



RPC Metrics at Google

JBD, Google (@rakyll)



gRPC Metrics at Google

JBD, Google (@rakyll)



Request Metrics at Google

JBD, Google (@rakyll)

"100% is the wrong
reliability target for
basically everything."

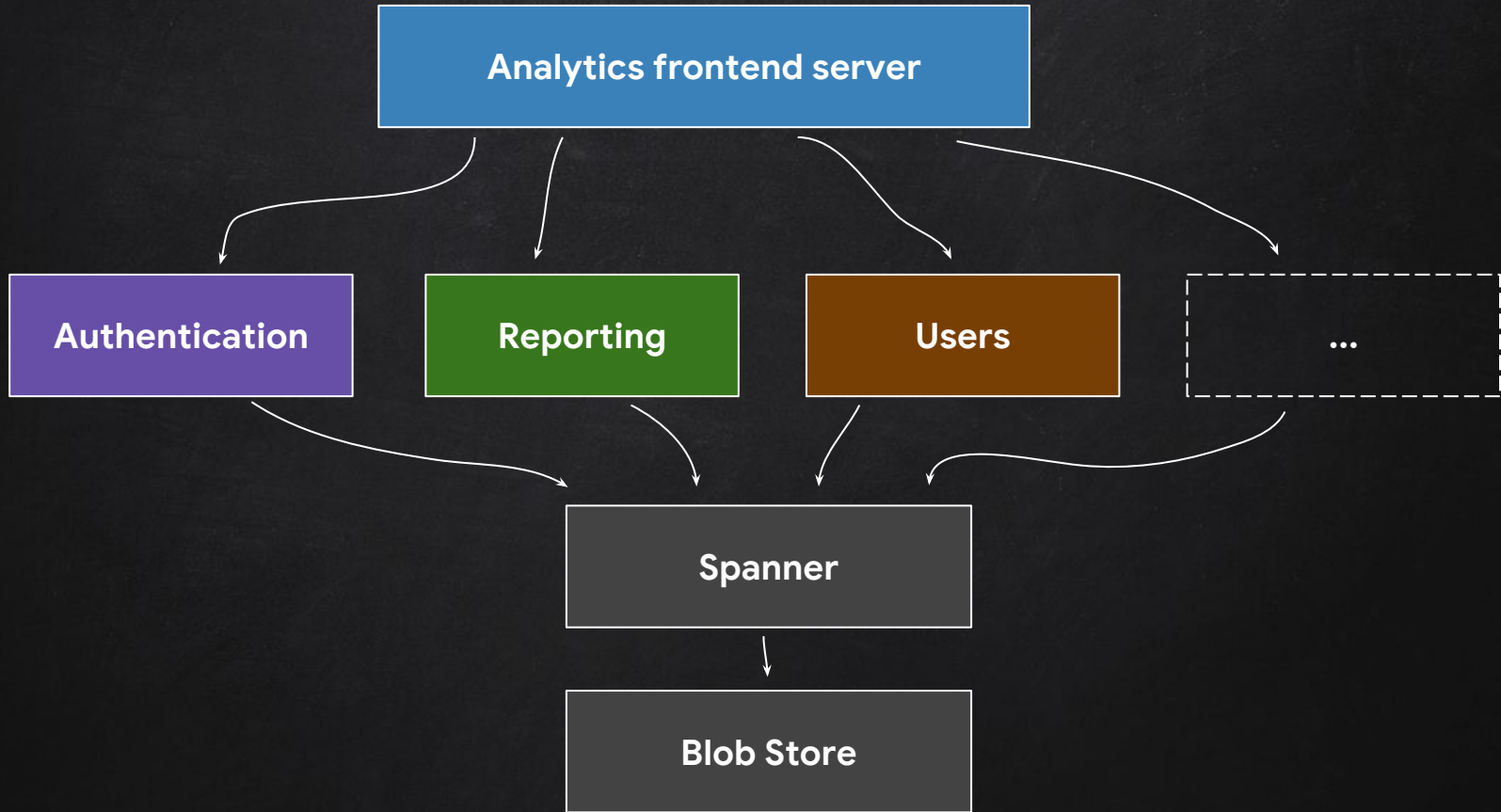
-- Benjamin Treynor Sloss, VP of Engineering, Google

"A service is available
if users cannot tell that
there was an outage."

SLOs

Principled way of saying what level of downtime is acceptable.

- Error rate
- Latency expectations



Questions infra teams want to ask:

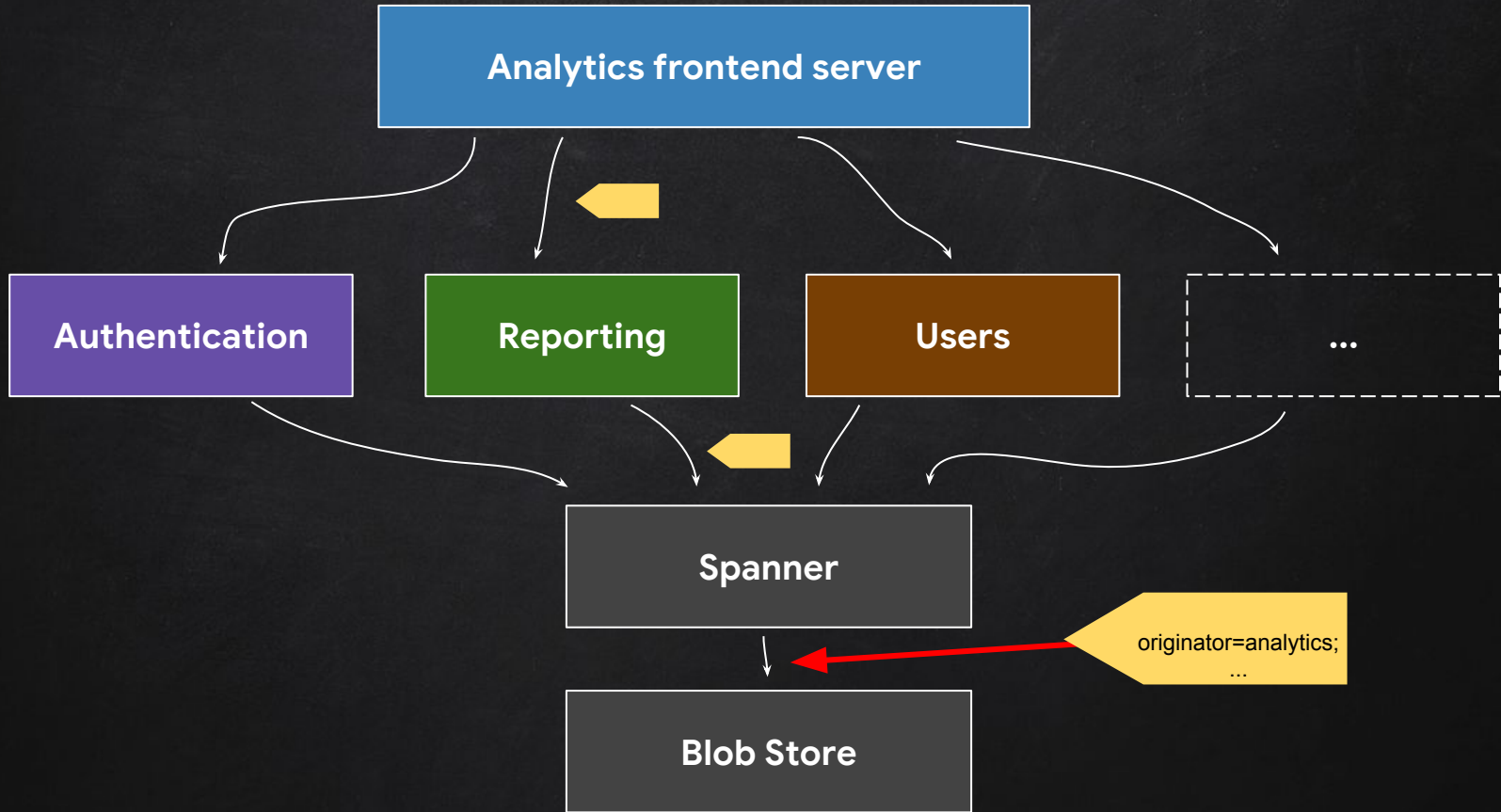
- Are we meeting the SLO for the other team?
- What's the impact of a product on infra?
- How much do we need to scale up if product grows 10%?

High-Cardinality

Breaking down the metrics data...

Query the collected data in various ways:

- Latency distribution for RPCs originated at Google Analytics.
- Requests take took more than 100ms for the customer #123.
- Compare the request latency initiated at web vs mobile frontend.

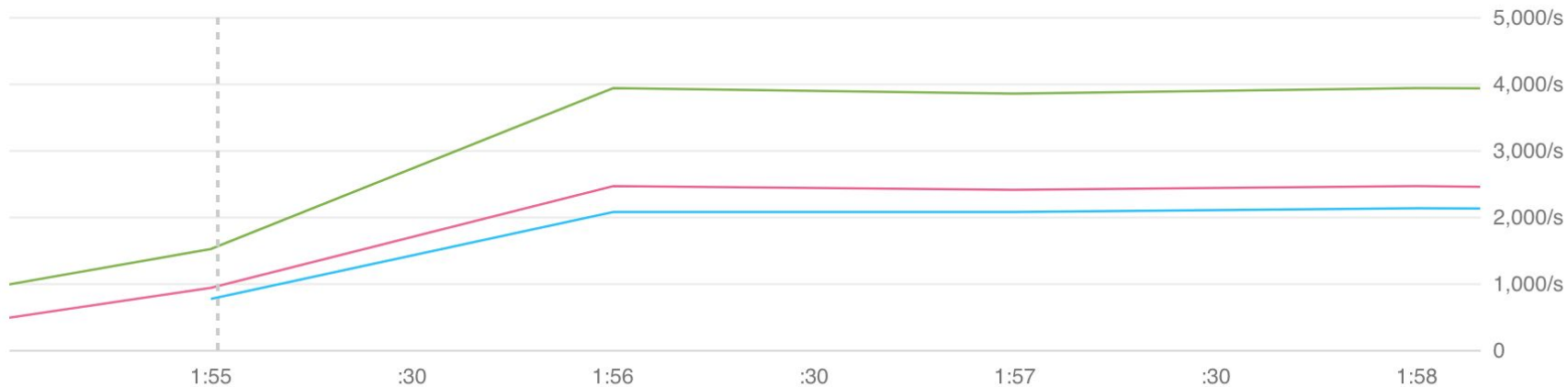


Blob store read errors by originator



by originator (sum)

1 min interval (rate)

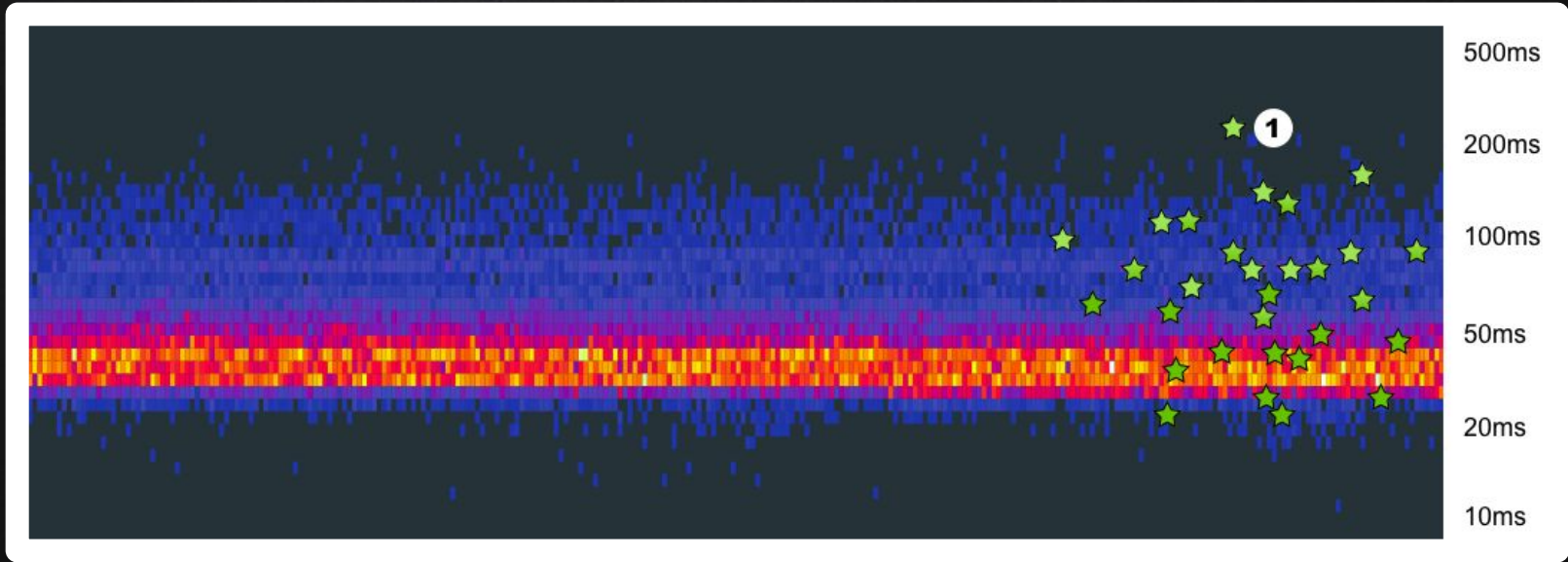


Name	Value
● analytics	941.99/s
● gmail	772.99/s
● inbox	0
● spanner	1,510.23/s

Dynamically choose aggregation

(split between recording and aggregation)

Exemplars



/rpz and /statz

Export?

Monarch, Prometheus, and more.

import “cloud.google.com/go/pubsub”

```
# HELP cloud_google_com_go_pubsub_pull_count Number of PubSub
messages pulled
# TYPE cloud_google_com_go_pubsub_pull_count counter
cloud_google_com_go_pubsub_pull_count{subscription="projects/
jbdlabs/subscriptions/tasks-sub"} 108

# HELP cloud_google_com_go_pubsub_stream_open_count Number of
calls opening a new streaming pull
# TYPE cloud_google_com_go_pubsub_stream_open_count counter
cloud_google_com_go_pubsub_stream_open_count{subscription="pr
ojects/jbdlabs/subscriptions/tasks-sub"} 1
```

Google

+



OPENMETRICS

Thank you!

JBD, Google
jbd@google.com
@rakyll