



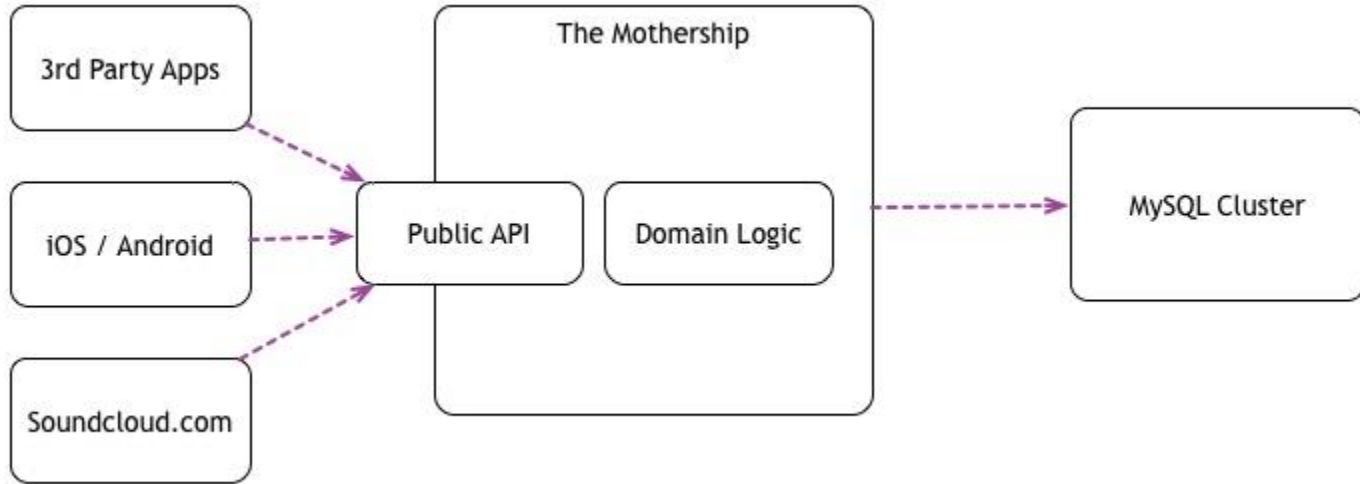
Prometheus: A Next-Generation Monitoring System

Björn Rabenstein, Julius Volz
SoundCloud

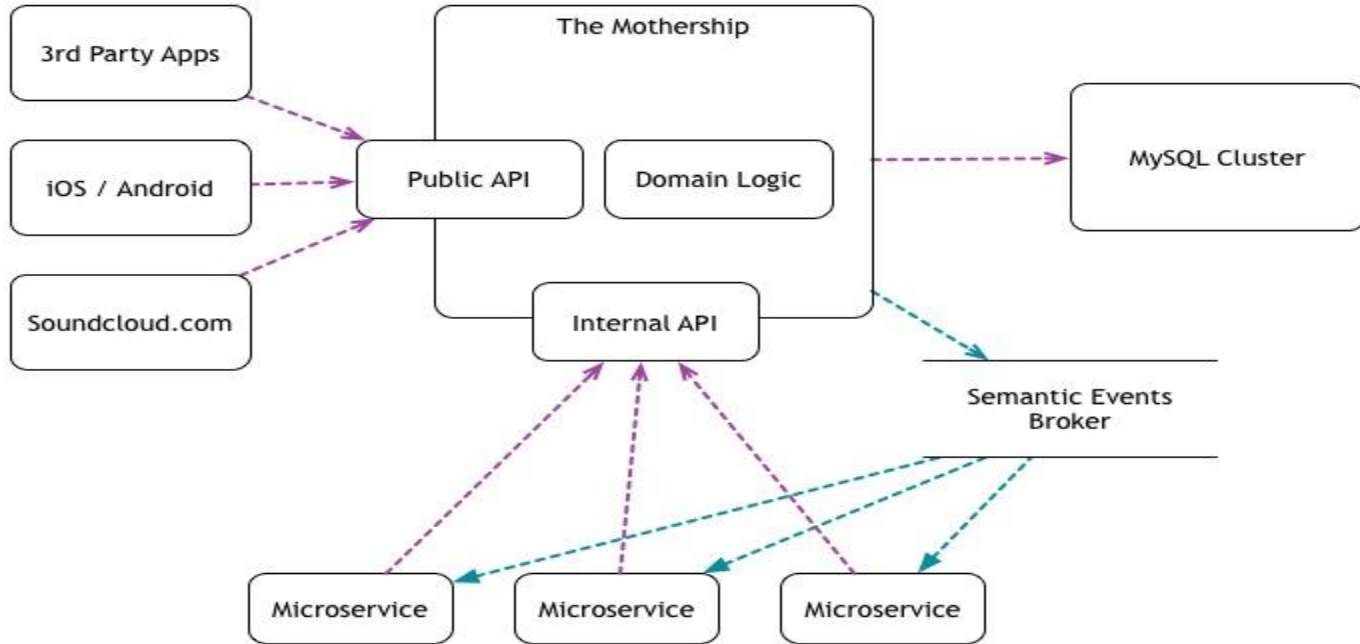
SREcon Dublin
May 14, 2015

History

From a monolith...



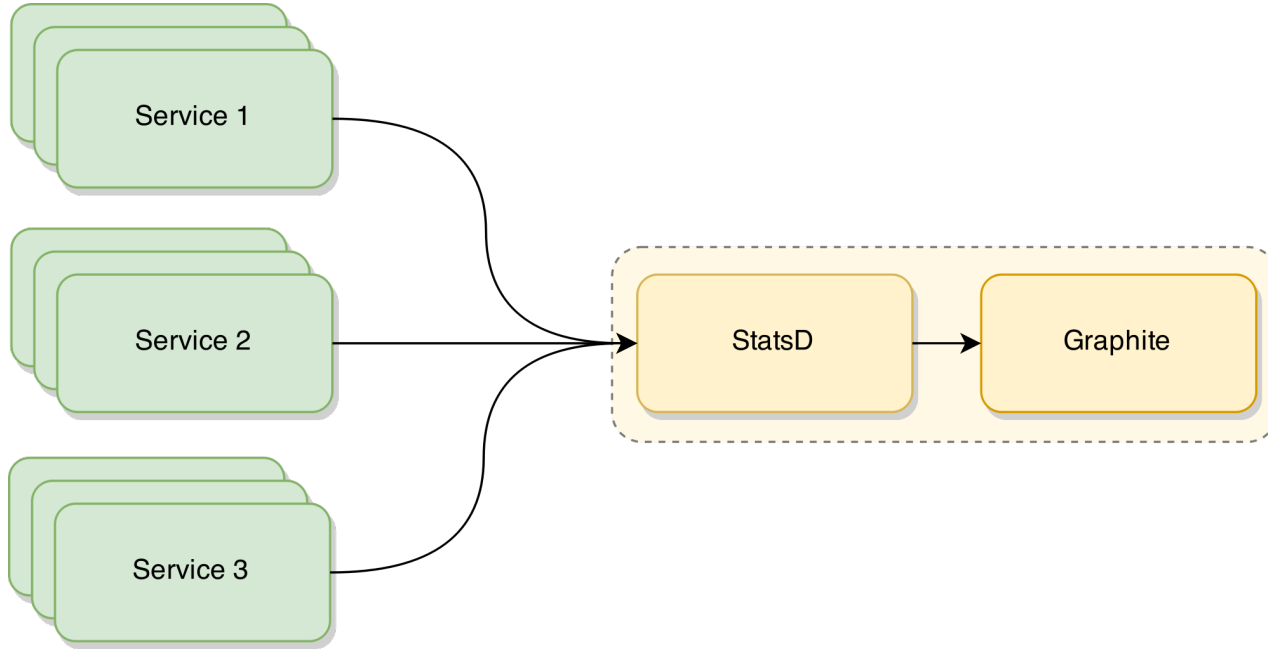
...to microservices



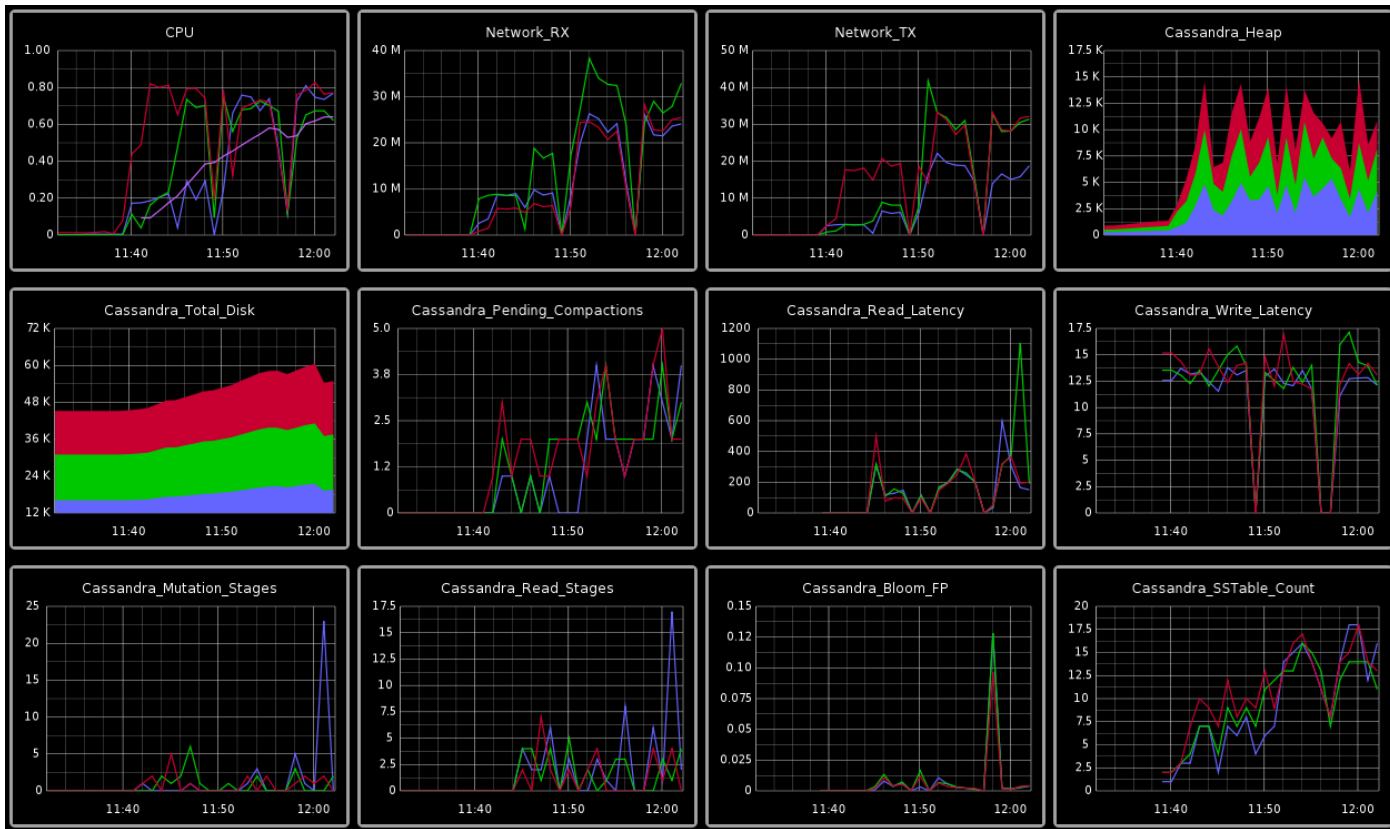
Existing monitoring setup

in 2012...

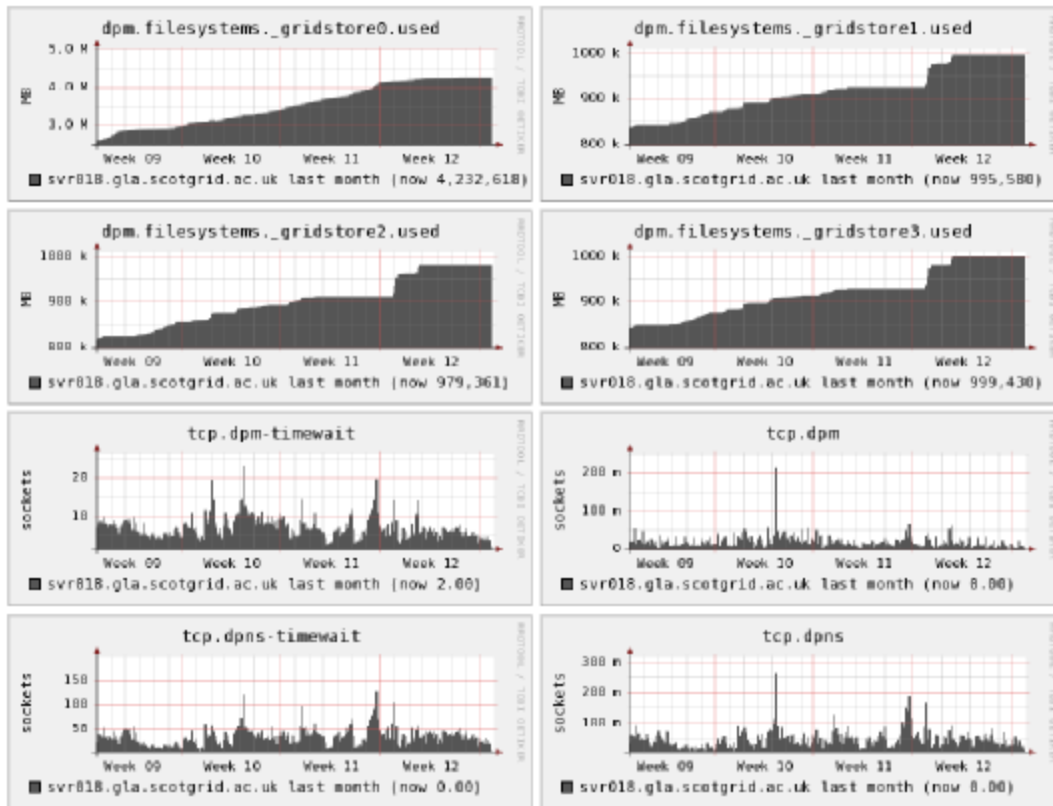
Service monitoring



Service monitoring



Host monitoring



Alerting

The screenshot displays the Nagios web interface in a Mozilla Firefox browser window. The address bar shows the URL `http://192.168.130.223/nagios/`. The interface includes a navigation sidebar on the left with categories like General, Monitoring, Service Problems, and Reporting. The main content area shows a table of monitored hosts and their services.

Host	Service	Status	Last Check	Next Check	Duration	Latency	Output
webprod03	Monitor	OK	01-26-2007 14:58:59	0d 4h 53m 23s	1/4		USERS OK - 1 users currently logged in
	Check Users	OK	01-26-2007 14:58:59	0d 4h 53m 23s	1/4		OK - load average: 0.21, 0.08, 0.05
	Current Load	OK	01-26-2007 14:59:54	0d 4h 53m 23s	1/4		OK: Memory Usage 56% - Total: 511 MB, Used: 287 MB, Free: 224 MB
	Memory Usage	OK	01-26-2007 14:55:29	0d 4h 53m 23s	1/4		PING OK - Packet loss = 0%, RTA = 0.16 ms
	PING	OK	01-26-2007 14:56:14	0d 4h 50m 23s	1/4		DISK OK [243816 kB (6%) free on /dev/sda2]
	Root Partition	OK	01-26-2007 14:57:09	0d 4h 50m 33s	1/4		Swap ok - (null) 0% (0 out of 16386)
	SWAP Usage	OK	01-26-2007 14:57:44	0d 4h 50m 33s	1/4		OK - 95 processes running
webprod04	Total Processes	OK	01-26-2007 14:58:29	0d 4h 50m 33s	1/4		OK - 95 processes running
	Xen Virtual Machine Monitor	CRITICAL	01-26-2007 14:59:04	0d 0h 44m 34s	4/4		Critical Xen VMs Usage - Total NB: 0 - detected VMs:
	Check Users	OK	01-26-2007 14:59:54	0d 0h 15m 33s	1/4		USERS OK - 2 users currently logged in
	Current Load	OK	01-26-2007 14:55:34	0d 0h 14m 53s	1/4		OK - load average: 0.30, 0.60, 0.44
	Memory Usage	OK	01-26-2007 14:56:19	0d 0h 14m 13s	1/4		OK: Memory Usage 37% - Total: 511 MB, Used: 190 MB, Free: 321 MB
	PING	OK	01-26-2007 14:57:10	0d 0h 13m 23s	1/4		PING OK - Packet loss = 0%, RTA = 0.27 ms
	Root Partition	OK	01-26-2007 14:57:49	0d 0h 12m 43s	1/4		DISK OK [3948940 kB (94%) free on /dev/sda2]
webprod05	SWAP Usage	OK	01-26-2007 14:58:34	0d 0h 11m 53s	1/4		Swap ok - (null) 0% (0 out of 16386)
	Total Processes	OK	01-26-2007 14:59:09	0d 0h 16m 22s	1/4		OK - 250 processes running
	Xen Virtual Machine Monitor	WARNING	01-26-2007 14:58:54	0d 0h 1m 33s	4/4		Warning Xen VMs Usage - Total NB: 1 - detected VMs: migrating-xen-vm4
	PING	OK	01-26-2007 14:55:39	0d 0h 24m 58s	1/4		PING OK - Packet loss = 0%, RTA = 0.25 ms
	Xen Virtual Machine Monitor	OK	01-26-2007 14:59:54	0d 0h 0m 33s	1/4		OK: Xen Hypervisor "webprod05" is running 4 Xen VMs: xen-vm1 xen-vm2 xen-vm3 xen-vm4
	Check Users	OK	01-26-2007 14:58:09	0d 0h 17m 23s	1/4		USERS OK - 1 users currently logged in
	Current Load	OK	01-26-2007 14:57:54	0d 3h 16m 21s	1/4		OK - load average: 1.54, 1.09, 0.48
xen-vm1	Memory Usage	OK	01-26-2007 14:58:39	0d 3h 15m 41s	1/4		OK: Memory Usage 8% - Total: 8195 MB, Used: 676 MB, Free: 7519 MB
	PING	OK	01-26-2007 14:59:15	0d 3h 15m 21s	1/4		PING OK - Packet loss = 0%, RTA = 0.49 ms
	Root Partition	OK	01-26-2007 14:59:59	0d 3h 14m 51s	1/4		DISK OK [4196280 kB (99%) free on udev]
	SWAP Usage	OK	01-26-2007 14:55:44	0d 3h 14m 1s	1/4		Swap ok - (null) 0% (0 out of 2055)
	Total Processes	OK	01-26-2007 14:57:29	0d 0h 18m 3s	1/4		OK - 88 processes running
	Check Users	OK	01-26-2007 14:57:15	0d 3h 7m 41s	1/4		USERS OK - 0 users currently logged in
	Current Load	OK	01-26-2007 14:57:59	0d 3h 7m 1s	1/4		OK - load average: 0.00, 0.00, 0.00
xen-vm2	Memory Usage	OK	01-26-2007 14:58:44	0d 3h 6m 21s	1/4		OK: Memory Usage 6% - Total: 1023 MB, Used: 64 MB, Free: 958 MB
	PING	OK	01-26-2007 14:59:19	0d 0h 48m 14s	1/4		PING OK - Packet loss = 0%, RTA = 0.43 ms
	Root Partition	OK	01-26-2007 15:00:05	0d 1h 15m 4s	1/4		DISK OK [524220 kB (99%) free on udev]
	SWAP Usage	OK	01-26-2007 14:55:49	0d 3h 9m 41s	1/4		Swap ok - (null) 0% (0 out of 2055)
	Total Processes	OK	01-26-2007 14:56:34	0d 3h 9m 1s	1/4		OK - 52 processes running

Not fun

Prometheus

Multi-dimensional data model

```
api_http_requests_total{method="GET", endpoint="/api/tracks", status="200"} 2034834
```

(like OpenTSDB)

Operational simplicity

```
$ go build
```

```
$ ./prometheus
```

(not like OpenTSDB)

Scalable data collection

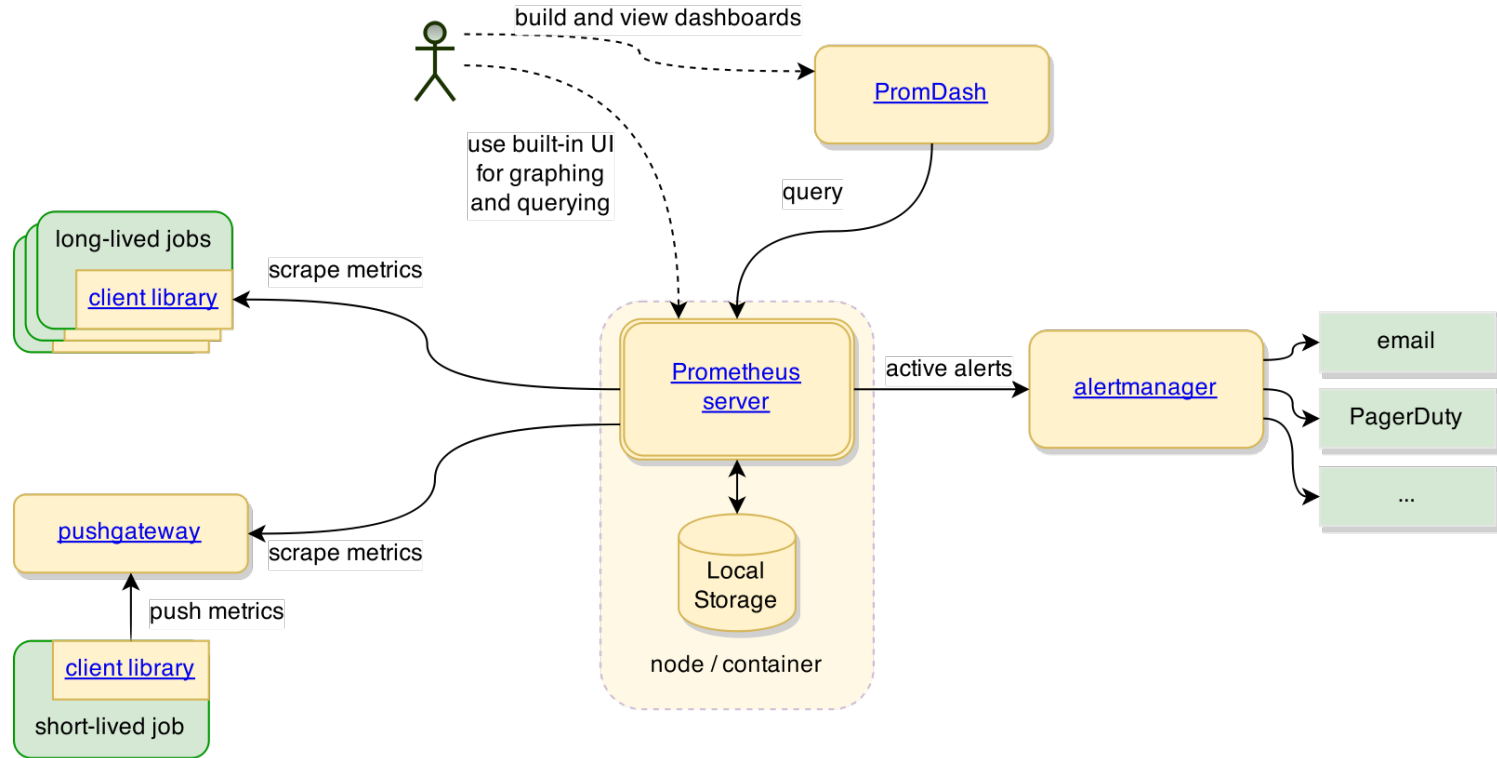
Thousands of targets.
Hundreds of thousands of samples per second.
Millions of time series.
On a single monitoring server.
Running many servers is easy, too...
Pull, not push.

Powerful query language

```
topk(3, sum(rate(bazooka_instance_cpu_time_ns[5m])) by (app, proc))
```

```
sort_desc(sum(bazooka_instance_memory_limit_bytes -  
    bazooka_instance_memory_usage_bytes) by (app, proc))
```

Architecture



Expression browser

[Prometheus](#)[Alerts](#)[Graph](#)[Status](#)[Help](#)

```
sort_desc(sum(bazooka_instance_memory_limit_bytes - bazooka_instance_memory_usage_bytes) by (app, proc)) / 1024 / 1024 / 1024
```

[Execute](#)[Graph](#)[Console](#)

Element	Value
{app="harsh-dagger",proc="api"}	132.720802
{app="quality-locomotive",proc="web"}	89.547081
{app="husky-long-oyster",proc="web"}	68.982738
{app="vital-albatross",proc="api"}	48.033772
{app="autopsy-gutsy",proc="widget"}	47.410583
{app="western-python",proc="cruncher"}	40.126926
{app="harsh-dagger",proc="api"}	28.527714
{app="outstanding-dagger",proc="api"}	26.119423
{app="gruesome-waterbird",proc="web"}	17.666714
{app="gutsy-square",proc="public"}	15.296242
{app="harsh-dagger",proc="web"}	14.738327
{app="northern-electron",proc="api"}	13.349815

Built-in graphing

Prometheus

Alerts

Graph

Status

Help

```
topk(3, sum(rate(bazooka_instance_cpu_time_ns[5m])) by (app, proc))
```

Execute

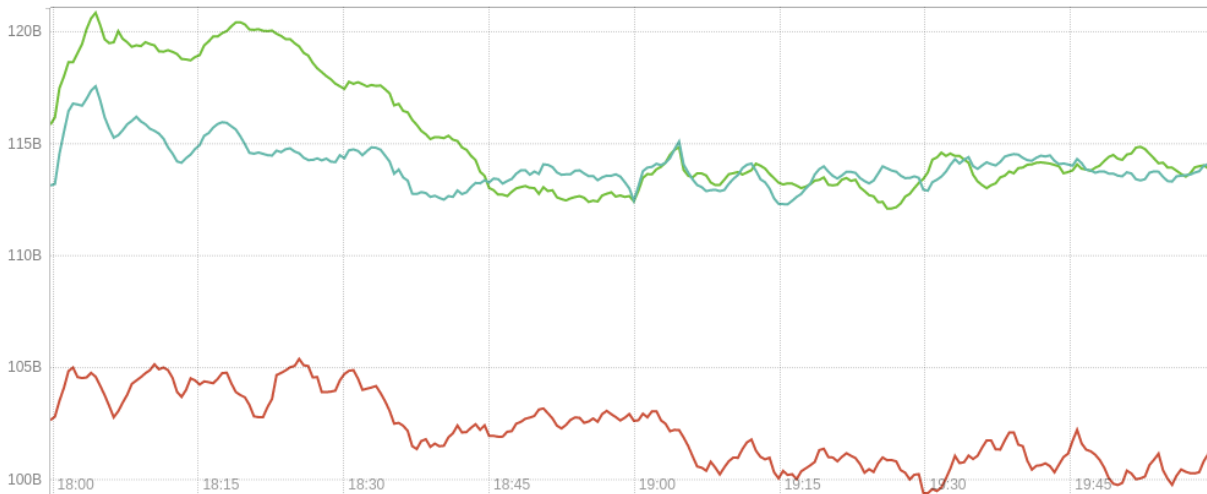
Graph Console

- 2h +

◀ Until ▶

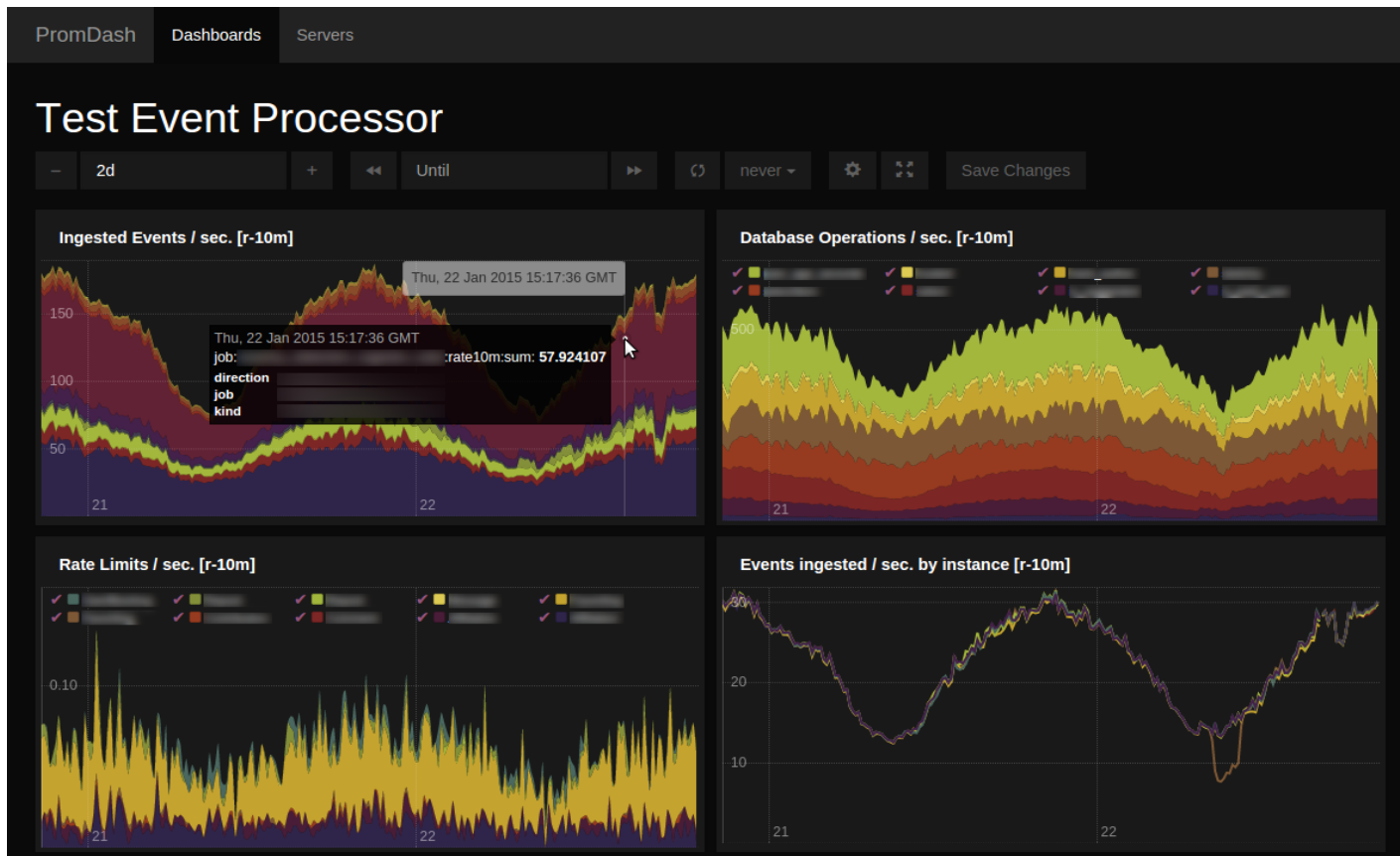
Res. (s)

stacked



{app="curious-turtle",proc="api"}
 {app="fast-bunny",proc="web"}
 {app="lunatic-fox",proc="web"}

PromDash



Alertmanager

Alert Manager

Alerts

Silences

Status

Alerts

Alert	Labels	Active Since	Last Refreshed	Generated By	Alerting Rule	Silenced
SomethingIsDown Silence Alert	<code>job="node"</code> <code>service="foobar"</code> <code>severity="page"</code> Silence Instance	8m0.4s ago	400ms ago	http://desktop:9090/graph#%5B%7B%22expr...	ALERT SomethingIsDown IF (COUNT((up == 0)) BY (job) > 1) FOR 1m WITH {service="foobar", severity="page"}	not silenced
HighRPCRate Silence Alert	<code>group="production"</code> <code>instance="localhost:8081"</code> <code>job="example-random"</code> <code>monitor="tutorial-monitor"</code> <code>service="exponential"</code> <code>severity="page"</code> Silence Instance	30.4s ago	400ms ago	http://desktop:9090/graph#%5B%7B%22expr...	ALERT HighRPCRate IF (rate(rpc_durations_microseconds_count[5m]) > 12) FOR 15s WITH {severity="page"}	not silenced
HighRPCRate Silence Alert	<code>group="production"</code> <code>instance="localhost:8080"</code> <code>job="example-random"</code> <code>monitor="tutorial-monitor"</code> <code>service="exponential"</code> <code>severity="page"</code> Silence Instance	30.4s ago	400ms ago	http://desktop:9090/graph#%5B%7B%22expr...	ALERT HighRPCRate IF (rate(rpc_durations_microseconds_count[5m]) > 12) FOR 15s WITH {severity="page"}	not silenced
HighRPCRate Silence Alert	<code>group="canary"</code> <code>instance="localhost:8082"</code> <code>job="example-random"</code> <code>monitor="tutorial-monitor"</code> <code>service="exponential"</code> <code>severity="page"</code> Silence Instance	30.4s ago	400ms ago	http://desktop:9090/graph#%5B%7B%22expr...	ALERT HighRPCRate IF (rate(rpc_durations_microseconds_count[5m]) > 12) FOR 15s WITH {severity="page"}	not silenced

Exporters

vs.

direct instrumentation

Client libraries

Official

Go
Java (JVM)
Ruby
Python

Unofficial

.NET / C#
Node.js
Bash

Integrations

Official exporters

Node/system metrics exporter
Graphite exporter
Collectd exporter
JMX exporter
HAProxy exporter
StatsD bridge
AWS CloudWatch exporter
Hystrix metrics publisher
Mesos task exporter
Consul exporter

Unofficial exporters

RethinkDB exporter
Redis exporter
collector exporter
MongoDB exporter
Django exporter
Google's mtail log data extractor
Minecraft exporter module
Meteor JS web framework exporter

Direct instrumentation

cAdvisor
Kubernetes
Kubernetes-Mesos
EtcD
gokit
go-metrics instrumentation library
...

Further reading

prometheus.io

developers.soundcloud.com/blog

boxever.com/tech-blog

Acknowledgements

Many contributors at SoundCloud...
...and by now so many more from everywhere.

Special thanks to:

Matt T. Proud (founding father)

Brian Brazil (Boxever)

Johannes “fish” Ziemke (Docker)