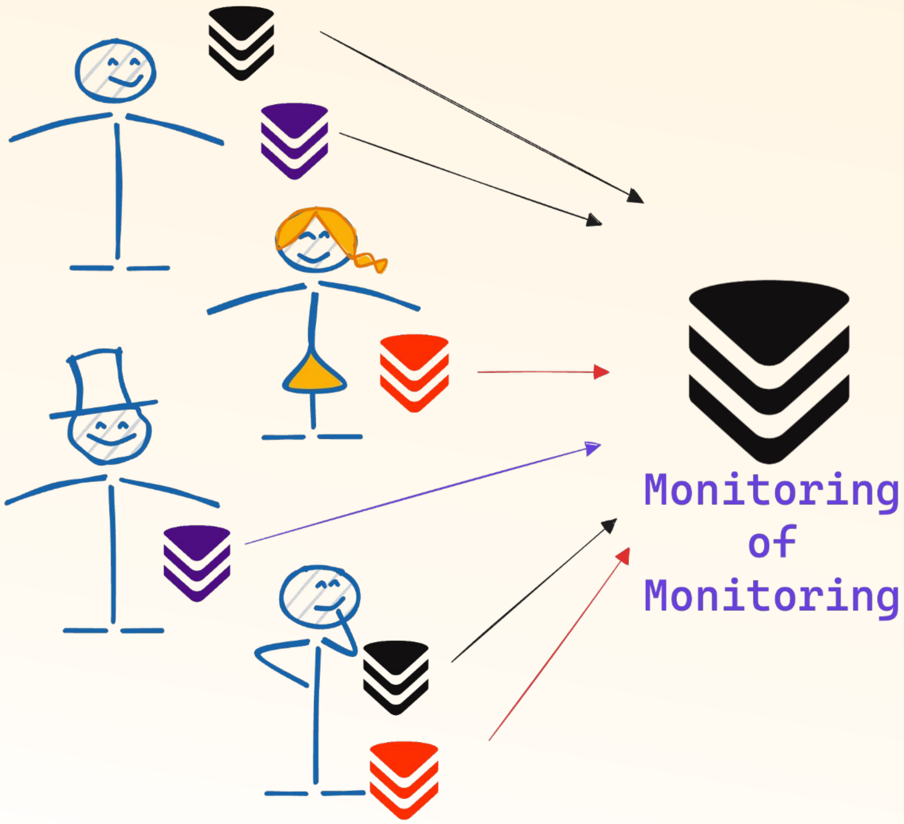
We use links to documentation in the following places:

- Error messages and logs
- Alerting rules
- Dashboards
- Github issues
- Community platforms (slack, reddit, stackoverflow)

We do care about broken links and always try to keep backward compatibility.

We do care about our docs, dashboards, alerts because we use them every day!

Monitoring of Monitoring (MoM)

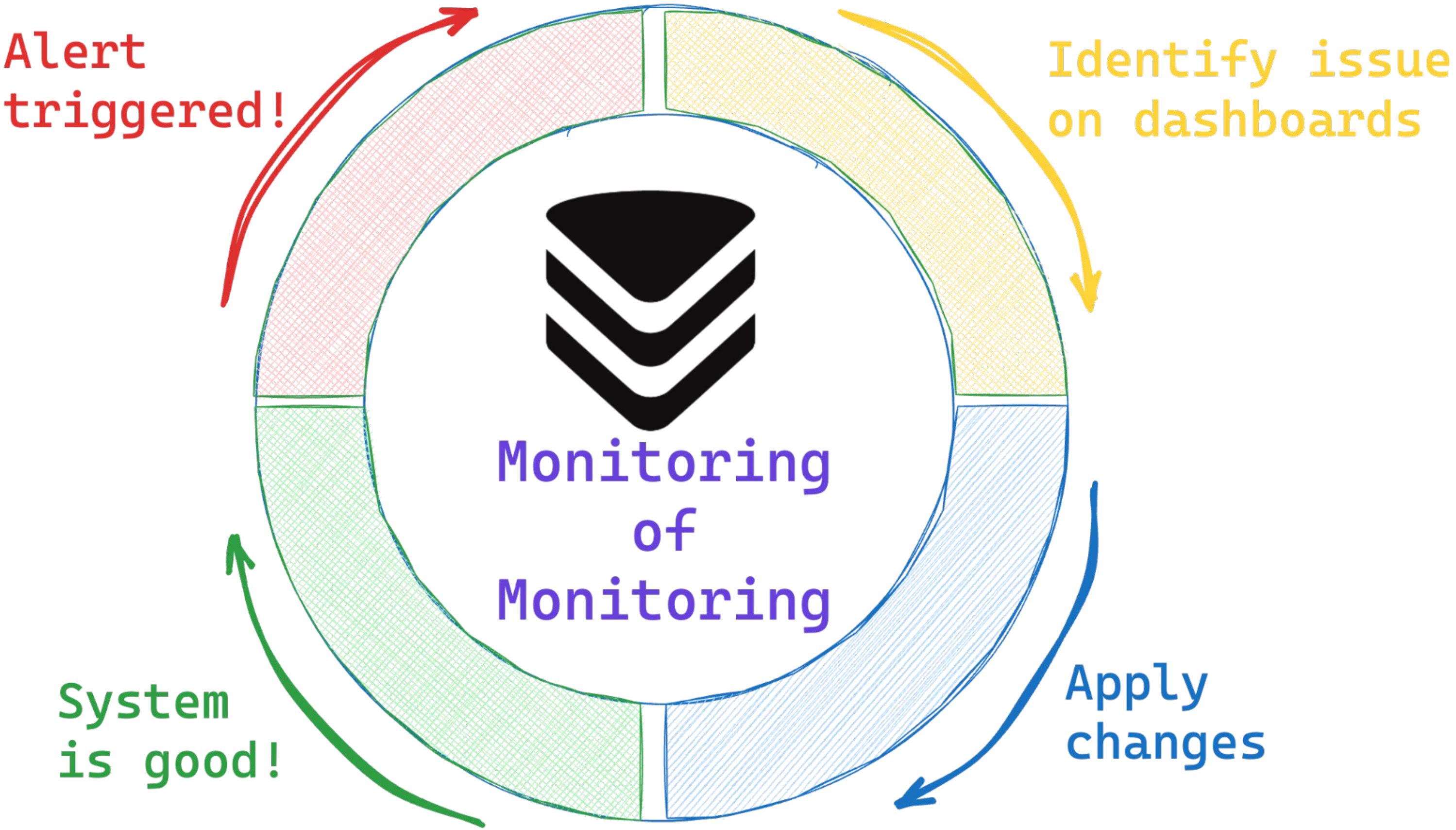


Monitoring of Monitoring

Dashboards

Alerts

VictoriaMetrics Docs



Monitoring of Monitoring

Alert triggered!

Identify issue on dashboards

Apply changes

System is good!

Additional materials

1. Live Grafana dashboard from our playground
2. List of Grafana dashboards for VictoriaMetrics components
3. List of alerting rules
4. Documentation

Questions?

- https://github.com/VictoriaMetrics
- https://github.com/hagen1778