From Alert Notification to Comparison of Good and Bad Requests In One Click

Kubecon EU, 2020-08-18
Shreyas Srivatsan

Who am I?

Shreyas Srivatsan

Technical Lead @Chronosphere

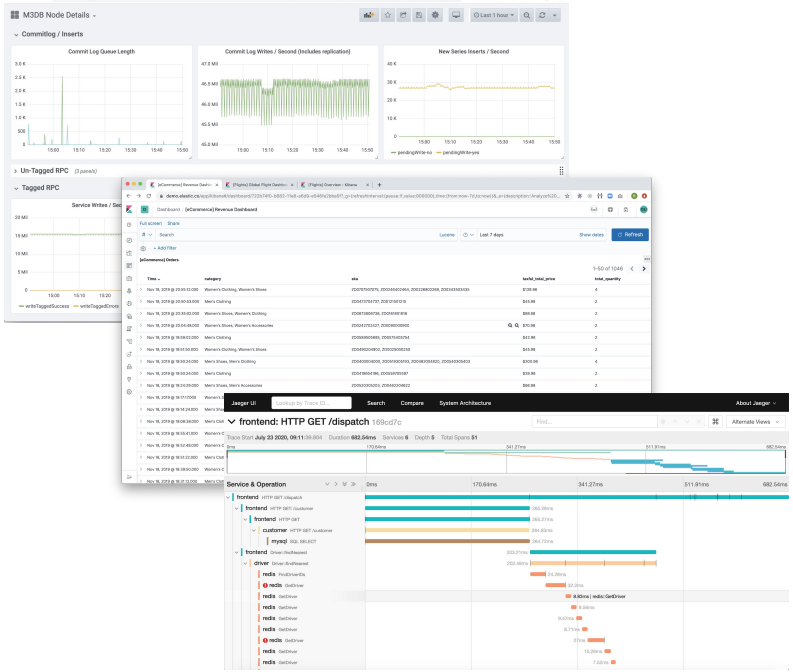
- Hosted metrics & monitoring platform
- Large scale, high throughput use cases
- Built on M3

Previously Observability @Uber

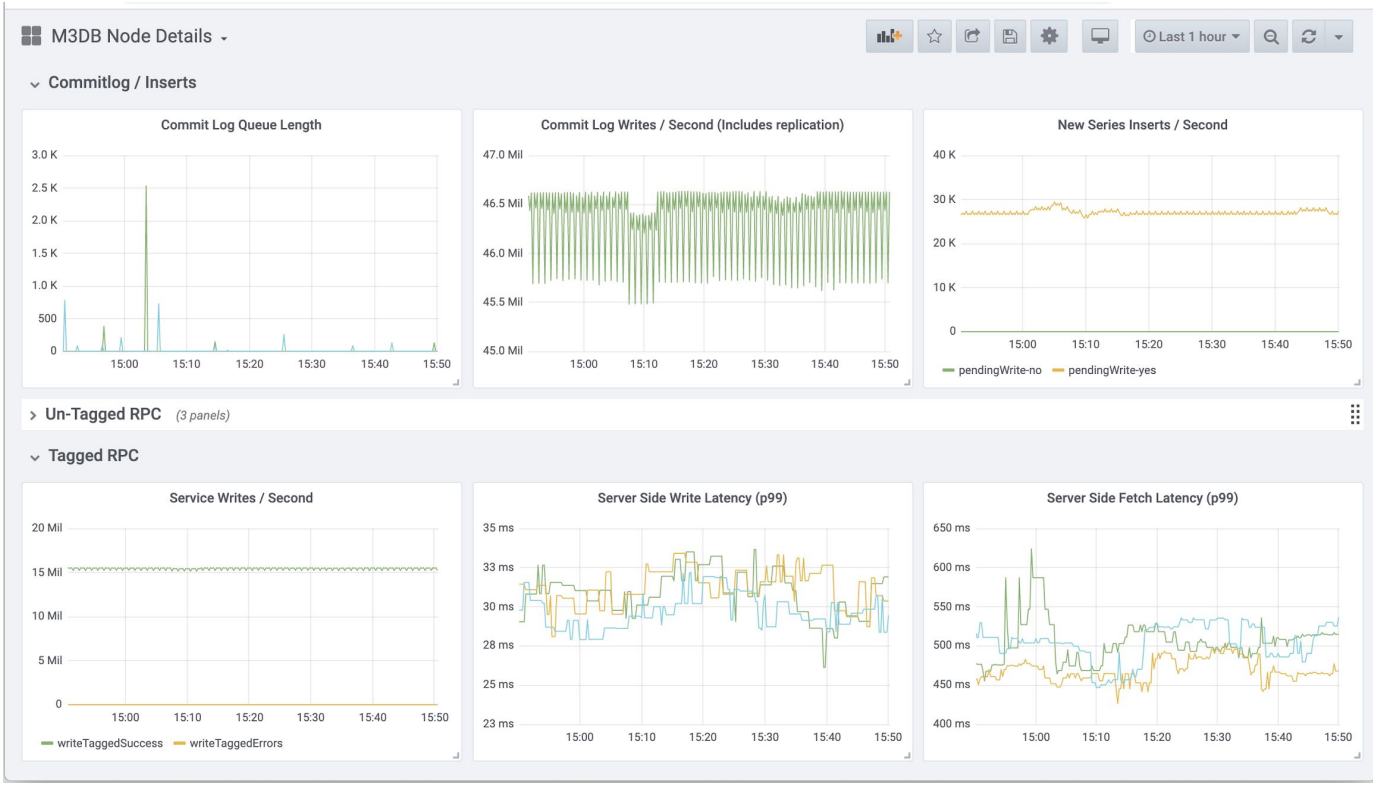
Agenda

1. Today:
   a. Observability Signals
   b. What Happens When You Get Alerted?
2. The Journey: Deep Linking Metrics and Traces
3. Tomorrow: Jumping from an Alert to a Request Comparison

Today: Observability Signals



Metrics..


M3DB Node Details
Last 1 hour
Commitlog / Inserts
Commit Log Queue Length
3.0 K
2.5 K
2.0 K
1.5 K
1.0 K
500
0
15:00
15:10
15:20
15:30
15:40
15:50
Commit Log Writes / Second (Includes replication)
47.0 Mil
46.5 Mil
46.0 Mil
45.5 Mil
45.0 Mil
New Series Inserts / Second
40 K
30 K
20 K
10 K
0
pendingWrite-no
pendingWrite-yes
Un-Tagged RPC
(3 panels)
Tagged RPC
Service Writes / Second
20 Mil
15 Mil
10 Mil
5 Mil
writeTaggedSuccess
writeTaggedErrors
Server Side Write Latency (p99)
35 ms
33 ms
30 ms
28 ms
25 ms
23 ms
Server Side Fetch Latency (p99)
650 ms
600 ms
550 ms
500 ms
450 ms
400 ms

Logs..

[eCommerce] Revenue Dashbo | [Flights] Global Flight Dashbo | [Flights] Overview - Kibana

demo.elastic.co/app/kibana#/dashboard/722b74f0-b882-11e8-a6d9-e546fe2bba5f?_g=(refreshInterval:(pause:!f,value:900000),time:(from:now-7d,to:now))&_a=(description:'Analyze%20...

Dashboard / [eCommerce] Revenue Dashboard

Full screen Share

Search Lucene Last 7 days Show dates Refresh

+ Add filter

[eCommerce] Orders

1–50 of 1046

Time	category	sku	taxful_total_price	total_quantity
Nov 18, 2019 @ 20:55:12.000	Women's Clothing, Women's Shoes	ZO0707507075, ZO0246402464, ZO0226802268, ZO0343503435	$139.96	4
Nov 18, 2019 @ 20:50:53.000	Men's Clothing	ZO0473704737, ZO0121501215	$45.98	2
Nov 18, 2019 @ 20:35:02.000	Women's Shoes, Women's Clothing	ZO0673606736, ZO0161801618	$88.98	2
Nov 18, 2019 @ 20:04:48.000	Women's Shoes, Women's Accessories	ZO0242702427, ZO0090000900	$70.98	2
Nov 18, 2019 @ 19:59:02.000	Men's Clothing	ZO0589505895, ZO0575405754	$42.98	2
Nov 18, 2019 @ 19:51:50.000	Women's Clothing, Women's Shoes	ZO0490204902, ZO0025000250	$45.98	2
Nov 18, 2019 @ 19:50:24.000	Men's Shoes, Men's Clothing	ZO0400004000, ZO0519305193, ZO0482004820, ZO0540305403	$300.96	4
Nov 18, 2019 @ 19:50:24.000	Men's Clothing	ZO0419604196, ZO0559705597	$39.98	2
Nov 18, 2019 @ 19:24:29.000	Men's Shoes, Men's Accessories	ZO0520305203, ZO0462204622	$66.98	2
Nov 18, 2019 @ 19:17:17.000	Women's Shoes, Women's Clothing	ZO0216502165, ZO0327503275	$78.98	2
Nov 18, 2019 @ 19:14:24.000	Men's Shoes, Men's Clothing	ZO0257002570, ZO0455404554	$85.98	2
Nov 18, 2019 @ 19:08:38.000	Men's Clothing	ZO0547905479, ZO0583305833	$32.98	2
Nov 18, 2019 @ 18:55:41.000	Women's Clothing	ZO0341103411, ZO0648406484	$60.98	2
Nov 18, 2019 @ 18:52:48.000	Women's Clothing	ZO0100901009, ZO0235102351	$53.98	2
Nov 18, 2019 @ 18:51:22.000	Men's Clothing	ZO0575305753, ZO0540605406	$58.98	2
Nov 18, 2019 @ 18:39:50.000	Women's Clothing, Women's Shoes	ZO0266902669, ZO0244202442	$105.98	2
Nov 18, 2019 @ 18:31:12.000	Men's Clothing	ZO0279702797, ZO0573705737	$48.98	2

Traces..

Jaeger UI | Lookup by Trace ID... | Search | Compare | System Architecture | About Jaeger

frontend: HTTP GET /dispatch 169cd7c

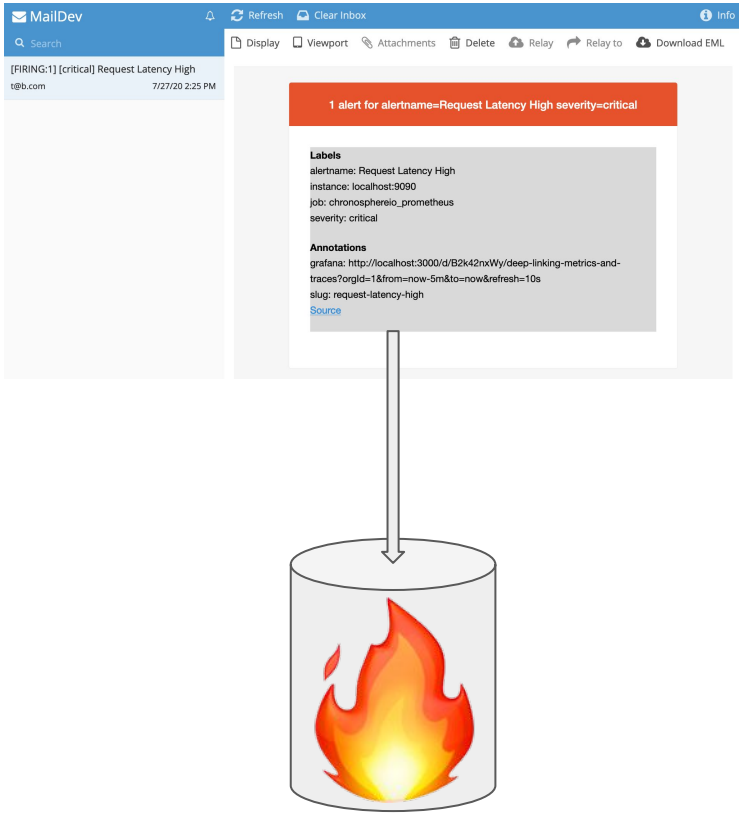
Find... | Alternate Views

Trace Start July 23 2020, 09:11:39.604 | Duration 682.54ms | Services 6 | Depth 5 | Total Spans 51

0ms | 170.64ms | 341.27ms | 511.91ms | 682.54ms

Service & Operation	0ms	170.64ms	341.27ms	511.91ms	682.54ms
frontend HTTP GET /dispatch					
frontend HTTP GET: /customer	265.28ms				
frontend HTTP GET	265.27ms				
customer HTTP GET /customer	264.83ms				
mysql SQL SELECT	264.72ms				
frontend Driver::findNearest	203.21ms				
driver Driver::findNearest	202.49ms				
redis FindDriverIDs	24.28ms				
redis GetDriver	32.2ms				
redis GetDriver	8.83ms | redis::GetDriver				
redis GetDriver	8.56ms				
redis GetDriver	9.47ms				
redis GetDriver	8.71ms				
redis GetDriver	27ms				
redis GetDriver	10.28ms				
redis GetDriver	7.52ms				

Today: What happens when you get alerted?



MySQL Down in region:us-west1
cluster:demo

On-call Experience Today

Received an alert email notification

MailDev

Search

Refresh | Clear Inbox | Info

Display | Viewport | Attachments | Delete | Relay | Relay to | Download EML

[FIRING:1] [critical] Request Latency High

t@b.com 7/27/20 2:25 PM

1 alert for alertname=Request Latency High severity=critical

Labels

alertname: Request Latency High

instance: localhost:9090

job: chronosphereio_prometheus

severity: critical

Annotations

grafana: http://localhost:3000/d/B2k42nxWy/deep-linking-metrics-and-traces?orgId=1&from=now-5m&to=now&refresh=10s

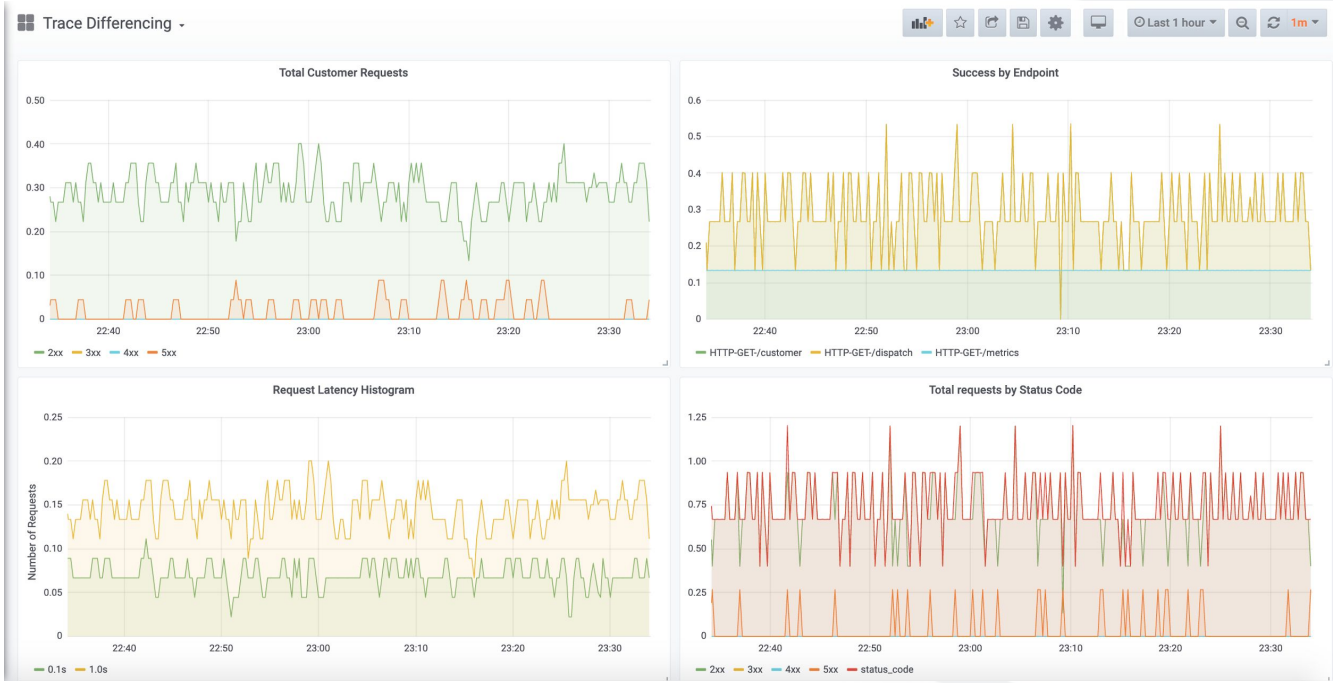
slug: request-latency-high

Source

Link to Grafana

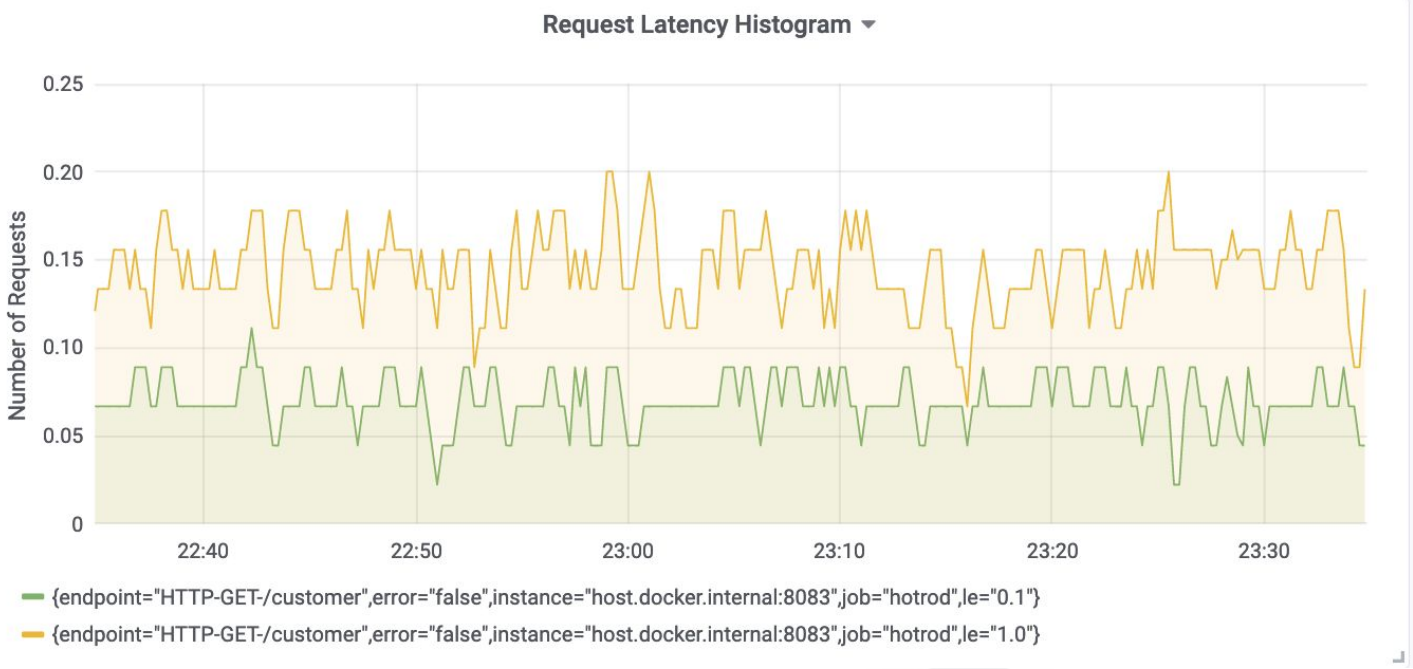
On-call Experience Today

Navigate to the related dashboard



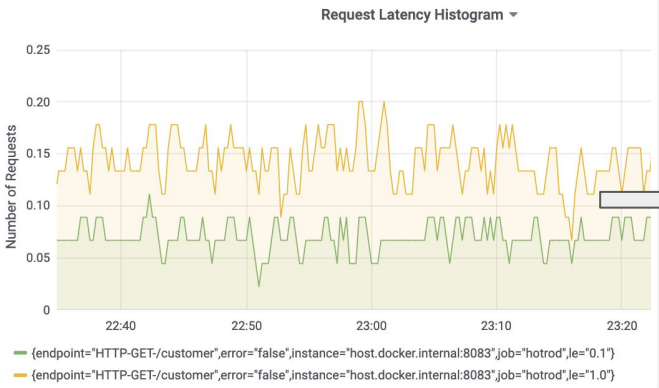
On-call Experience Today

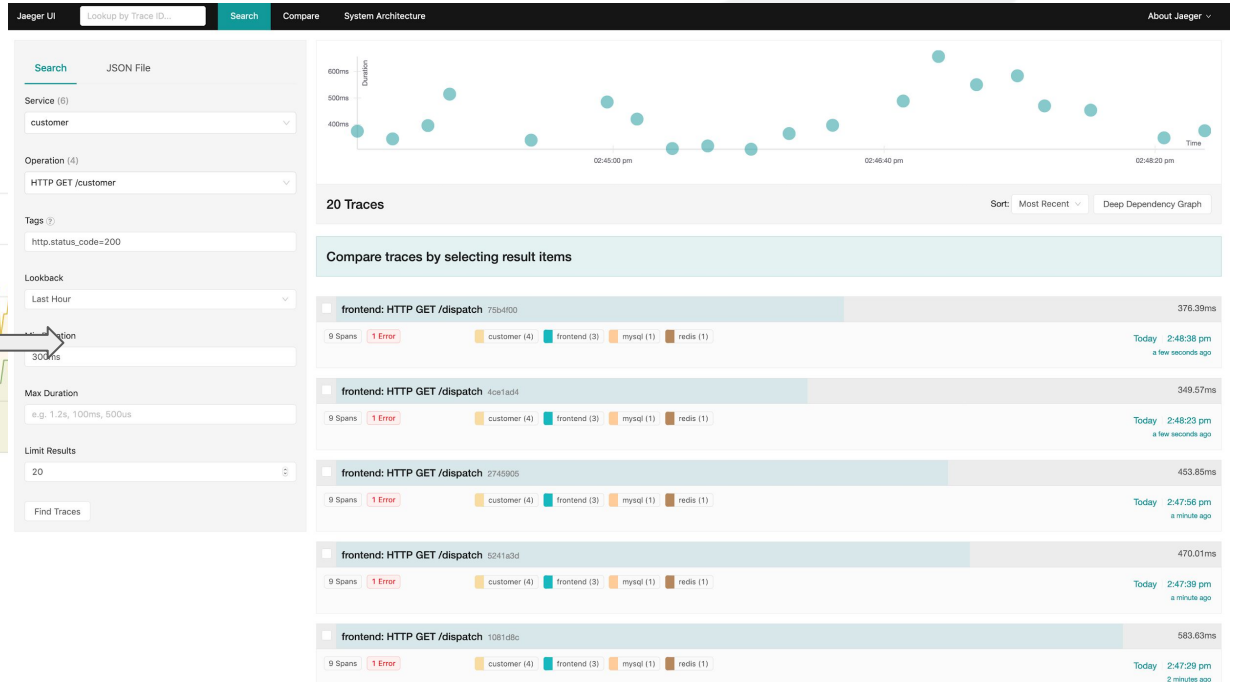
Get the tags on metric to search related traces / logs



On-call Experience Today

Can investigate using trace/logs. Let’s talk about using traces..





On-call Experience Today

Find a trace which is showing higher latency

Jaeger UI | Lookup by Trace ID... | Search | Compare | System Architecture | About Jaeger

frontend: HTTP GET /dispatch 7b0f3f1

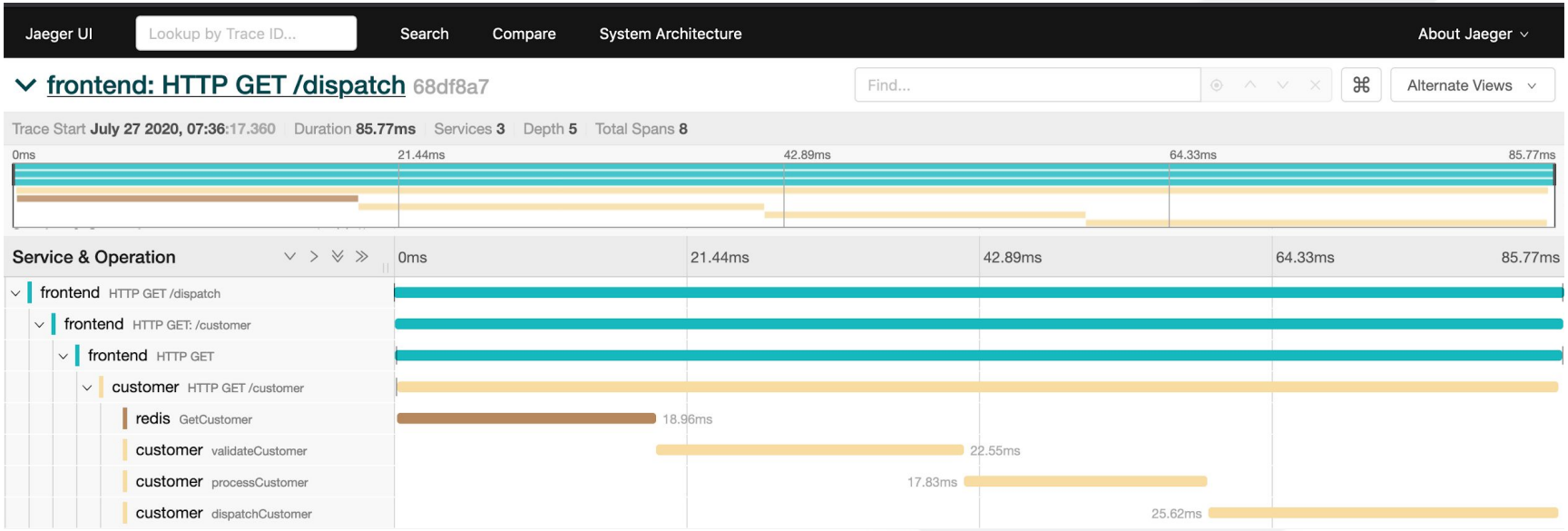
Find... | Alternate Views

Trace Start **July 27 2020, 07:36:14.197** | Duration **522.58ms** | Services **4** | Depth **5** | Total Spans **9**

Service & Operation	0ms	130.65ms	261.29ms	391.94ms	522.58ms
frontend HTTP GET /dispatch					
frontend HTTP GET: /customer					
frontend HTTP GET					
customer HTTP GET /customer					
redis GetCustomer	45.38ms				
mysql SQL SELECT				408.47ms	
customer validateCustomer				23.6ms	
customer processCustomer				23.84ms	
customer dispatchCustomer				20.28ms	

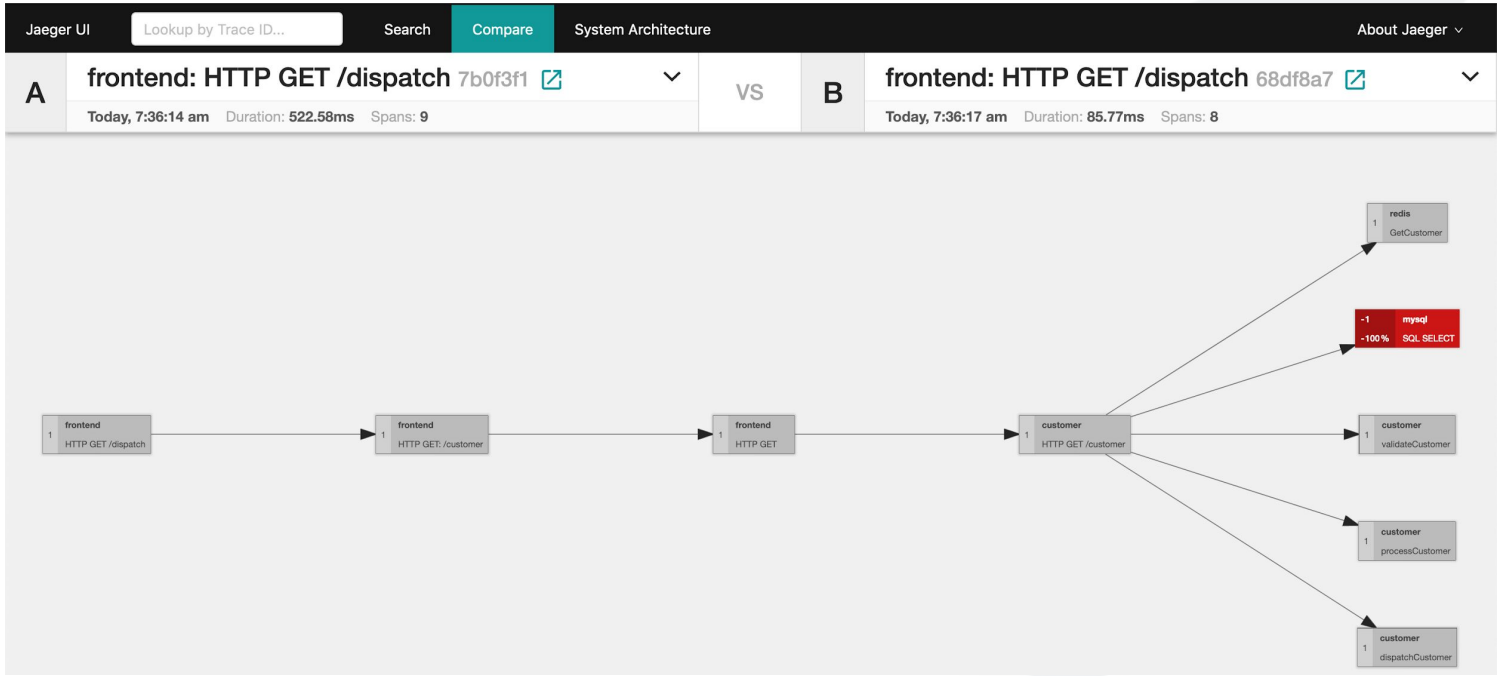
On-call Experience Today

Find a trace which is showing lower latency

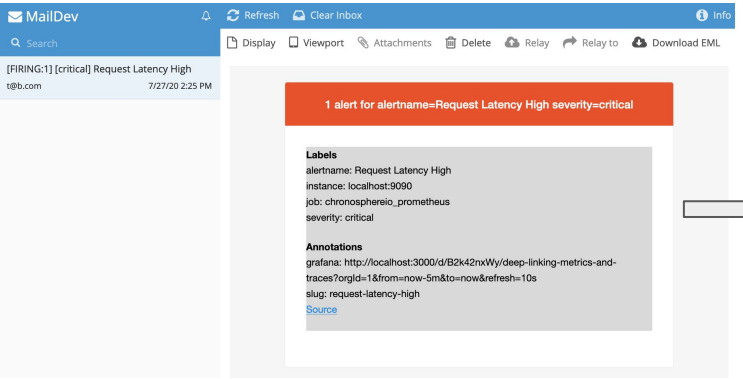


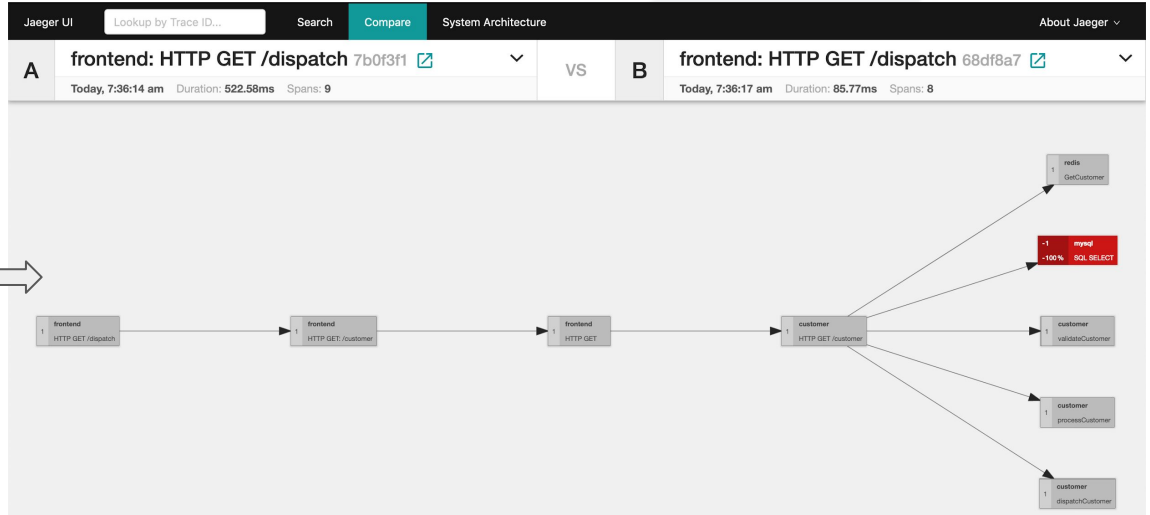
On-call Experience Today

Trace differences are a powerful tool, MySQL is being slow!

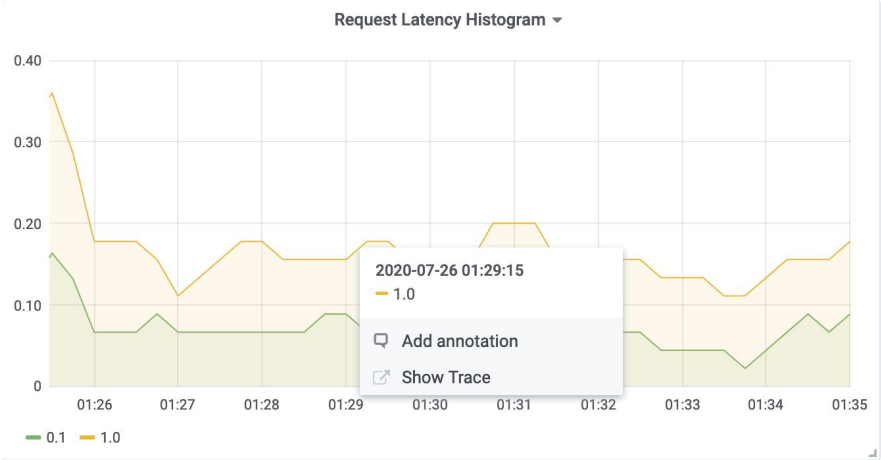


Can we jump there automatically?





The Journey: Deep Linking Metrics and Traces

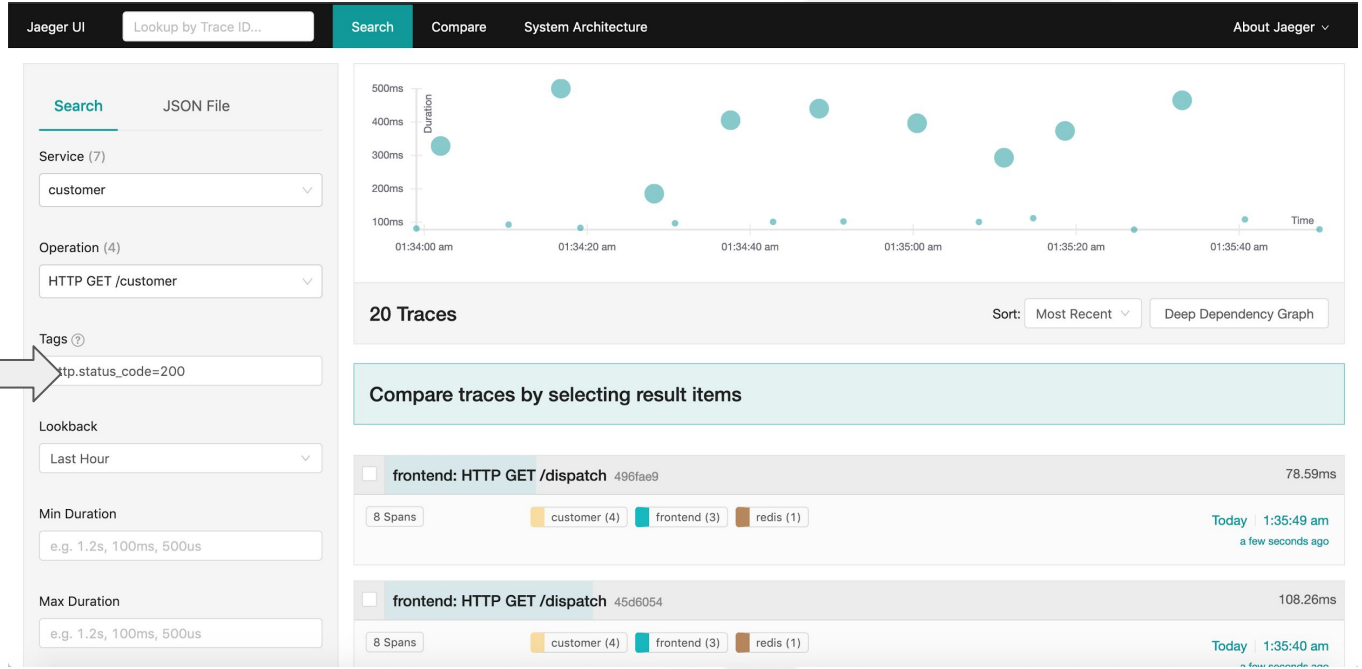


Tracing and Metrics

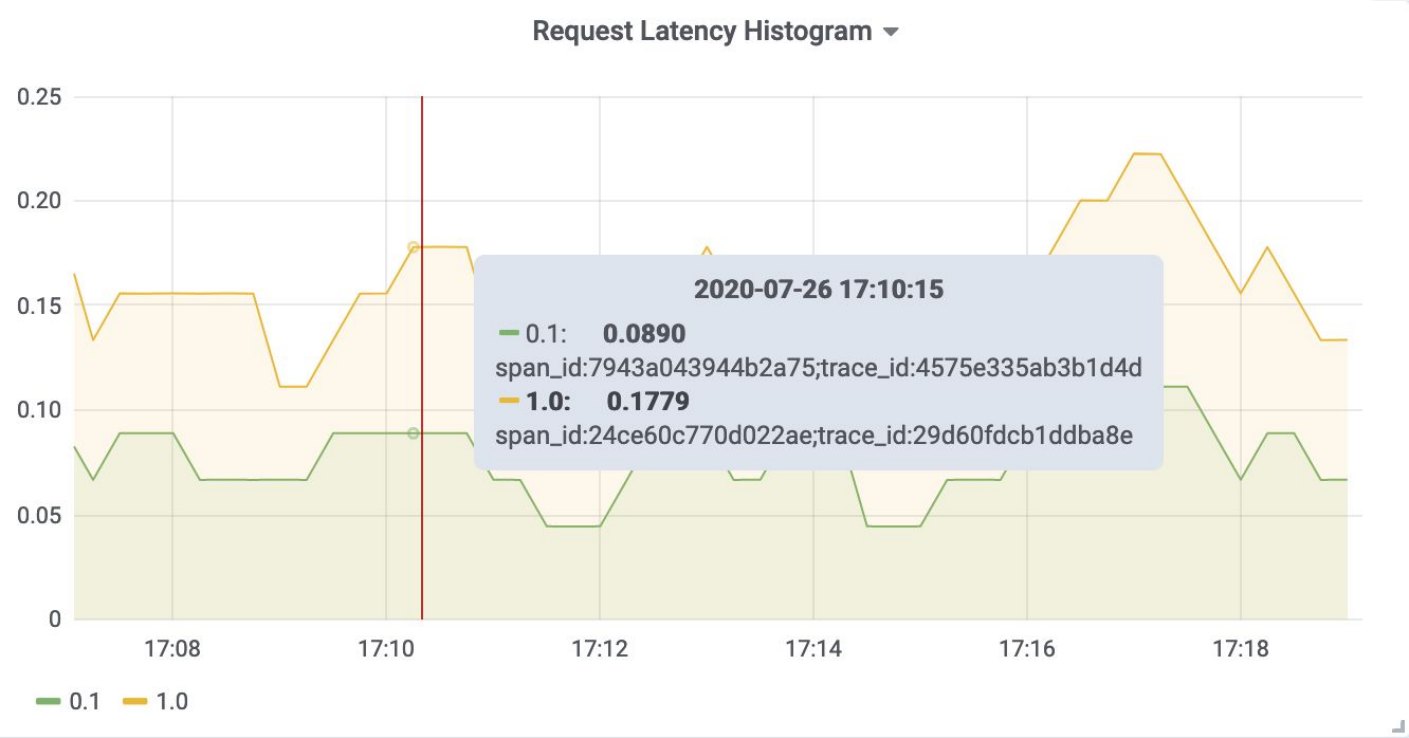
Generally linked by common or similar tags.

Tracing and Metrics

Generally linked by common or similar tags.



We can actually jump to the trace directly...



Open Metrics and Exemplars

- Open Metrics allows augmenting context information

```
# HELP http_requests_total http_requests
# TYPE http_requests_total counter
http_requests_total{endpoint="/search",status_code="2xx"} 1725 # {trace_id="b096e71d..."} 1
http_requests_total{endpoint="/search",status_code="4xx"} 4 # {trace_id="944a6d97..."} 1
http_requests_total{endpoint="/search",status_code="5xx"} 27 # {trace_id="50785260..."} 1
http_request_latency_bucket{endpoint="/search",le="0.1"} 7 # {trace_id="7f78deda..."} 1
http_request_latency_bucket{endpoint="/search",le="0.2"} 7 # {trace_id="5ad53ac9..."} 1
http_request_latency_bucket{endpoint="/search",le="0.3"} 7 # {trace_id="c78493ec..."} 1
```

OpenTelemetry: Instrumentation SDK

- OpenTelemetry provides a single set of APIs to emit metrics and traces
- Metrics can now be emitted with tracing context, with an ability to choose which metrics actually get that context
- Use OpenMetrics format support to ensure trace ID information is sent to the metrics datastore

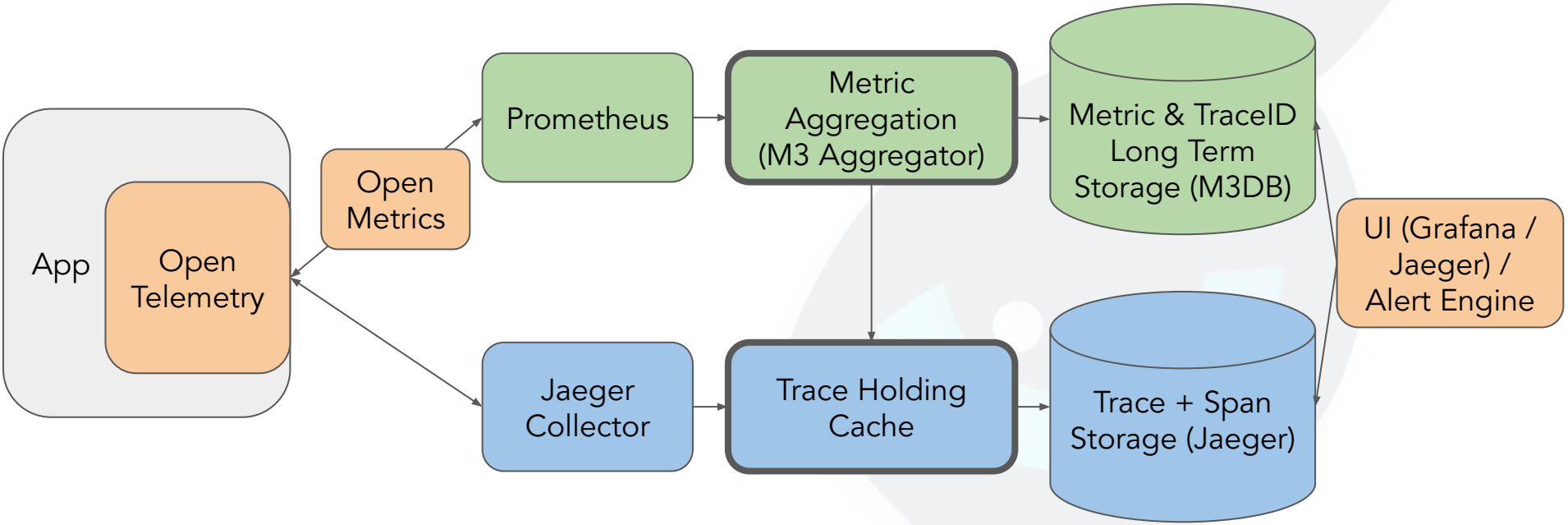
Prometheus / M3

- Prometheus support to scrape metrics with exemplars
- M3 has the ability to store exemplars alongside metric datapoints
  - Durable and stored for lifetime of datapoint
- M3 query support to return exemplars alongside datapoints
- M3 query ensures at least one representative exemplar is present even after applying aggregation functions like sum(...), max(...)

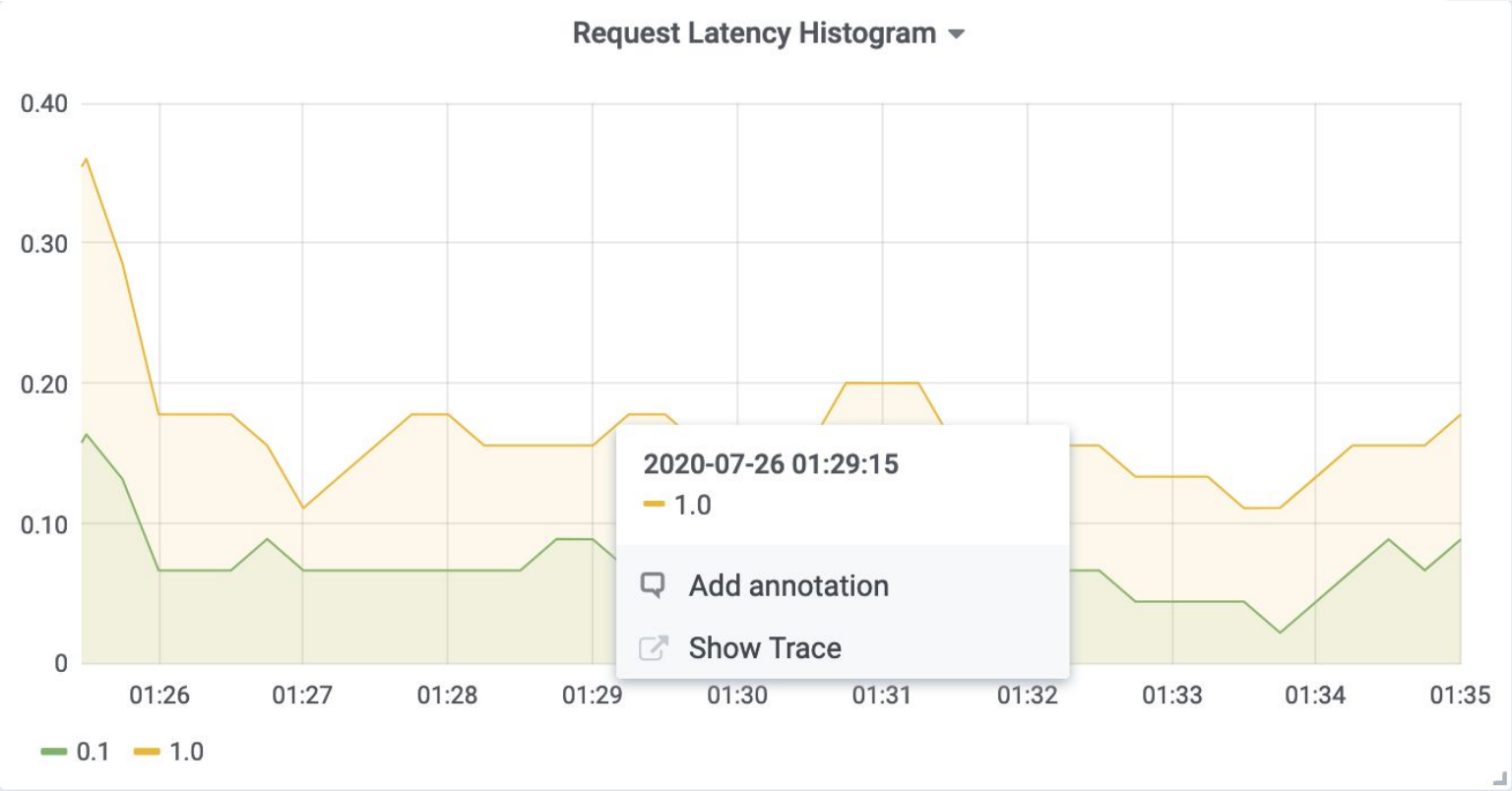
Trace Sampling

- Traditional trace sampling techniques insufficient
- We need to store the specific traces that were emitted as exemplars with the metrics
- A trace holding tier can hold all traces for a short duration, with the M3 aggregation layer indicating which traces to actually persist

A Complete Ingestion Pipeline



What That Enables..



Tomorrow: Getting from an Alert to a Request Comparison

Demo

This is what the on-call experience can look like..

How?

- Leverage OpenTelemetry and OpenMetrics to emit metrics with trace IDs as exemplars
- Leverage Prometheus and M3's support to scrape and store exemplars alongside metric datapoints
- Building contextual links into the systems consuming trace and metric information

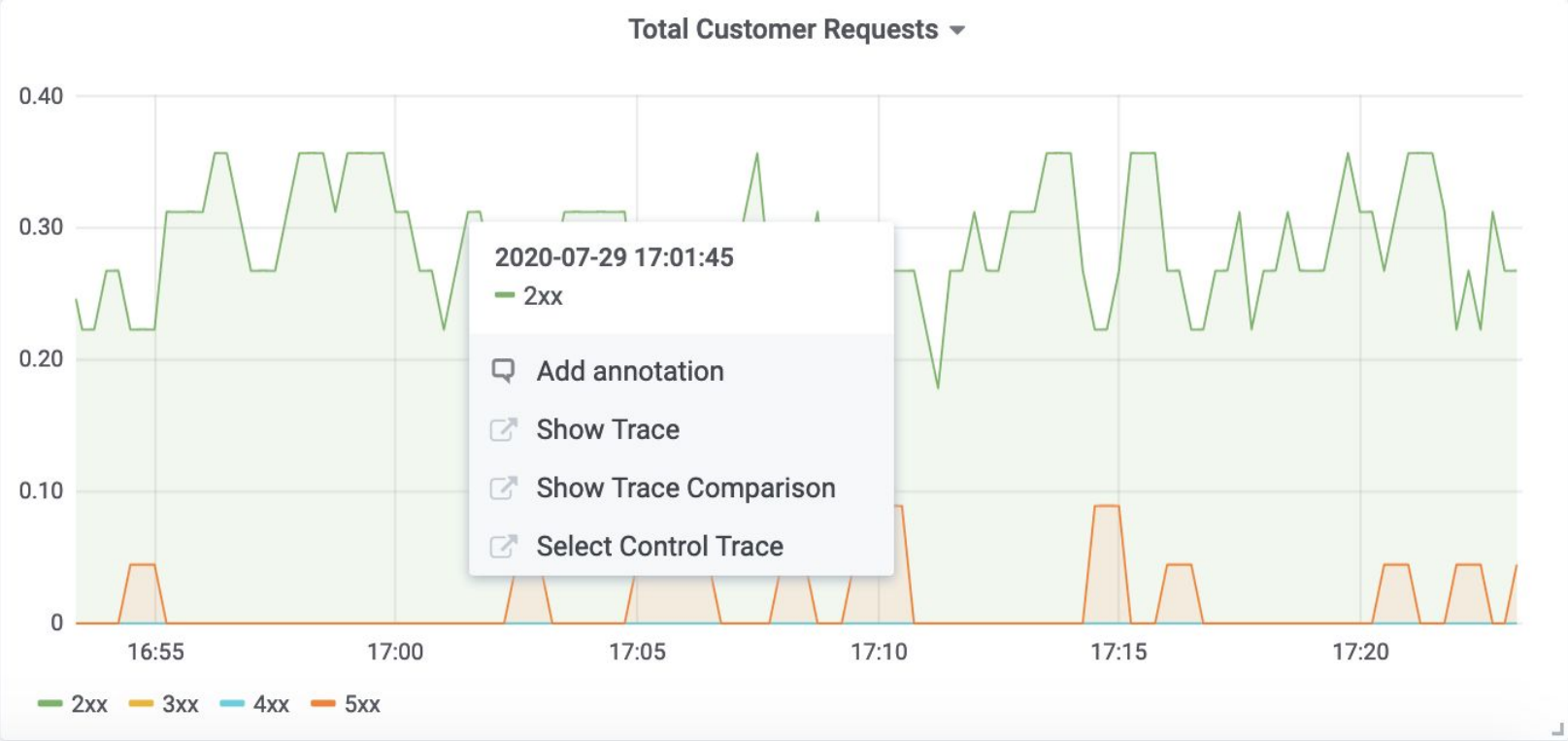
M3 Query and Exemplars

M3 Query response augments exemplar alongside metric value.

Ensures a representative exemplar on aggregation functions.


```
{
  "metric": {
    "endpoint": "HTTP-GET-/customer",
    "error": "false",
    "instance": "host.docker.internal:8083",
    "job": "hotrod",
    "le": "0.1"
  },
  "values": [
    [
      1595949555,
      "0.041792222222222224",
      "span_id:516f48571f0f0082;trace_id:7510b68f10714f10"
    ],
```

Selecting a good/bad source for traces?



Building Contextual Links

- For graphing integrations, can configure a metric that can act as source of good exemplars
- For alerting integrations, provide ability to configure a metric that can act as a source of good exemplars
- For standard well named metrics, like RPC metrics, can build plugins that can automatically detect and provide comparisons based on knowledge of metrics emitted

Summary

- Trace differences can be a powerful tool to debug issues
- Using deep linking support between metric datapoints and traces we can build integrations that can speed up root cause

Where are we on this journey?

Current end-to-end demo at:

https://github.com/chronosphereio/demo-trace-differencing

Merged: Add exemplar support to OpenMetrics:
https://github.com/prometheus/prometheus/pull/6292

Merged: Add exemplar support in Prometheus Client (@beorn7):
https://github.com/prometheus/client_golang/pull/707

Open(needs discussion): Store exemplars in Prometheus memory, forward on remote write:
https://github.com/prometheus/prometheus/pull/6309

Resources

Talk Deep Linking Metrics and Traces with OpenTelemetry, OpenMetrics and M3.

Rob Skillington, Kubecon San Diego, 2019 [Video]

OpenMetrics https://github.com/OpenObservability/OpenMetrics

OpenTelemetry https://github.com/open-telemetry/opentelemetry-specification

Prometheus https://github.com/prometheus/prometheus

M3 https://github.com/m3db/m3

Grafana https://github.com/grafana/grafana

Thank you and Q&A

Come say hi at our virtual booth!

http://bit.ly/m3slack