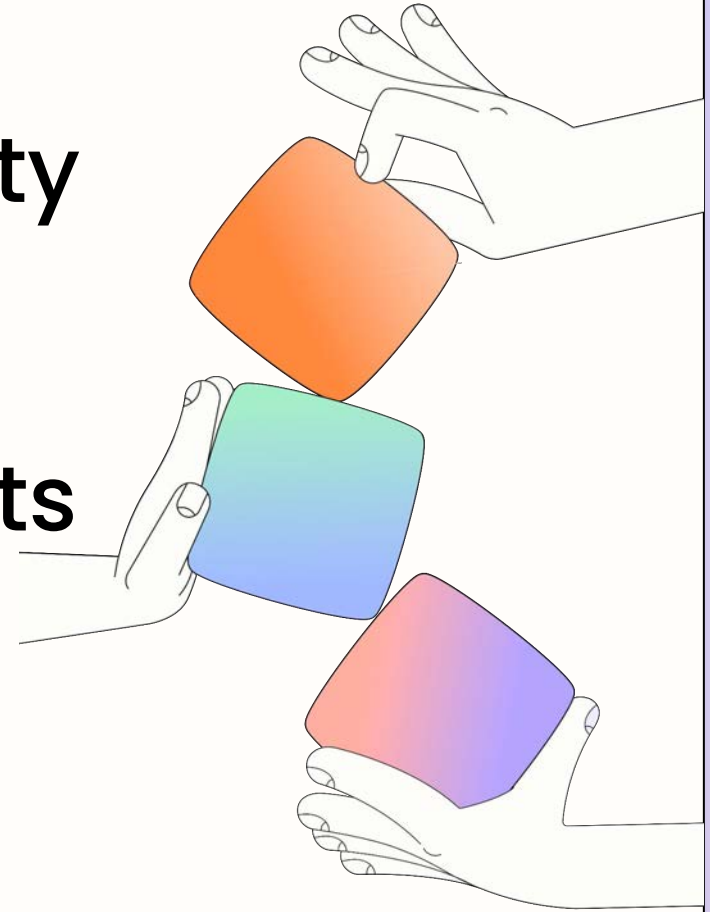




Optimizing Observability with OpenTelemetry Collector for Budget-Friendly Insights

Bruno Ferreira - DevOps Engineer





2023 Tech Layoffs

👤 = 100

84,714



CLOUD SERVICES / OBSERVABILITY / OPERATIONS / SOFTWARE DEVELOPMENT

Datadog's \$65M Bill and Why Developers Should Care

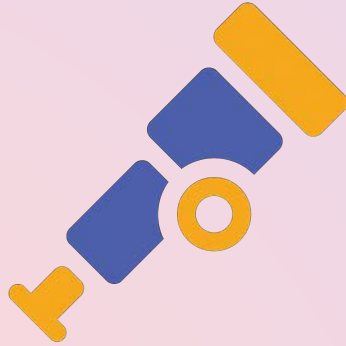
Observability bills can surprise companies, and developers may be on the hook for changing code if bills grow too large.

May 17th, 2023 10:56am by [Loraine Lawson](#)



Source: [TechCrunch](#), [Layoffs.fyi](#) • This chart was last updated January 8, 2024.

Some numbers in this chart may have been estimated based off percent of employees being laid off and company headcount that's publicly available.



OpenTelemetry





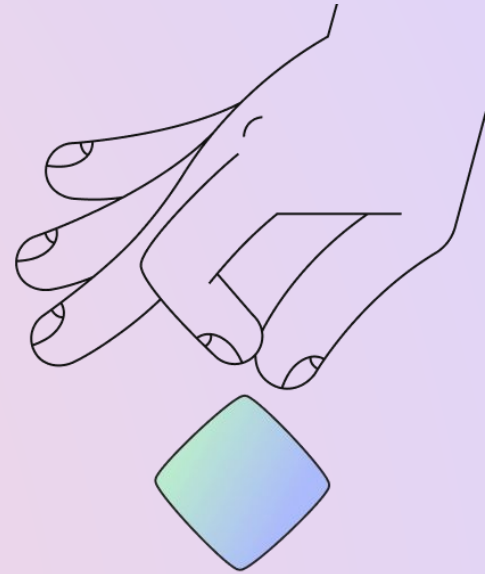
**Do we really need
all that data?** 🤔

Probably not... 🙄



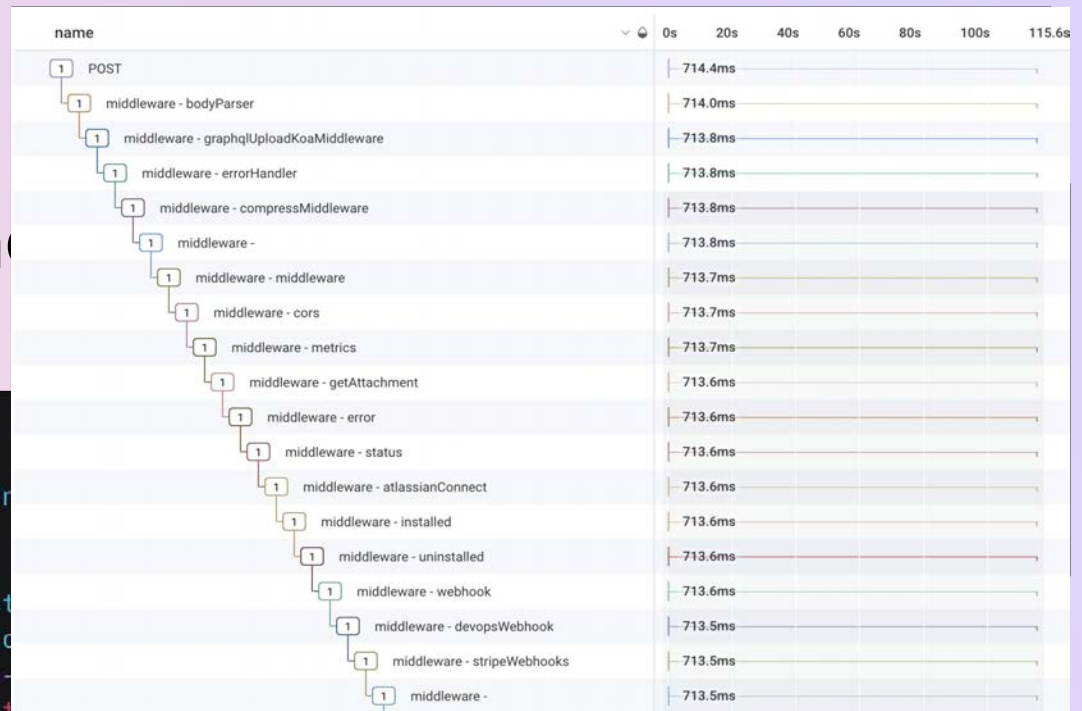


Pick the data that is useful





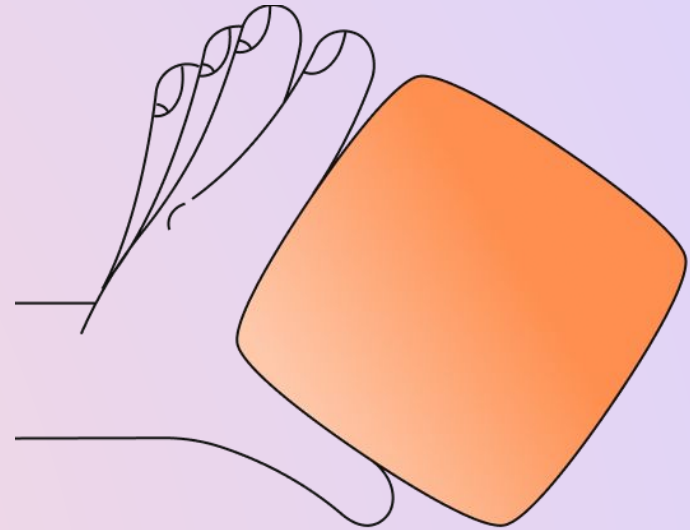
Auto-instrument



```
1 import { getNodeAutoInstrumentations, InstrumentationNodeModuleDefinition } from '@opentelemetry/instrumentations-node'
2
3 export const initTracing = (nodeAutoInstrumentations, instrumentationNodeModuleDefinition) => {
4   const defaultNodeAutoInstrumentationConfig = {
5     '@opentelemetry/instrumentation-aws-lambda': {
6       suppressInternalInstrumentation: true
7     },
8     '@opentelemetry/instrumentation-koa': {
9       ignoreLayersType: [KoaLayerType.MIDDLEWARE]
10    }
11  }
12 }
```



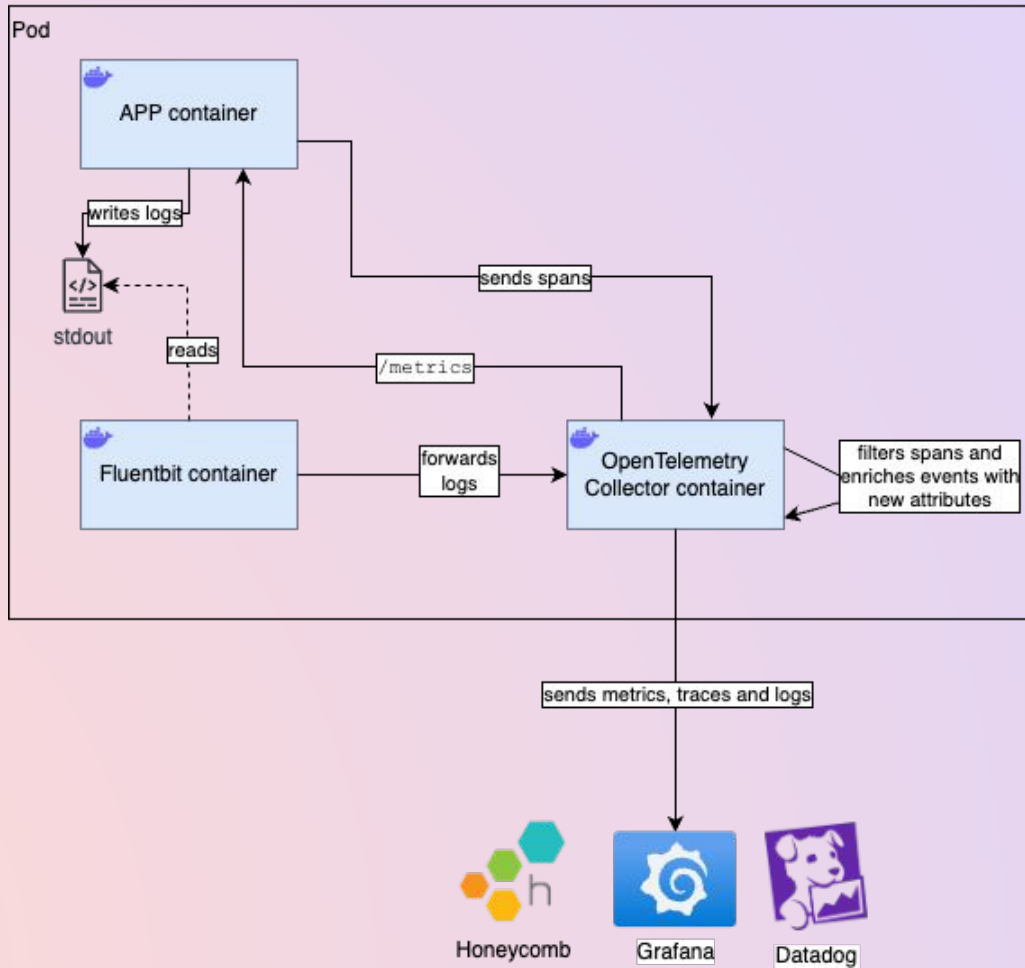
Tail sampling

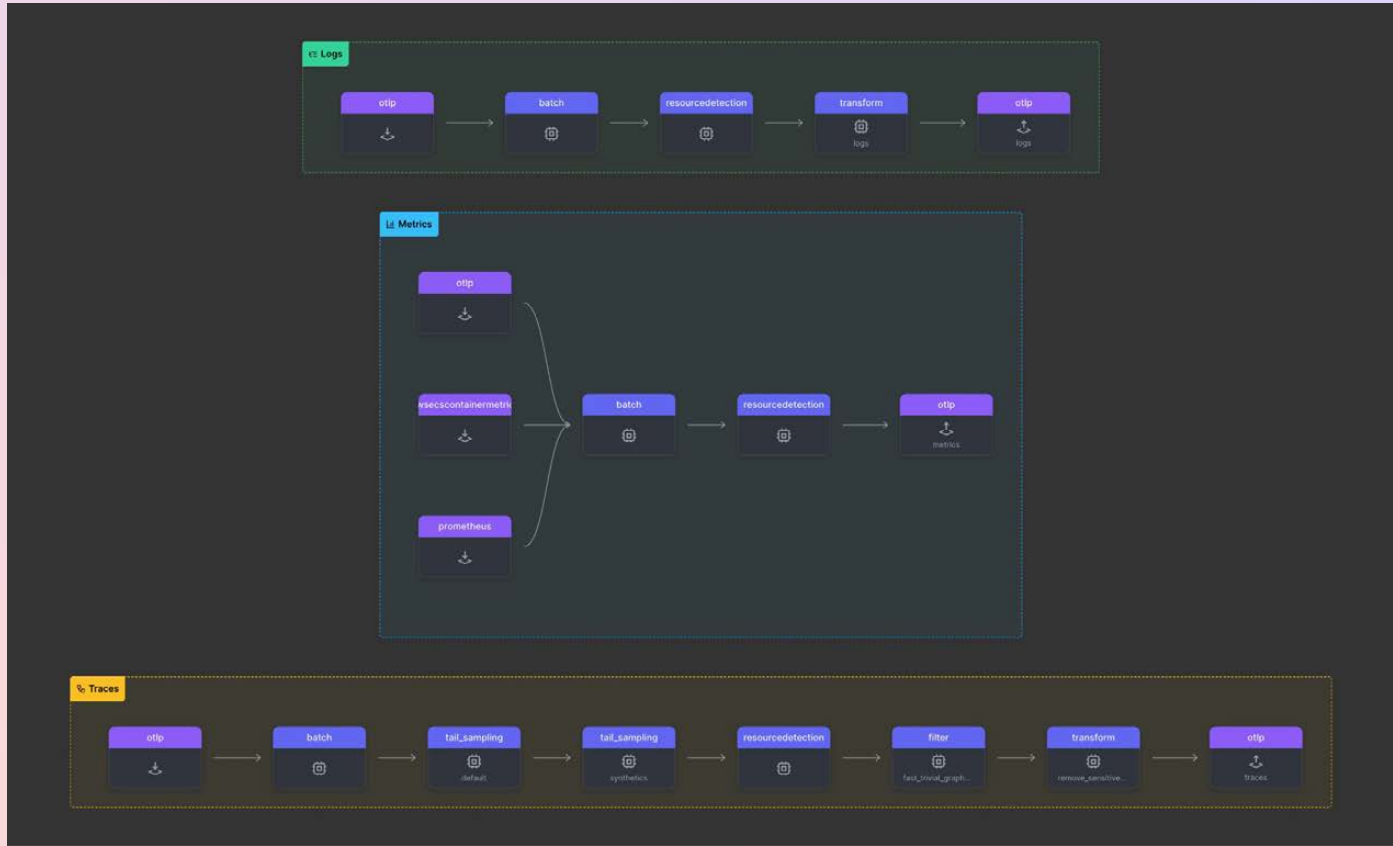




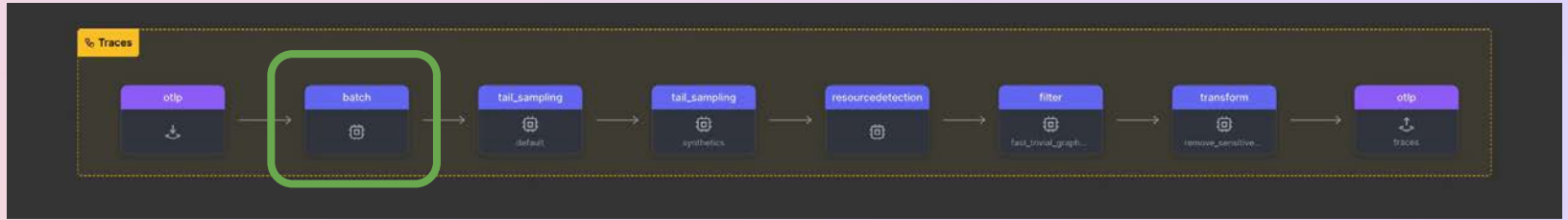
Opentelemetry collector to the rescue! ☐

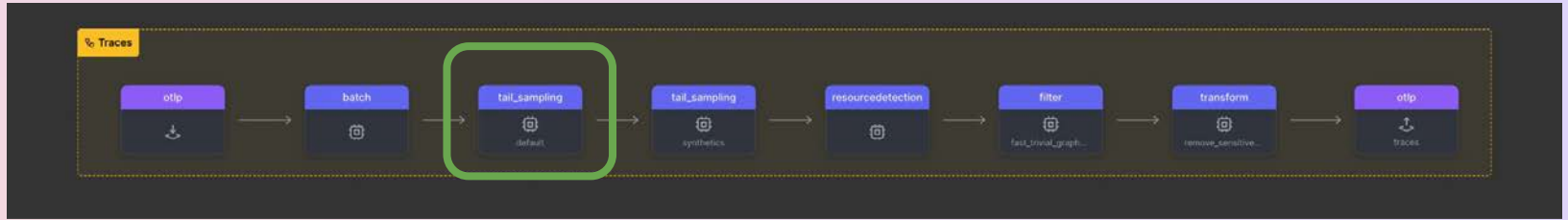
- AWS Distro for OpenTelemetry (ADOT): <https://aws-otel.github.io/>
 - Able to load the configuration file from S3 👍
 - Does not support the [transform processor](#):
<https://github.com/aws-observability/aws-otel-collector/issues/2300>
 - Didn't support logs at the time, [now it does](#) 🎉
- OpenTelemetry Core:
<https://github.com/open-telemetry/opentelemetry-collector>
- **OpenTelemetry Contrib:**
<https://github.com/open-telemetry/opentelemetry-collector-contrib>
 - How to build a secure opentelemetry collector:
<https://martinjt.me/2023/12/20/building-a-secure-opentelemetry-collector/>
- Refinery: <https://github.com/honeycombio/refinery>

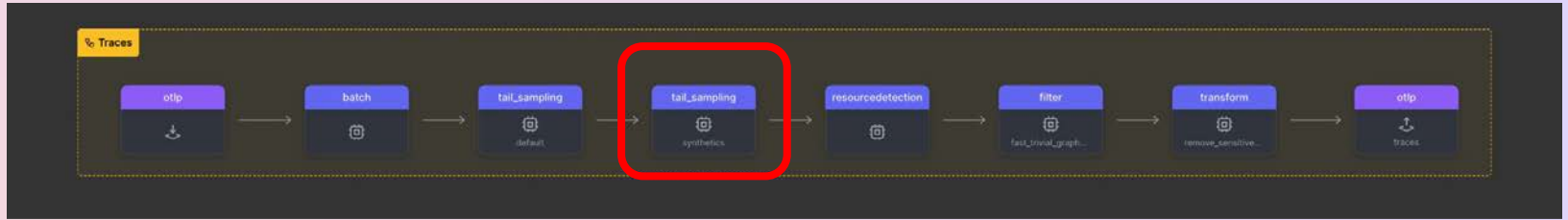




generated on <https://www.otelbin.io/>

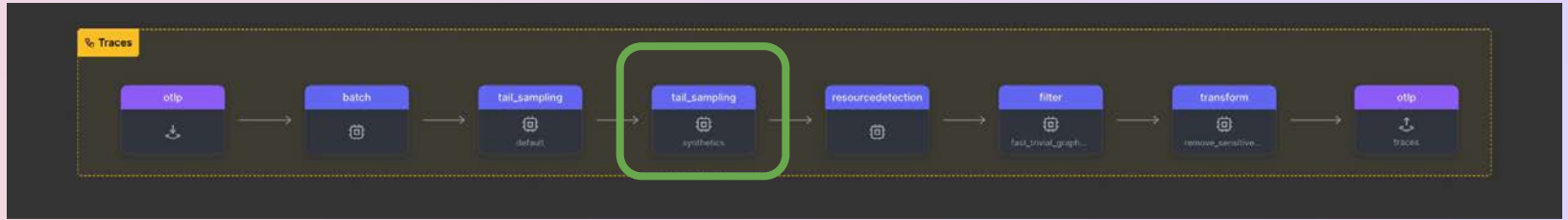








```
tail_sampling/default:
  policies: [
    # sample any request that has an error status
    {
      name: errors-policy,
      type: status_code,
      status_code: { status_codes: [ERROR] },
    },
    # sample any request that takes longer than 100ms
    {
      name: latency-policy,
      type: latency,
      latency: { threshold_ms: 100 },
    },
    # sample traces with ResyncProject as graphql.operation.name
    {
      name: resync-project-policy,
      type: string_attribute,
      string_attribute: { key: graphql.operation.name, values: [ResyncProject] },
    },
  ]
}
```



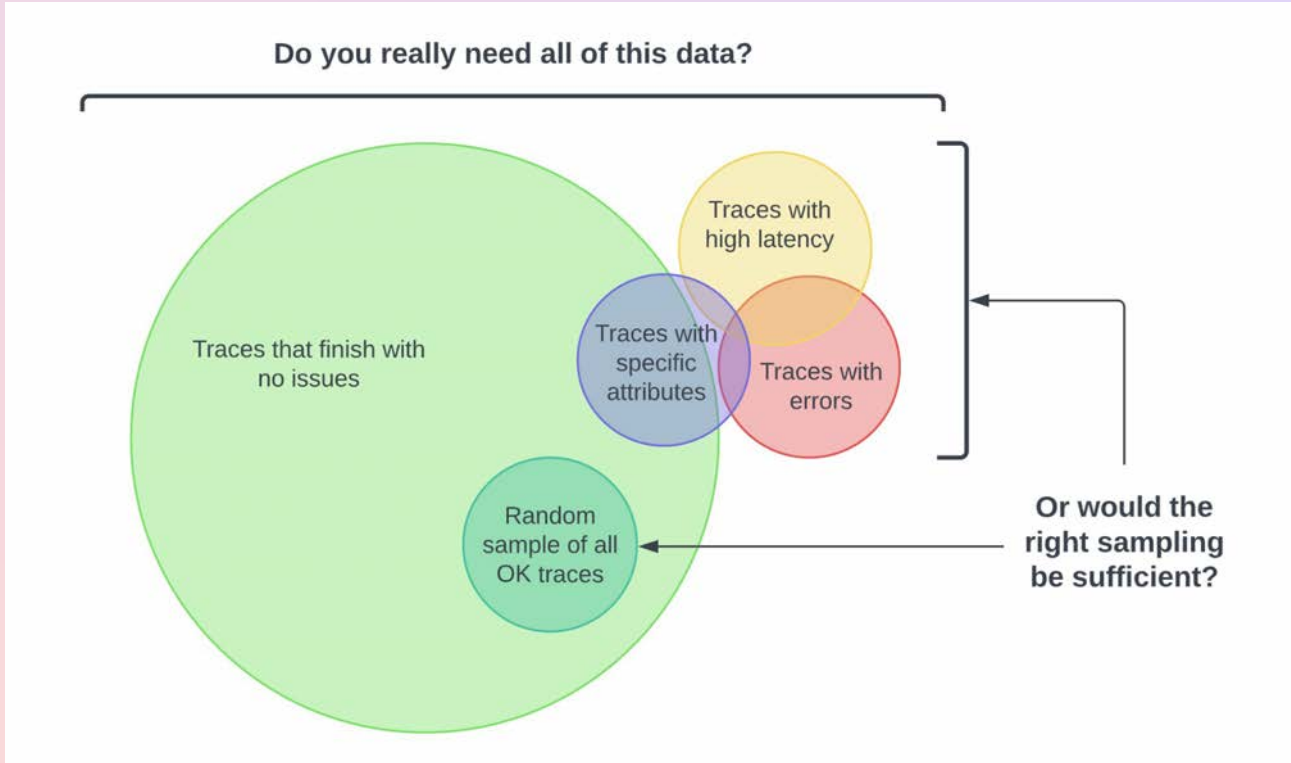


```
tail_sampling/synthetics:
  policies: [
    {
      name: errors-policy,
      type: status_code,
      status_code: { status_codes: [ERROR] },
    },
    {
      name: latency-policy,
      type: latency,
      latency: { threshold_ms: 1000 },
    },
    # sample 1% of synthetic requests (i.e. requests with a user agent that matches CloudWatch.*)
    {
      name: synthetic-policy,
      type: and
      and: {
        and_sub_policy: [
          {
            name: user-agent-policy,
            type: string_attribute,
            string_attribute: { key: http.user_agent, values: [CloudWatch.*], enabled_regex_matching: true },
          },
          {
            name: probabilistic-policy,
            type: probabilistic,
            probabilistic: { sampling_percentage: 1 }
          },
        ],
      },
    },
    # sample 100% of non-synthetic requests
    {
      name: default-policy,
      type: and
      and: {
        and_sub_policy: [
          {
            name: user-agent-policy-inverted,
            type: string_attribute,
            string_attribute: { key: http.user_agent, values: [CloudWatch.*], enabled_regex_matching: true
invert_match: true },
          },
          {
            name: default-always-sample,
            type: always_sample,
          },
        ],
      },
    },
  ],
}
```




```
filter/fast_trivial_graphql_resolve:
  error_mode: ignore
  spans:
    exclude:
      match_type: regexp
      span_names: ["graphql.resolve"]
      attributes:
        - key: "graphql.field.type"
          value:
```

```
(UUID!|UUID|String!|String|Float!|Float|DateTime!|DateTime|HexColor|SanitizedHtml|ThemeColor|CategorizableEntityType!|WorkflowCategory|CategoryType!)
```



source: <https://opentelemetry.io/blog/2022/tail-sampling/>, by Reese Lee.



Feedback and Questions?

 <https://linkedin.com/in/bmbferreira>

 <https://twitter.com/bmbferreira>

www.theydo.com/careers - We're hiring!

