Grafana Labs

# Loki: Cloud Native Logging

Grafana Labs

# What’s in it for you?

- Learn a different approach to logs, at scale
- Learn strategies for building cloud native systems

Grafana Labs

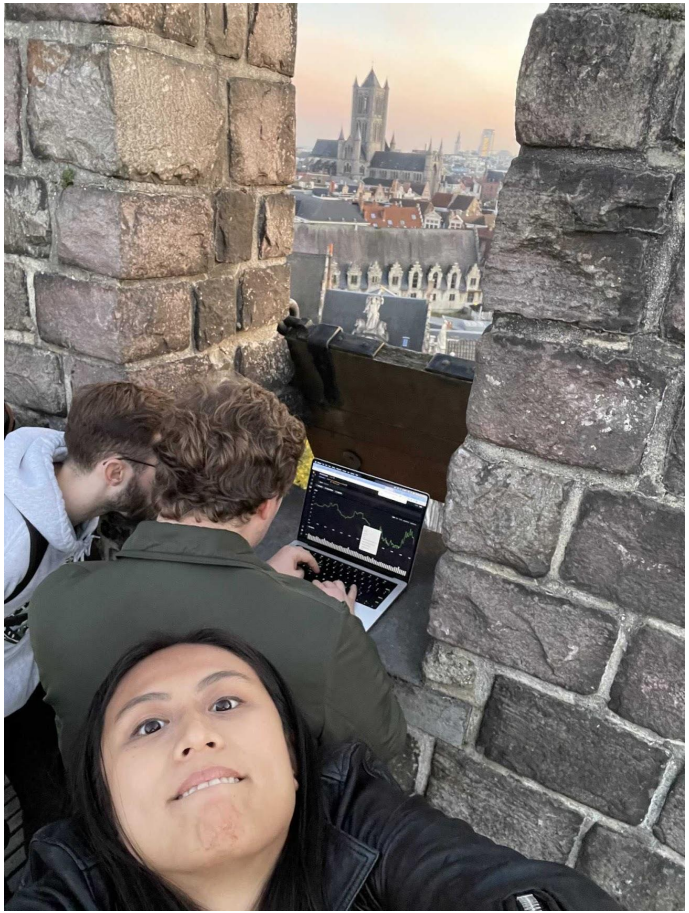
# Who we are?

- Engineers at Grafana Labs
- Grafana Loki Open Source maintainers

**Owen Diehl**
Eng

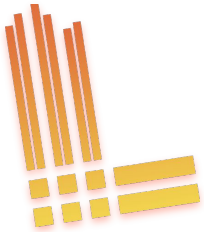

**Kaviraj Kanagaraj**
Eng

https://github.com/grafana/loki



# What is Loki?

Loki is a time series database, but for strings

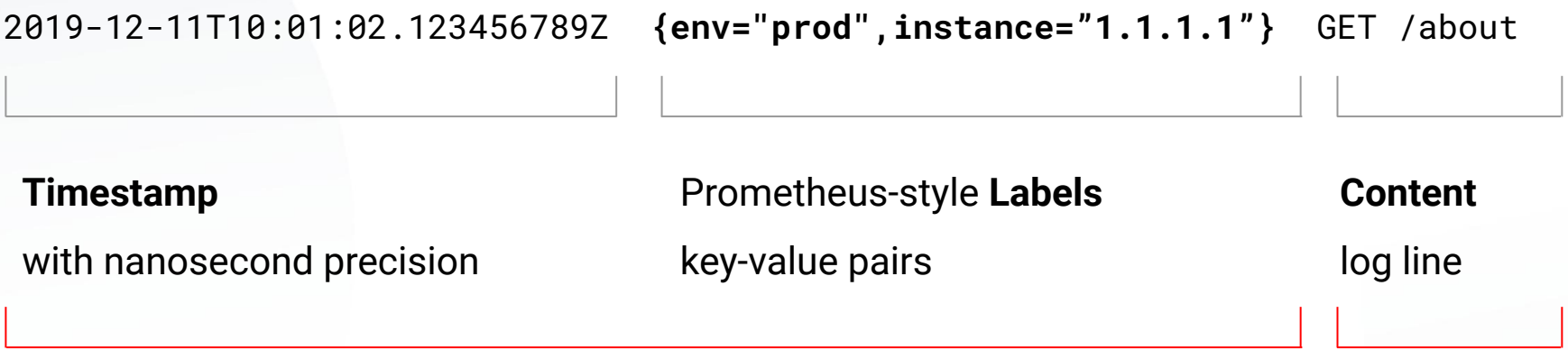


# What is Timeseries database?

```
identifier -> (t0, v0), (t1, v1), (t2, v2), (t3, v3), ....
```

Prometheus: `{app="nginx", cluster="us-central-0"} -> [(1653994269, 34.5)]`

Loki: `{app="nginx", cluster="us-central-0"} -> [(1653994269, "/ GET")]`


2019-12-11T10:01:02.123456789Z
{env="prod",instance="1.1.1.1"}
GET /about
Timestamp
with nanosecond precision
Prometheus-style Labels
key-value pairs
Content
log line
indexed
unindexed

All the logs in cluster us-central

```
{cluster="us-central"}
```

1TB of data in 24 hour period

# Only the Loki ingester logs

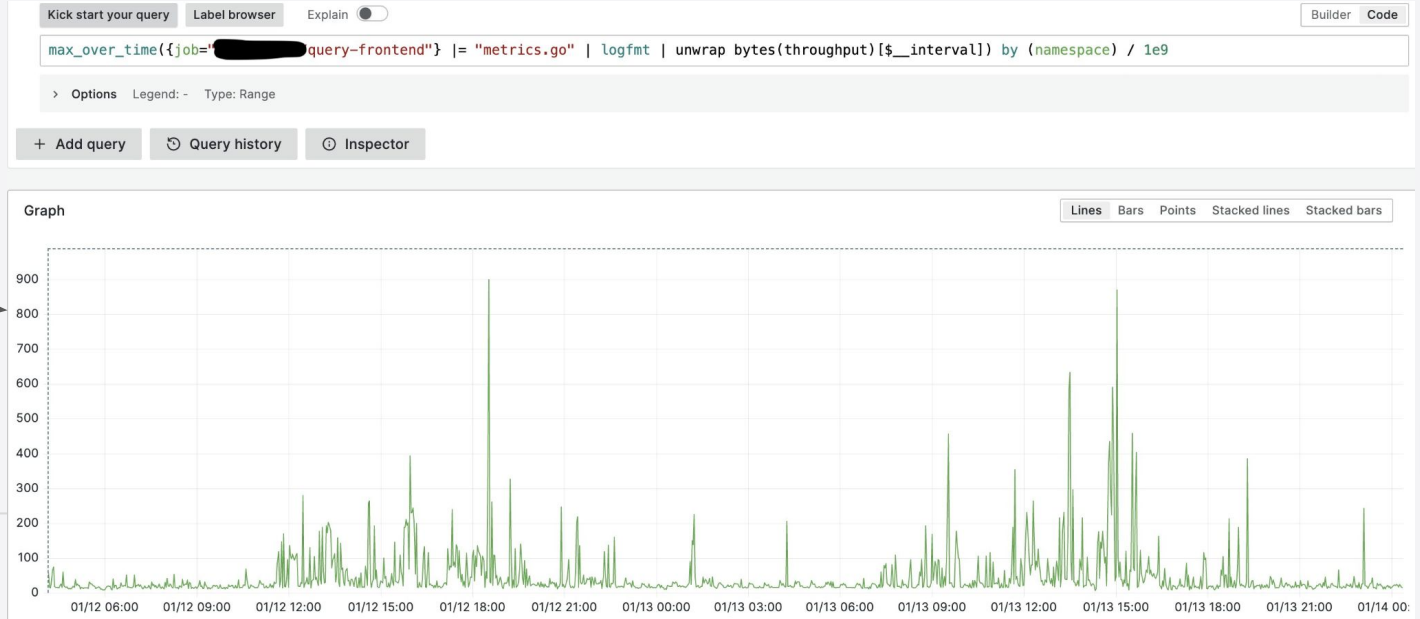
```
{container="ingester",namespace=~"loki.*"}
```

10GB of data in 24 hour period

# Loki @ Grafana Labs

- Largest user cluster (as of 2023-01): 50 TiB per day
- Queries peak at **900GB/s**
  - Query 10TB in 12 seconds, including complex processing of result sets

*These metrics are built with Loki using logs!*

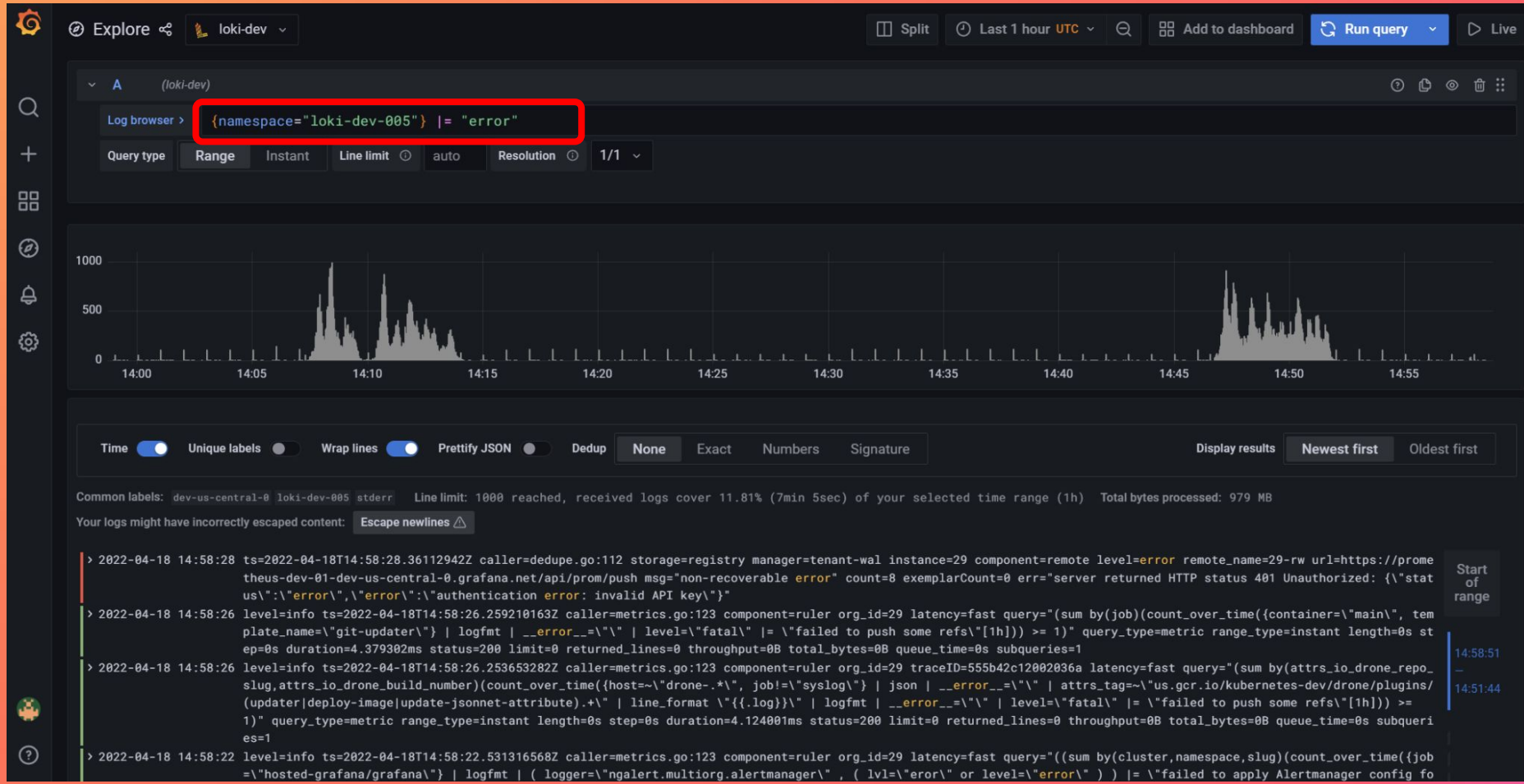


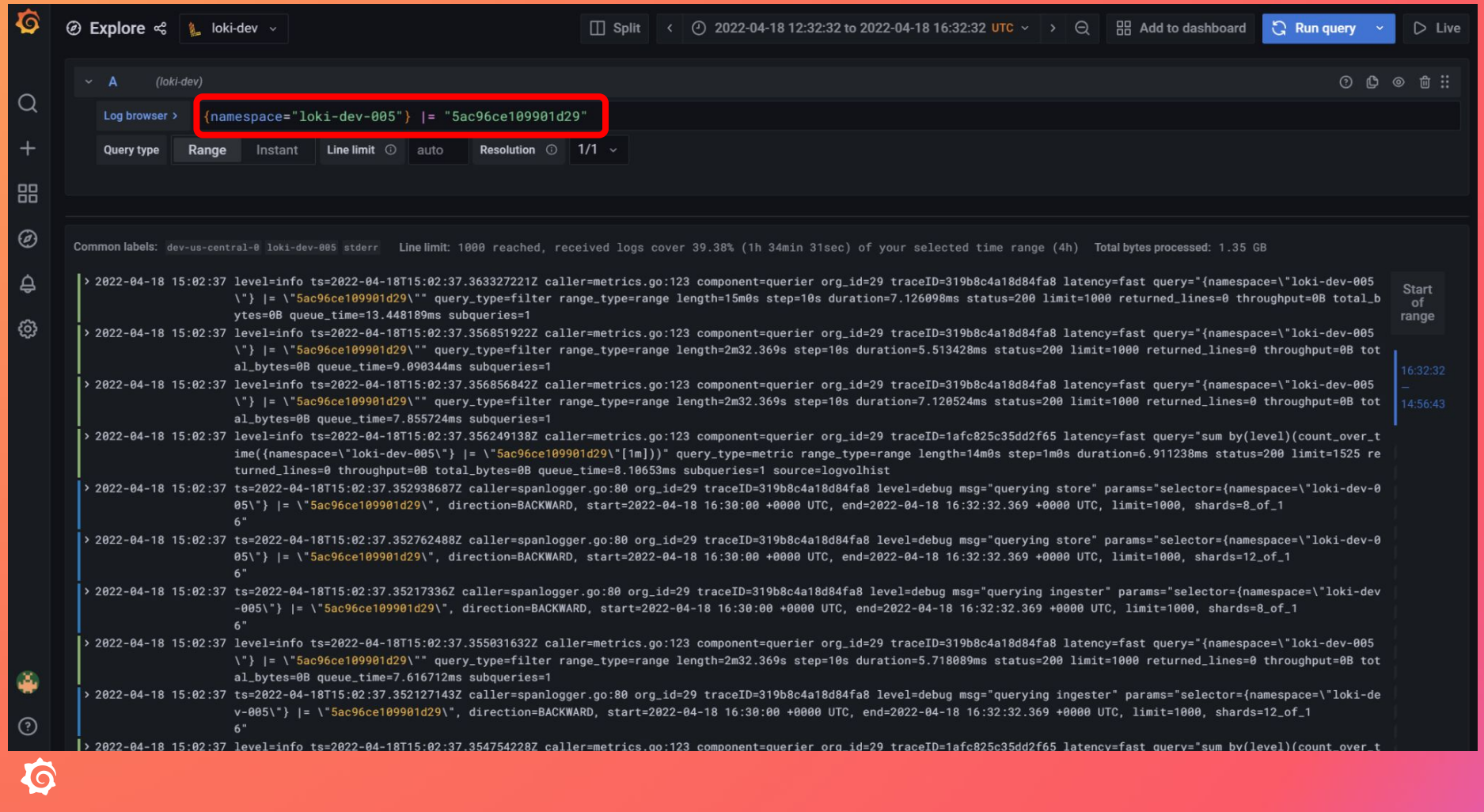
# Log aggregation: the problem

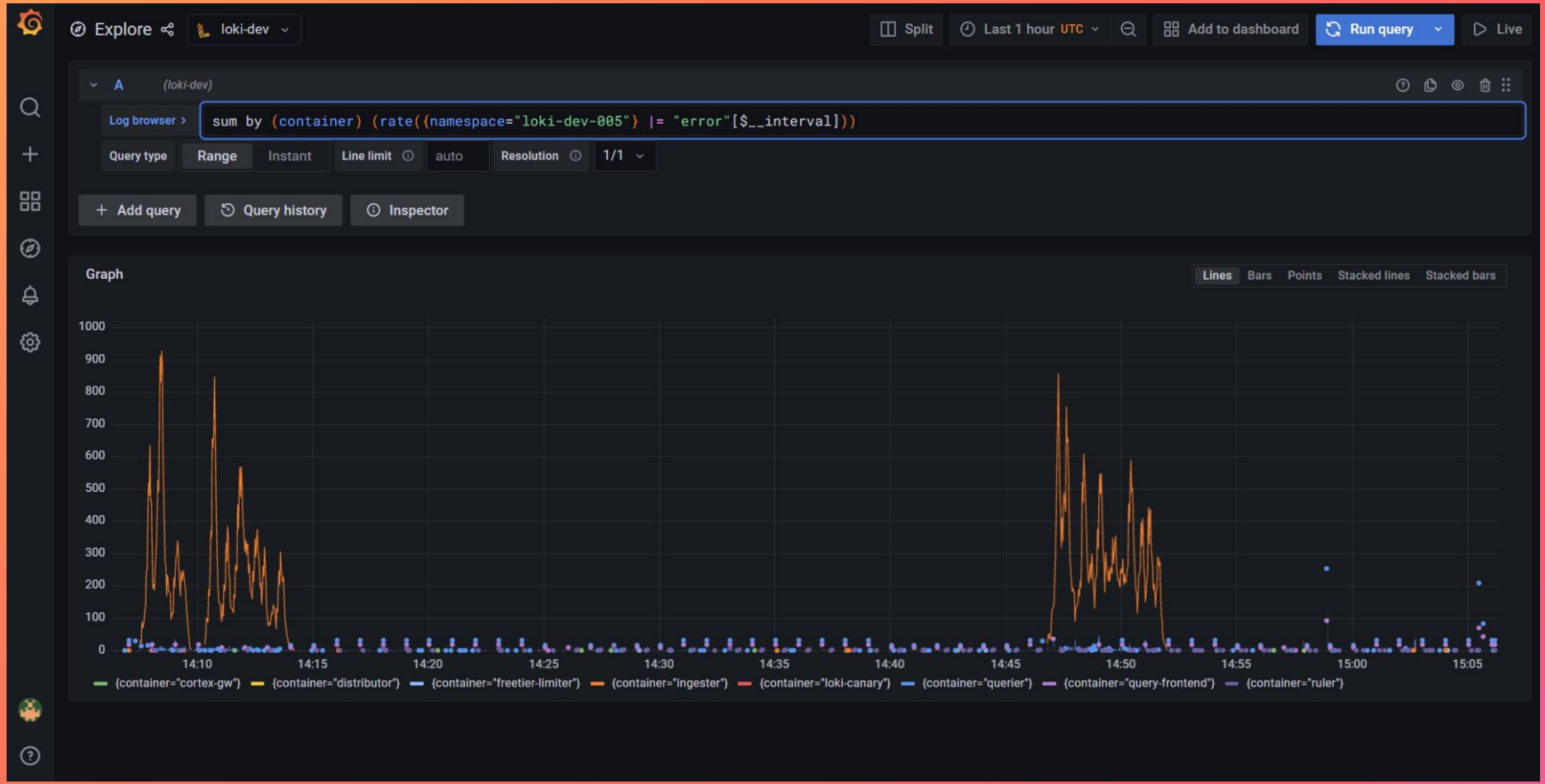
- Starts with `tail | grep`
- Doesn’t scale so well today

# Log aggregation: the solution

# Distributed Grep


Explore
loki-dev
Split
Last 1 hour UTC
Add to dashboard
Run query
Live
A
(loki-dev)
Log browser >
{namespace="loki-dev-005"} |= "error"
Query type
Range
Instant
Line limit
auto
Resolution
1/1
1000
500
0
14:00
14:05
14:10
14:15
14:20
14:25
14:30
14:35
14:40
14:45
14:50
14:55
Time
Unique labels
Wrap lines
Prettify JSON
Dedup
None
Exact
Numbers
Signature
Display results
Newest first
Oldest first
Common labels: dev-us-central-0 loki-dev-005 stderr
Line limit: 1000 reached, received logs cover 11.81% (7min 5sec) of your selected time range (1h)
Total bytes processed: 979 MB
Your logs might have incorrectly escaped content:
Escape newlines
> 2022-04-18 14:58:28 ts=2022-04-18T14:58:28.36112942Z caller=dedupe.go:112 storage=registry manager=tenant-wal instance=29 component=remote level=error remote_name=29-rw url=https://prometheus-dev-01-dev-us-central-0.grafana.net/api/prom/push msg="non-recoverable error" count=8 exemplarCount=0 err="server returned HTTP status 401 Unauthorized: {\"status\":\"error\",\"error\":\"authentication error: invalid API key\"}"
> 2022-04-18 14:58:26 level=info ts=2022-04-18T14:58:26.259210163Z caller=metrics.go:123 component=ruler org_id=29 latency=fast query="(sum by(job)(count_over_time({container=\"main\", template_name=\"git-updater\"} | logfmt | __error__=\"\" | level=\"fatal\" |= \"failed to push some refs\"[1h])) >= 1)" query_type=metric range_type=instant length=0s step=0s duration=4.379302ms status=200 limit=0 returned_lines=0 throughput=0B total_bytes=0B queue_time=0s subqueries=1
> 2022-04-18 14:58:26 level=info ts=2022-04-18T14:58:26.253653282Z caller=metrics.go:123 component=ruler org_id=29 traceID=555b42c12002036a latency=fast query="(sum by(attrs_io_drone_repo_slug,attrs_io_drone_build_number)(count_over_time({host=~\"drone-.*\", job!=\"syslog\"} | json | __error__=\"\" | attrs_tag=~\"us.gcr.io/kubernetes-dev/drone/plugins/(updater|deploy-image|update-jsonnet-attribute).+\" | line_format \"{{.log}}\" | logfmt | __error__=\"\" | level=\"fatal\" |= \"failed to push some refs\"[1h])) >= 1)" query_type=metric range_type=instant length=0s step=0s duration=4.124001ms status=200 limit=0 returned_lines=0 throughput=0B total_bytes=0B queue_time=0s subqueries=1
> 2022-04-18 14:58:22 level=info ts=2022-04-18T14:58:22.531316568Z caller=metrics.go:123 component=ruler org_id=29 latency=fast query="((sum by(cluster,namespace,slug)(count_over_time({job=\"hosted-grafana/grafana\"} | logfmt | ( logger=\"ngalert.multiorg.alertmanager\" , ( lvl=\"eror\" or level=\"error\" ) ) |= \"failed to apply Alertmanager config fo
Start of range
14:58:51
–
14:51:44


Explore
loki-dev
Split
2022-04-18 12:32:32 to 2022-04-18 16:32:32 UTC
Add to dashboard
Run query
Live
A
(loki-dev)
Log browser
{namespace="loki-dev-005"} |= "5ac96ce109901d29"
Query type
Range
Instant
Line limit
auto
Resolution
1/1
Common labels: dev-us-central-0 loki-dev-005 stderr
Line limit: 1000 reached, received logs cover 39.38% (1h 34min 31sec) of your selected time range (4h)
Total bytes processed: 1.35 GB
Start of range
16:32:32
14:56:43
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.363327221Z caller=metrics.go:123 component=querier org_id=29 traceID=319b8c4a18d84fa8 latency=fast query="{namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\"" query_type=filter range_type=range length=15m0s step=10s duration=7.126098ms status=200 limit=1000 returned_lines=0 throughput=0B total_bytes=0B queue_time=13.448189ms subqueries=1
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.356851922Z caller=metrics.go:123 component=querier org_id=29 traceID=319b8c4a18d84fa8 latency=fast query="{namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\"" query_type=filter range_type=range length=2m32.369s step=10s duration=5.513428ms status=200 limit=1000 returned_lines=0 throughput=0B total_bytes=0B queue_time=9.090344ms subqueries=1
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.356856842Z caller=metrics.go:123 component=querier org_id=29 traceID=319b8c4a18d84fa8 latency=fast query="{namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\"" query_type=filter range_type=range length=2m32.369s step=10s duration=7.120524ms status=200 limit=1000 returned_lines=0 throughput=0B total_bytes=0B queue_time=7.855724ms subqueries=1
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.356249138Z caller=metrics.go:123 component=querier org_id=29 traceID=1afc825c35dd2f65 latency=fast query="sum by(level)(count_over_time({namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\"[1m]))" query_type=metric range_type=range length=14m0s step=1m0s duration=6.911238ms status=200 limit=1525 returned_lines=0 throughput=0B total_bytes=0B queue_time=8.10653ms subqueries=1 source=logvolhist
2022-04-18 15:02:37 ts=2022-04-18T15:02:37.352938687Z caller=spanlogger.go:80 org_id=29 traceID=319b8c4a18d84fa8 level=debug msg="querying store" params="selector={namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\", direction=BACKWARD, start=2022-04-18 16:30:00 +0000 UTC, end=2022-04-18 16:32:32.369 +0000 UTC, limit=1000, shards=8_of_16"
2022-04-18 15:02:37 ts=2022-04-18T15:02:37.352762488Z caller=spanlogger.go:80 org_id=29 traceID=319b8c4a18d84fa8 level=debug msg="querying store" params="selector={namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\", direction=BACKWARD, start=2022-04-18 16:30:00 +0000 UTC, end=2022-04-18 16:32:32.369 +0000 UTC, limit=1000, shards=12_of_16"
2022-04-18 15:02:37 ts=2022-04-18T15:02:37.35217336Z caller=spanlogger.go:80 org_id=29 traceID=319b8c4a18d84fa8 level=debug msg="querying ingester" params="selector={namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\", direction=BACKWARD, start=2022-04-18 16:30:00 +0000 UTC, end=2022-04-18 16:32:32.369 +0000 UTC, limit=1000, shards=8_of_16"
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.355031632Z caller=metrics.go:123 component=querier org_id=29 traceID=319b8c4a18d84fa8 latency=fast query="{namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\"" query_type=filter range_type=range length=2m32.369s step=10s duration=5.718089ms status=200 limit=1000 returned_lines=0 throughput=0B total_bytes=0B queue_time=7.616712ms subqueries=1
2022-04-18 15:02:37 ts=2022-04-18T15:02:37.352127143Z caller=spanlogger.go:80 org_id=29 traceID=319b8c4a18d84fa8 level=debug msg="querying ingester" params="selector={namespace=\"loki-dev-005\"} |= \"5ac96ce109901d29\", direction=BACKWARD, start=2022-04-18 16:30:00 +0000 UTC, end=2022-04-18 16:32:32.369 +0000 UTC, limit=1000, shards=12_of_16"
2022-04-18 15:02:37 level=info ts=2022-04-18T15:02:37.354754228Z caller=metrics.go:123 component=querier org_id=29 traceID=1afc825c35dd2f65 latency=fast query="sum by(level)(count_over_t

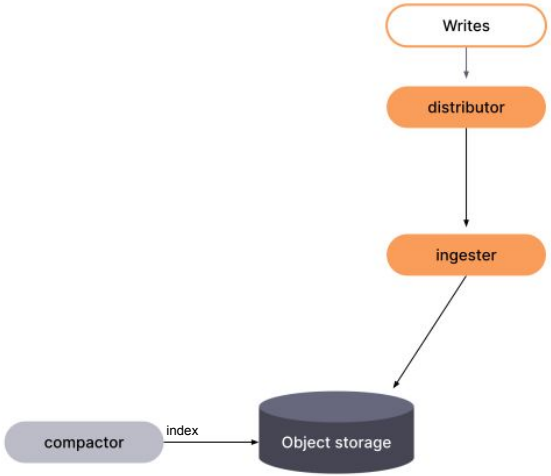

Explore
loki-dev
Split
Last 1 hour UTC
Add to dashboard
Run query
Live
A
(loki-dev)
Log browser ›
sum by (container) (rate({namespace="loki-dev-005"} |= "error"[$__interval]))
Query type
Range
Instant
Line limit
auto
Resolution
1/1
Add query
Query history
Inspector
Graph
Lines
Bars
Points
Stacked lines
Stacked bars
1000
900
800
700
600
500
400
300
200
100
0
14:10
14:15
14:20
14:25
14:30
14:35
14:40
14:45
14:50
14:55
15:00
15:05
{container="cortex-gw"}
{container="distributor"}
{container="freetier-limiter"}
{container="ingester"}
{container="loki-canary"}
{container="querier"}
{container="query-frontend"}
{container="ruler"}

# Low Footprint

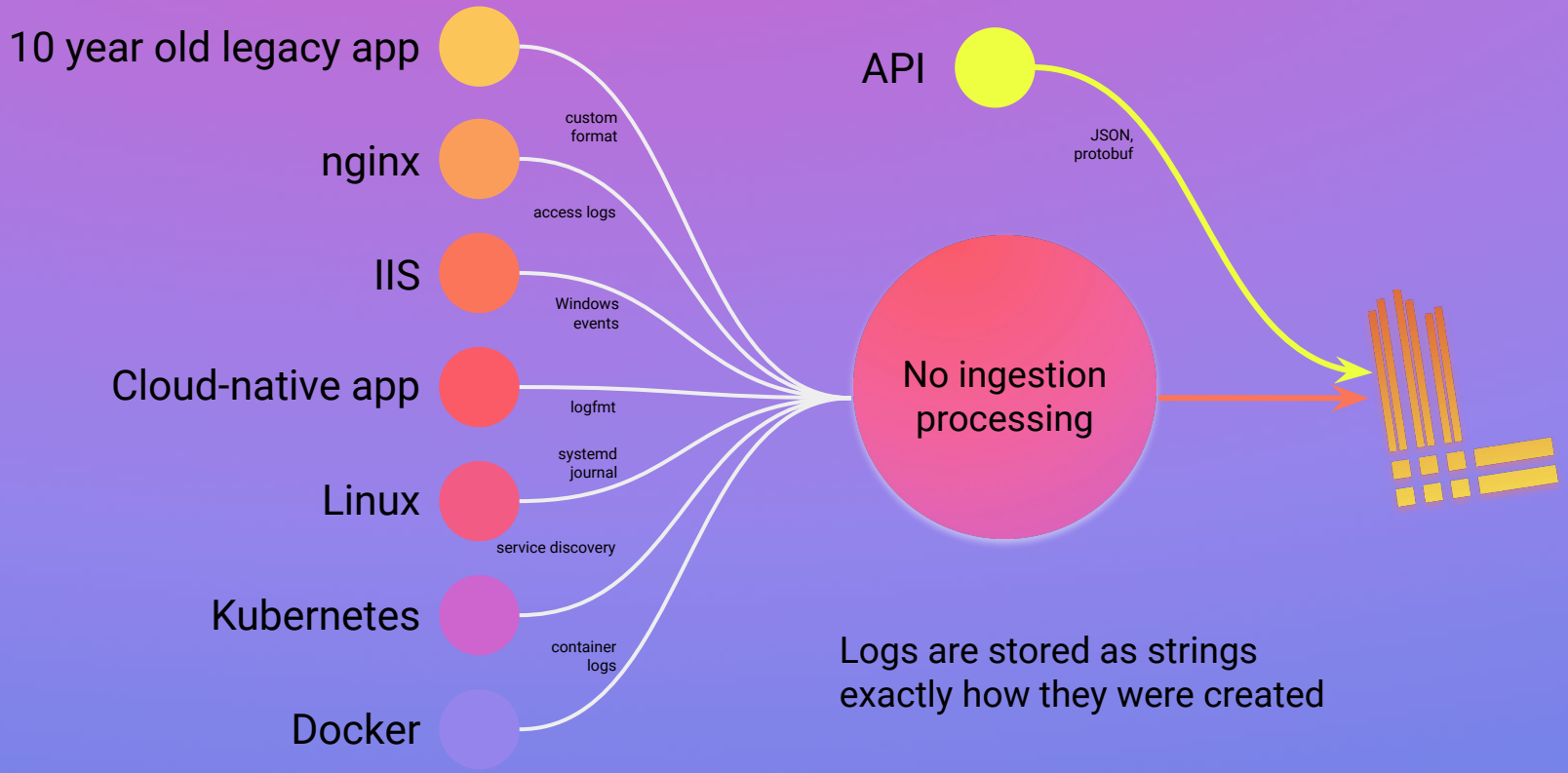
- Individually scalable read/write paths
- Commodity hardware
- Only dependency is object storage. cheap.

# Architecture

- Ingestion



# Loki accepts it all

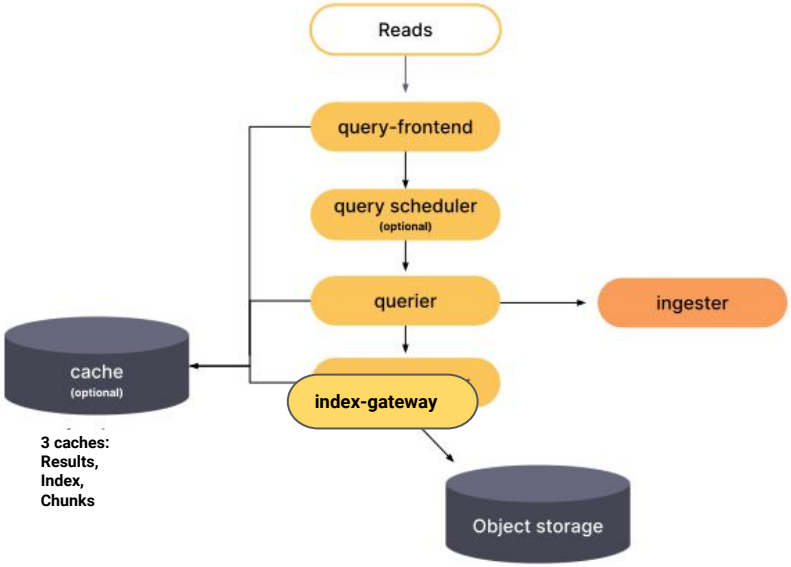



Logs are stored as strings exactly how they were created

## Simplified ingestion

# Architecture

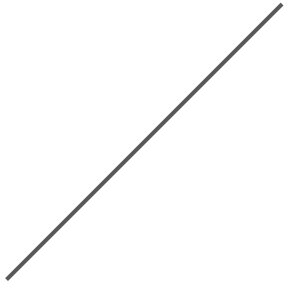
- Query



# Index

**140MB**

Index

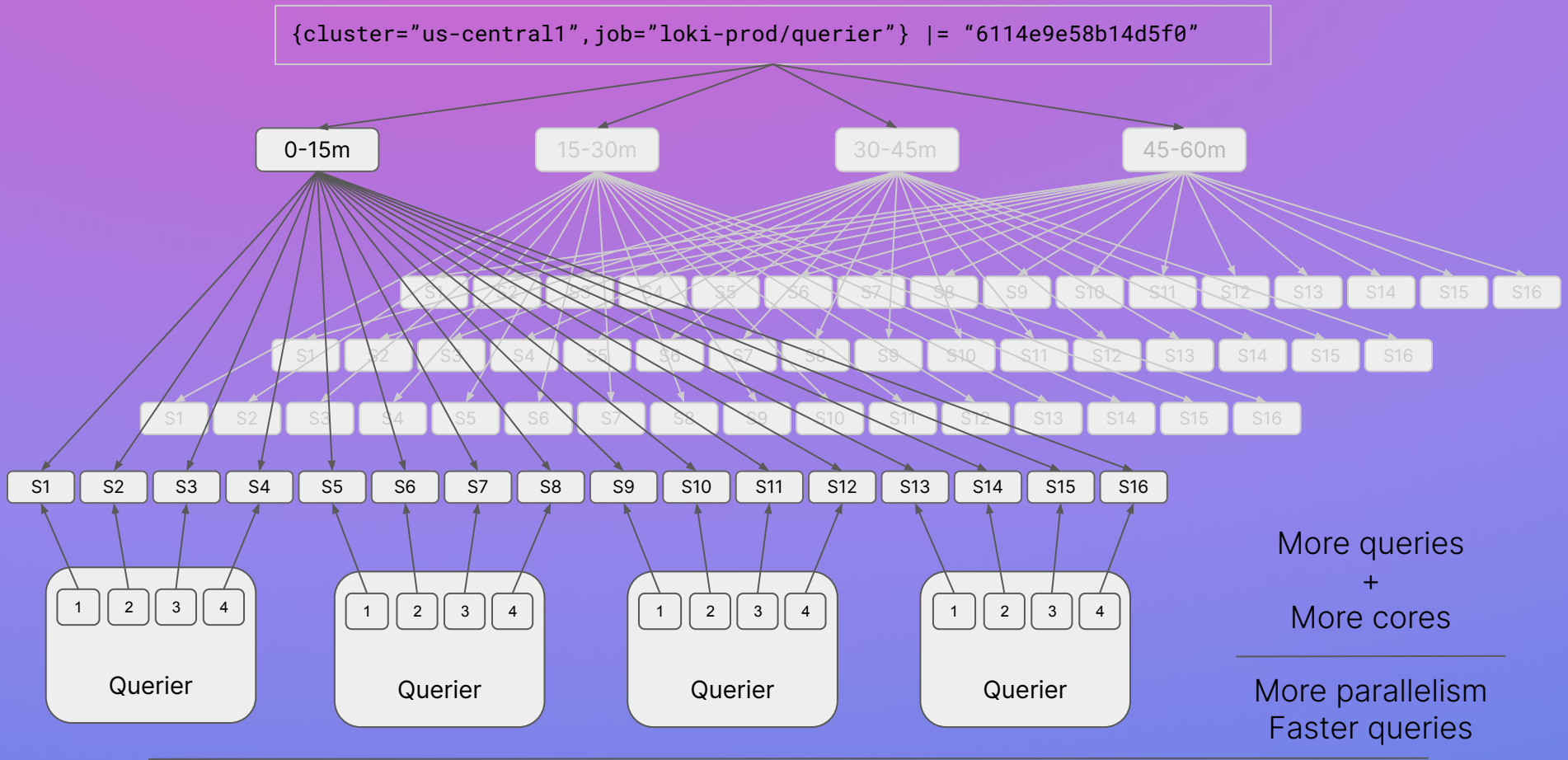


**37TB**

Log Data

270 *thousand* times smaller

How do we make Loki fast without a big index?

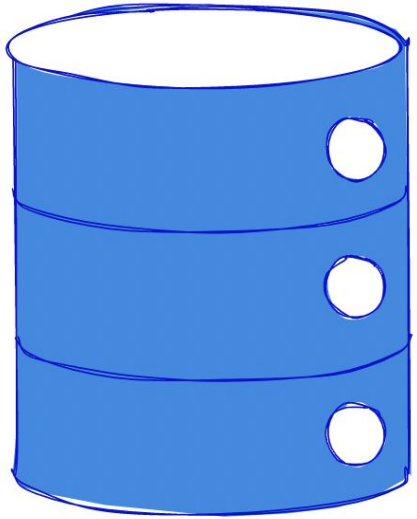

{cluster="us-central1",job="loki-prod/querier"} |= "6114e9e58b14d5f0"
0-15m
15-30m
30-45m
45-60m
S1
S2
S3
S4
S5
S6
S7
S8
S9
S10
S11
S12
S13
S14
S15
S16
1
2
3
4
Querier
Querier
Querier
Querier
More queries
+
More cores
More parallelism
Faster queries

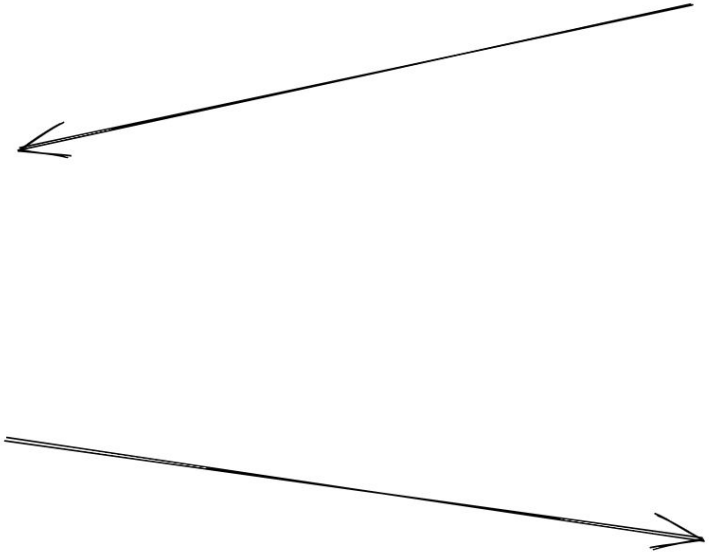
# Retention

- Without indices, retention is just long term storage cost (S3, GCS, etc)
- Application logs for 30d, audit logs for 1y? Easy and (almost) free

Index. 55GB/y

Object storage (low cost). 15PB/y





* All data is active

# Easy Operations

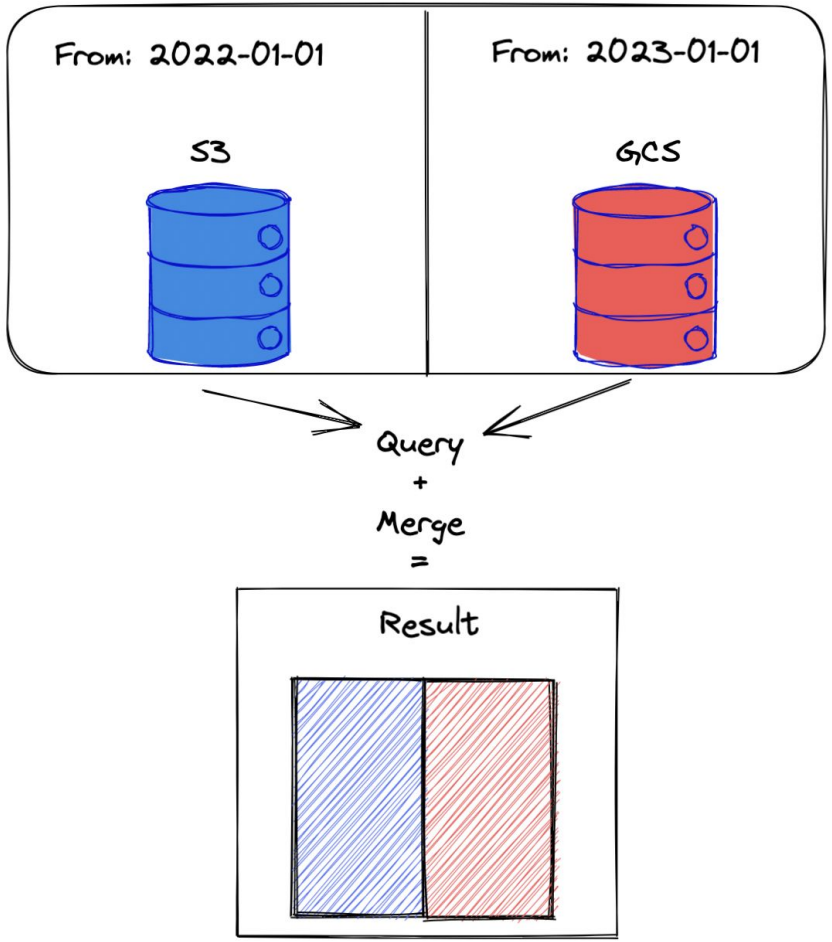
- Migrations
- Zero downtime

# As easy as

```yaml
schema_config:
  configs:
    - from: 2020-10-24
      store: boltdb-shipper
      object_store: filesystem
      schema: v11
      index:
        prefix: index_
        period: 24h
    - from: 2022-01-01
      store: tsdb
      object_store: filesystem
      schema: v12
      index:
        prefix: tsdb_index_
        period: 24h
```

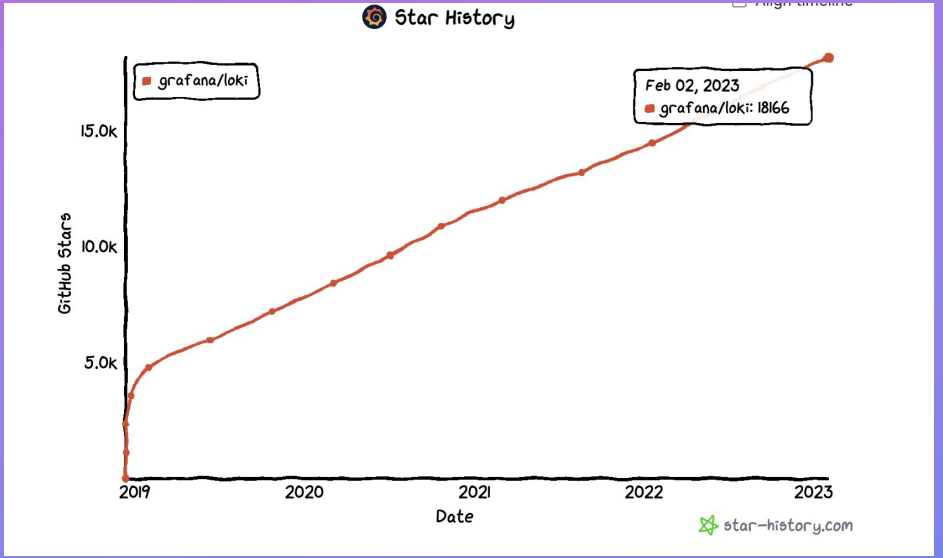
“Continue to query the old version”

“Starting in 2022, use new version”


From: 2022-01-01
From: 2023-01-01
S3
GCS
Query
+
Merge
=
Result

# Open Source Community

- Launched as open source in 2019
- Growing active community
- Talk to us
  - https://grafana.slack.com #loki
  - https://community.grafana.com/
  - Monthly community call with two different timezones
- Docs
  - https://grafana.com/docs/loki/latest/



# Thank you

- Distributed grep
- Low footprint & schema-free
- Low cost retention
- Easy operations
- Open Source Community

Want to talk more after? Will be in the hallway for 10m

@kvrajk

castle_vanity