Sweet Streams are Made of These

Data Driven Development for Stream Processing

@CAITO_200_OK | VERVERICA

Introduction

- Caito Scherr

@CAITO_200_OK

Introduction

- Caito Scherr
- Developer Advocate

@CAITO_200_OK

Introduction

- Caito Scherr
- Developer Advocate
- Ververica, GmbH

@CAITO_200_OK

Introduction

- Caito Scherr
- Developer Advocate
- Ververica, GmbH
- Portland, OR, USA

@CAITO_200_OK

Introduction

@CAITO_200_OK

Agenda

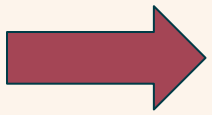
- The Challenge
- “Metrics-Driven Metrics”
- Metrics as a Shared Language

Agenda

- The Challenge
- “Metrics-Driven Metrics”
- Metrics as a Shared Language

Agenda

- The Challenge
- “Metrics-Driven Metrics”
- Metrics as a Shared Language



Before We Start...

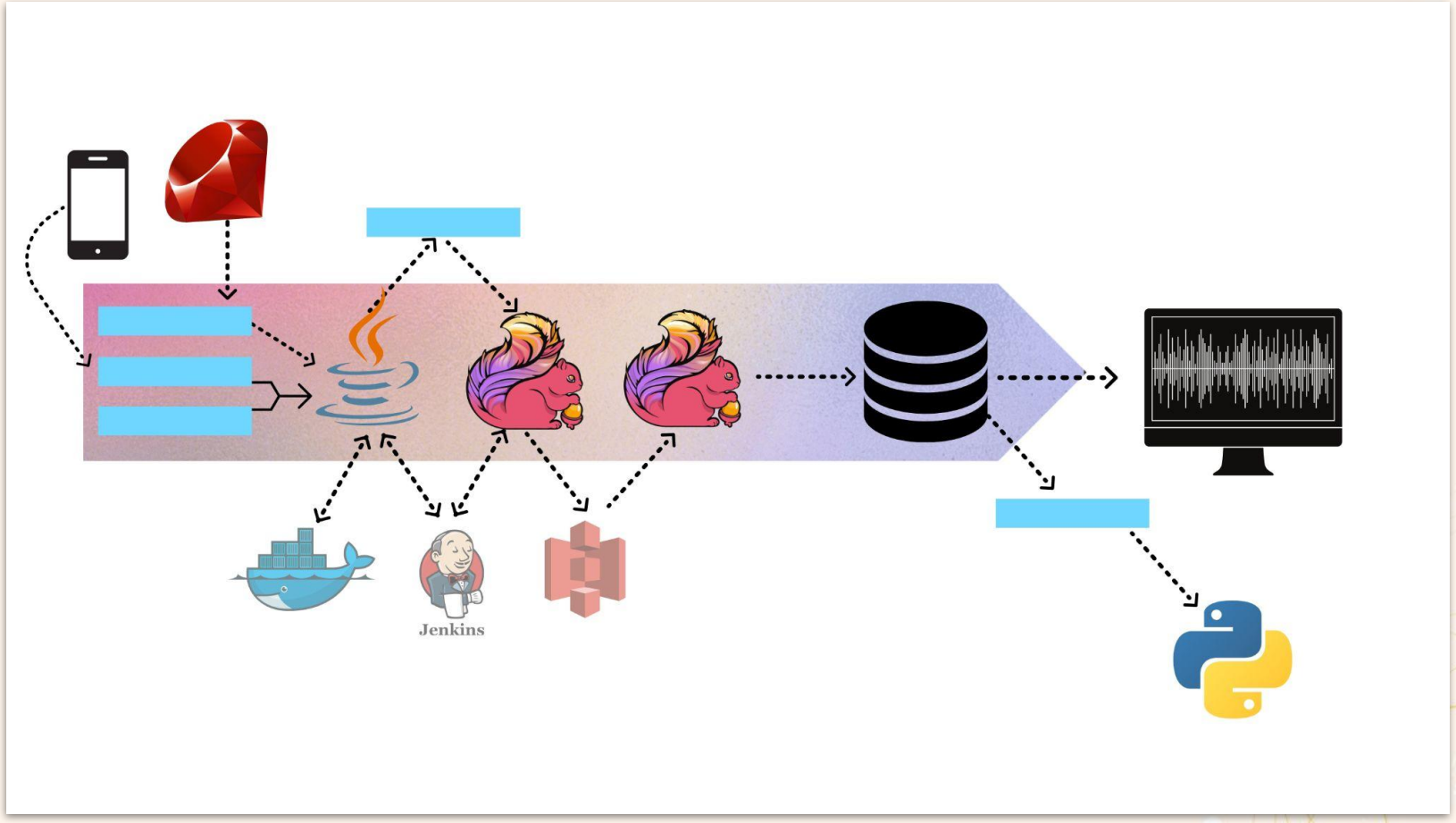
@CAITO_200_OK

Before We Start...

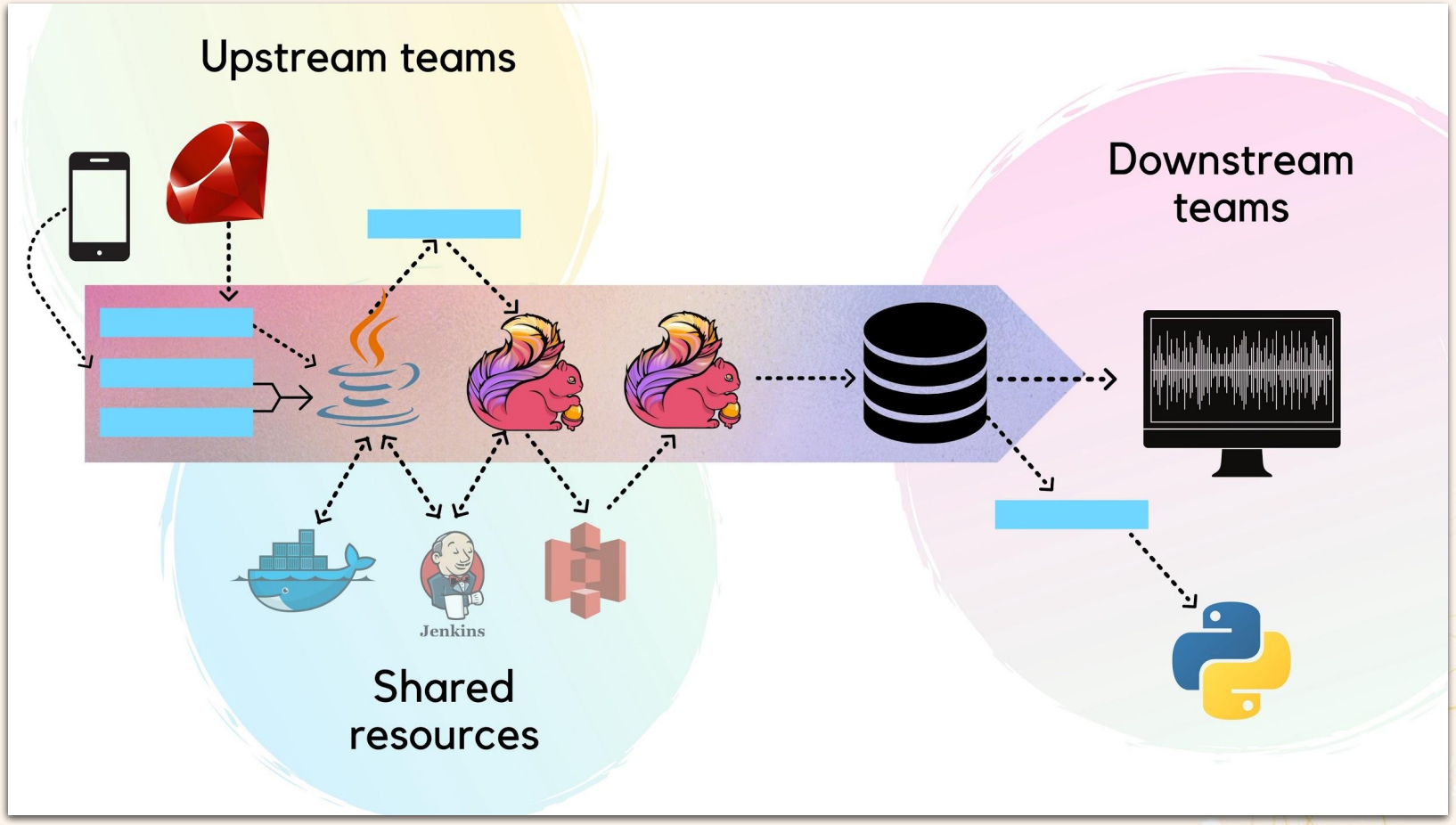
The Challenge

@CAITO_200_OK

The Challenge



The Challenge



"Metrics-Driven Metrics"

@CAITO_200_OK

Metrics-Driven Metrics

> “You can measure almost anything, but you cannot pay attention to everything”

Metrics-Driven Metrics

> “Any situation where people create their own dashboards without structure, quickly starts to look like the cockpit of a 747”

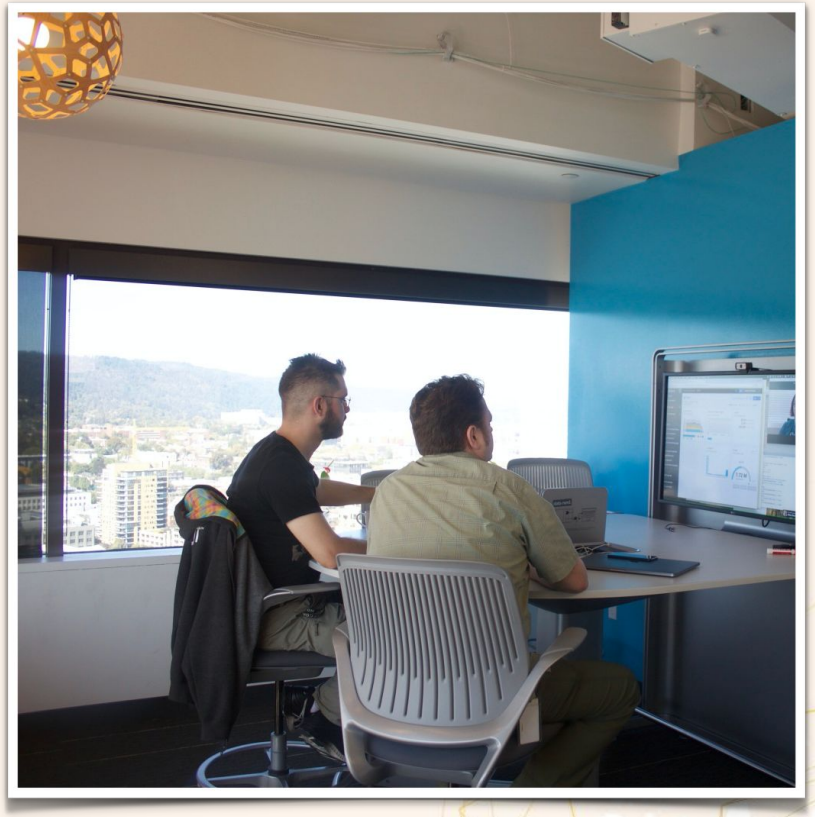
Metrics-Driven Metrics

Metrics-Driven Metrics

Metrics-Driven Metrics

The dashboard should **be**:

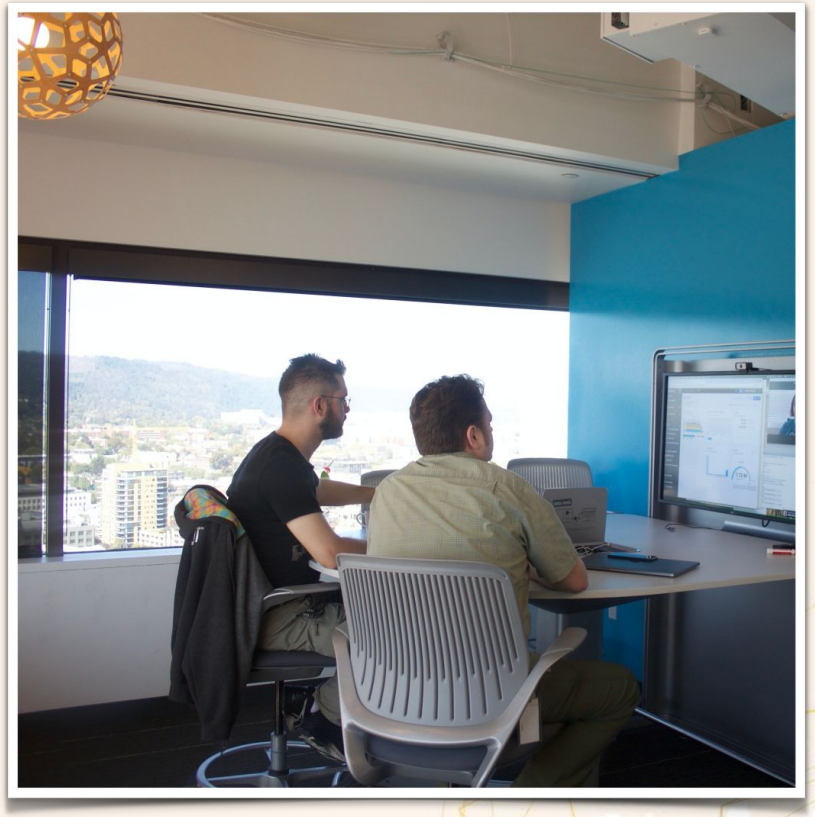
- Meaningful
- Iterative
- Accessible



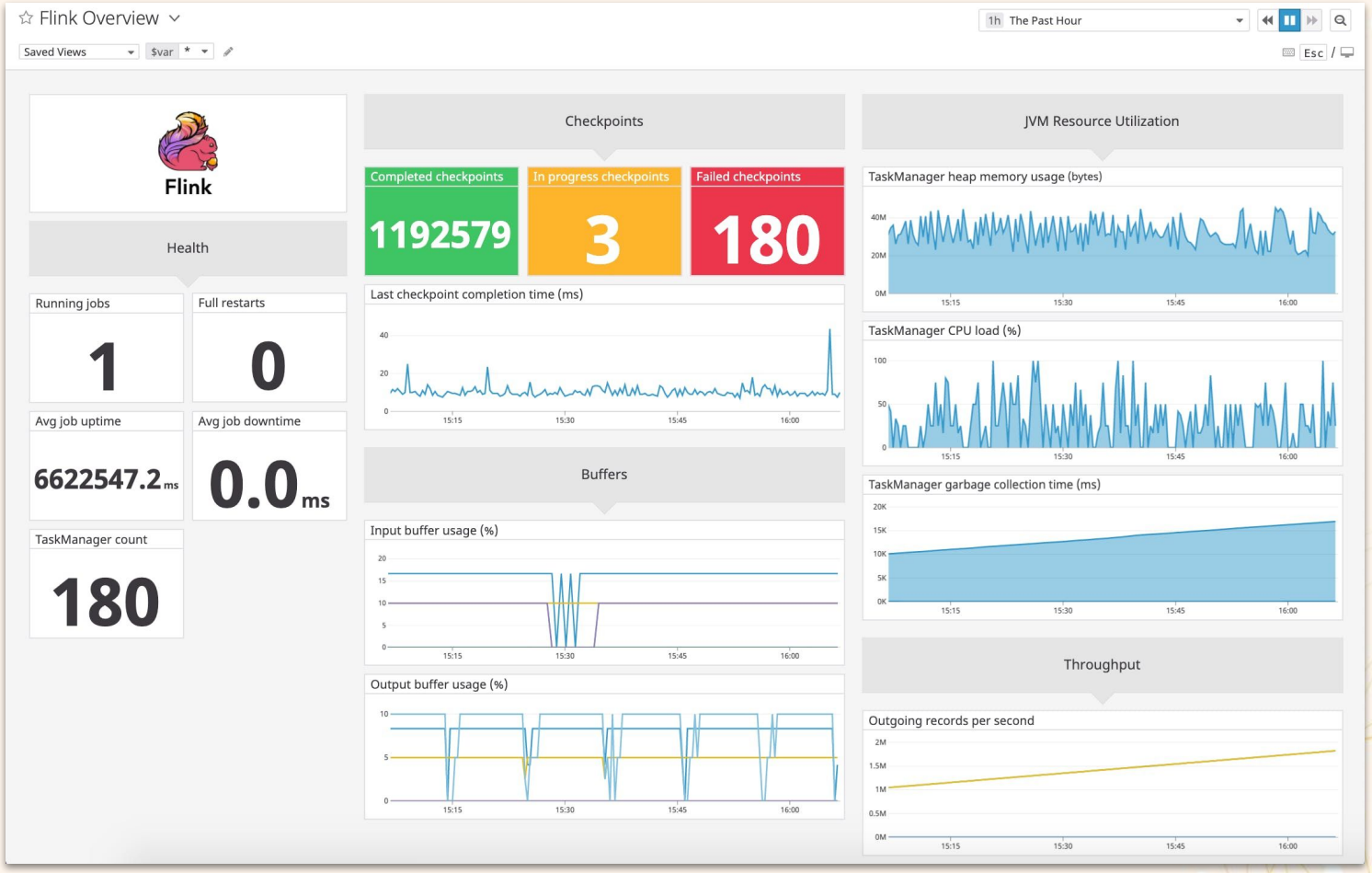
Metrics-Driven Metrics

The dashboard should **reflect** your:

- Roadmap
- Highest Risk
- Most uncertain metrics

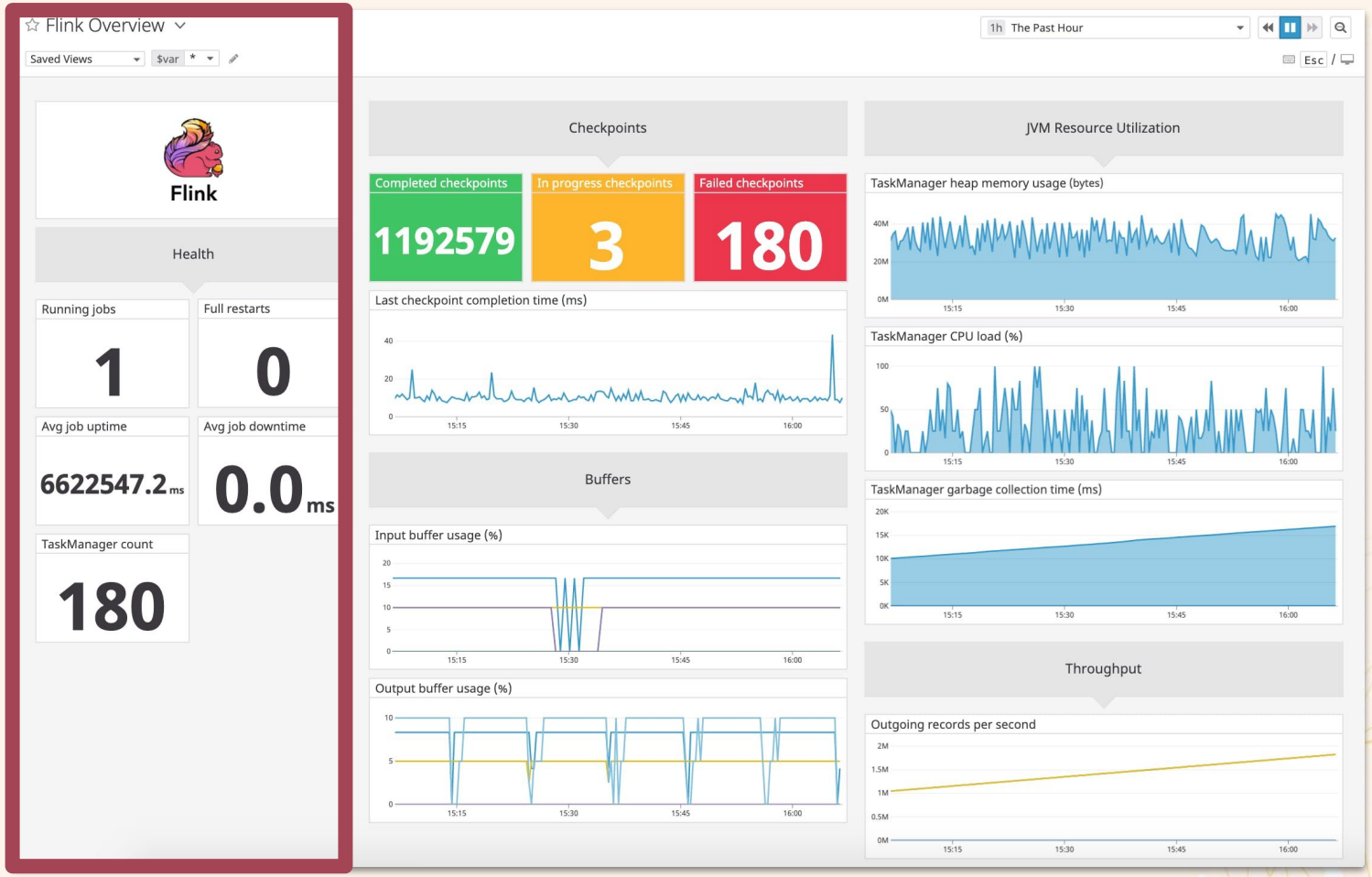


Metrics-Driven Metrics



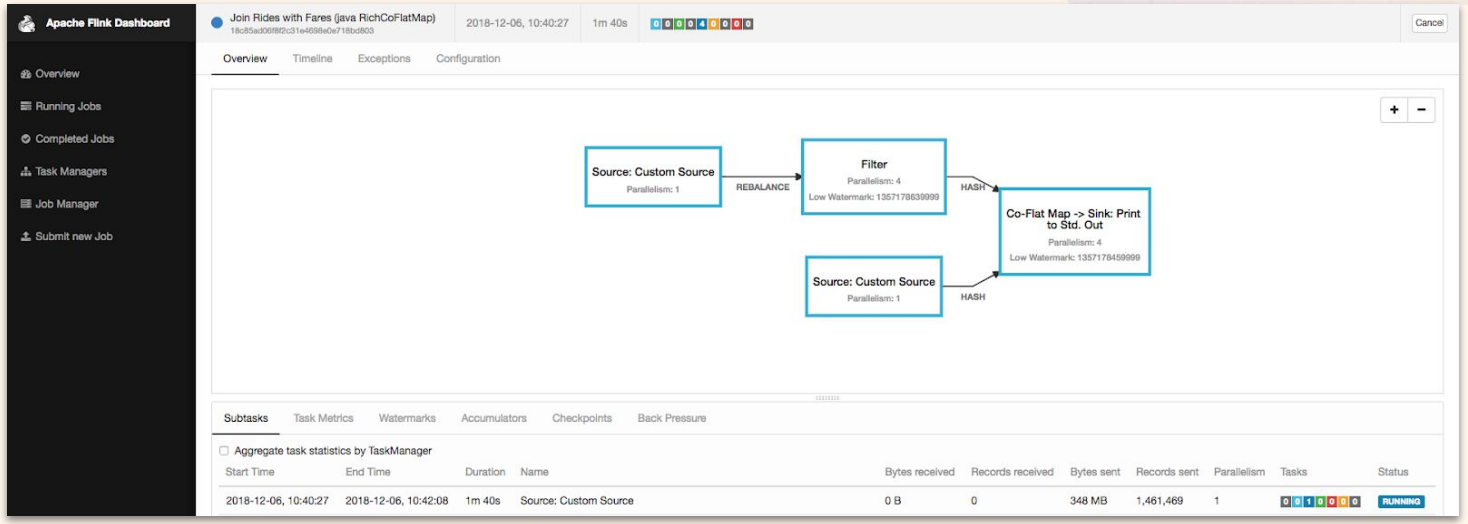
@CAITO_200_OK

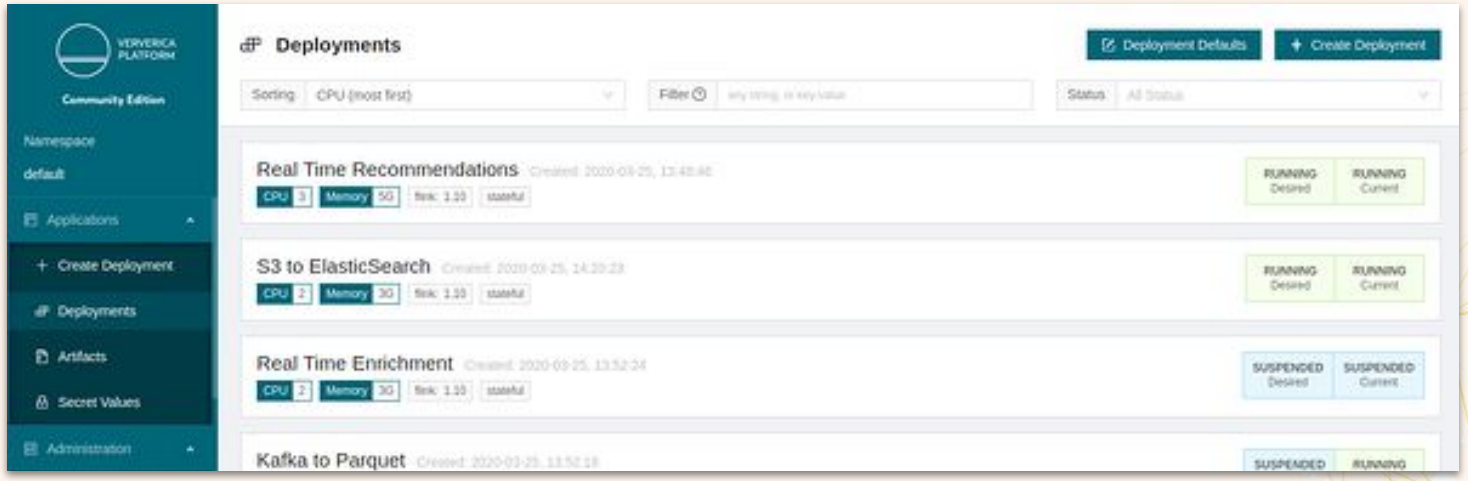
Metrics-Driven Metrics



@CAITO_200_OK

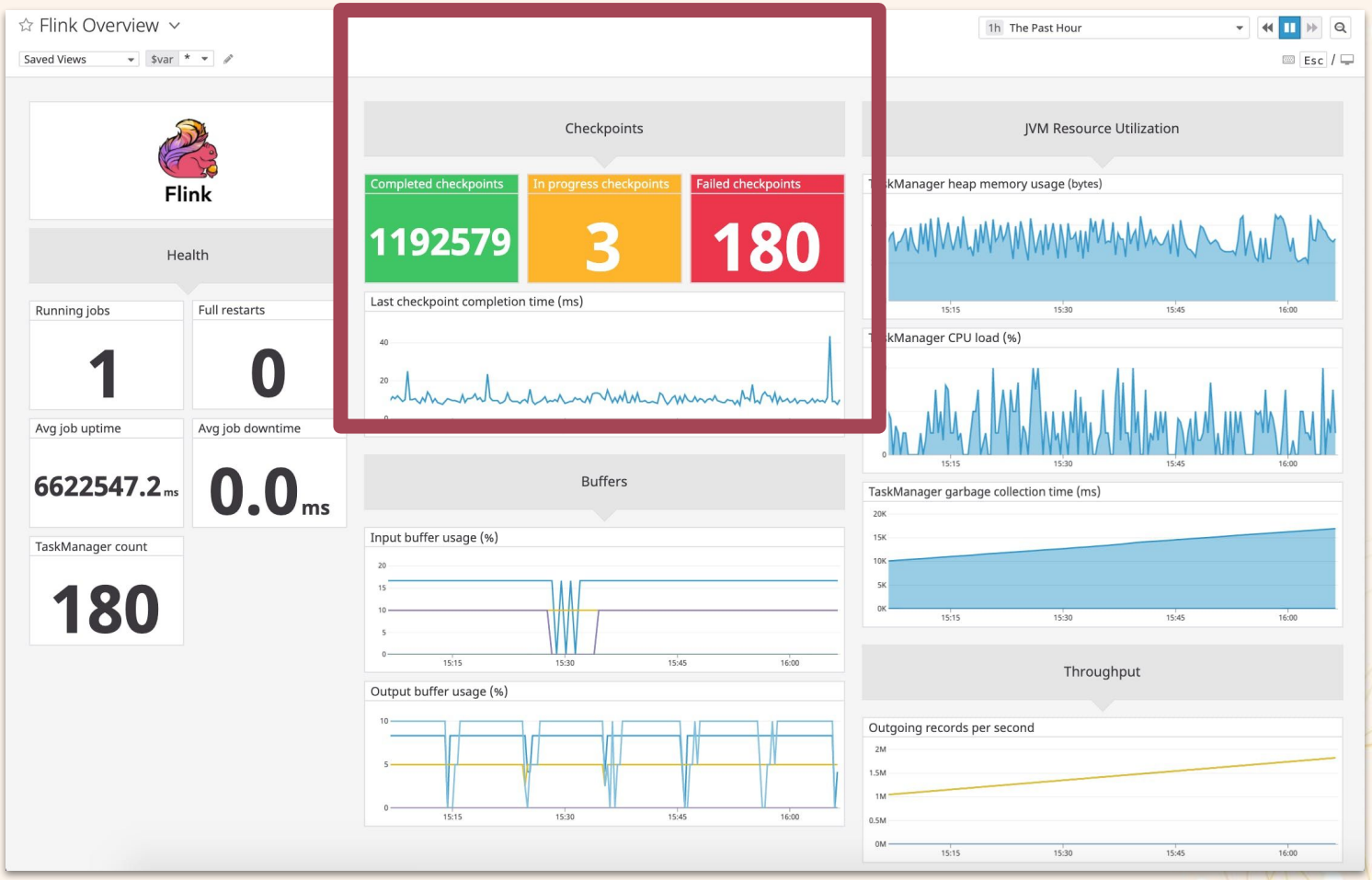
Metrics-Driven Metrics





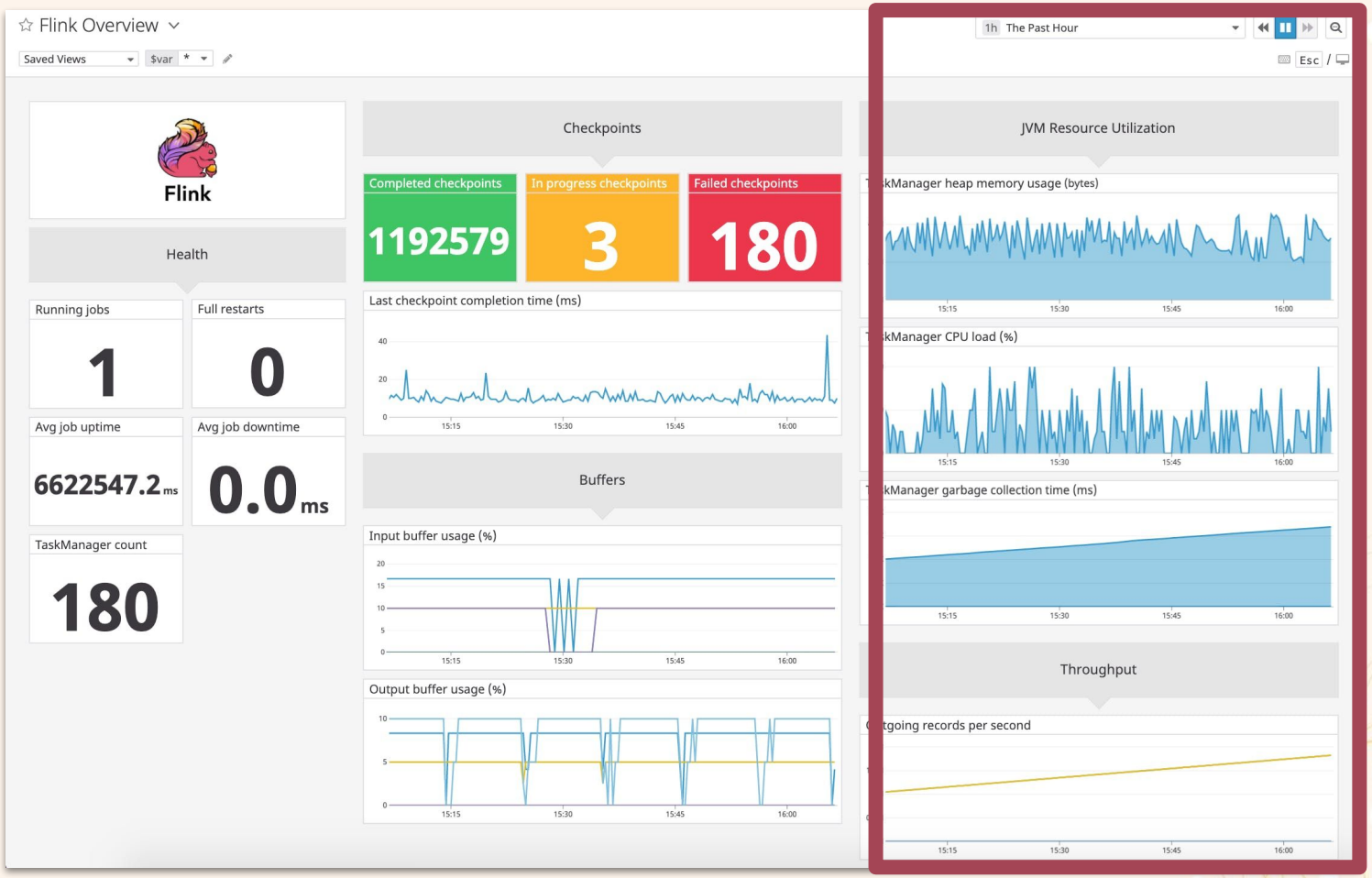
@CAITO_200_OK

Metrics-Driven Metrics



@CAITO_200_OK

Metrics-Driven Metrics



@CAITO_200_OK

Metrics-Driven Metrics

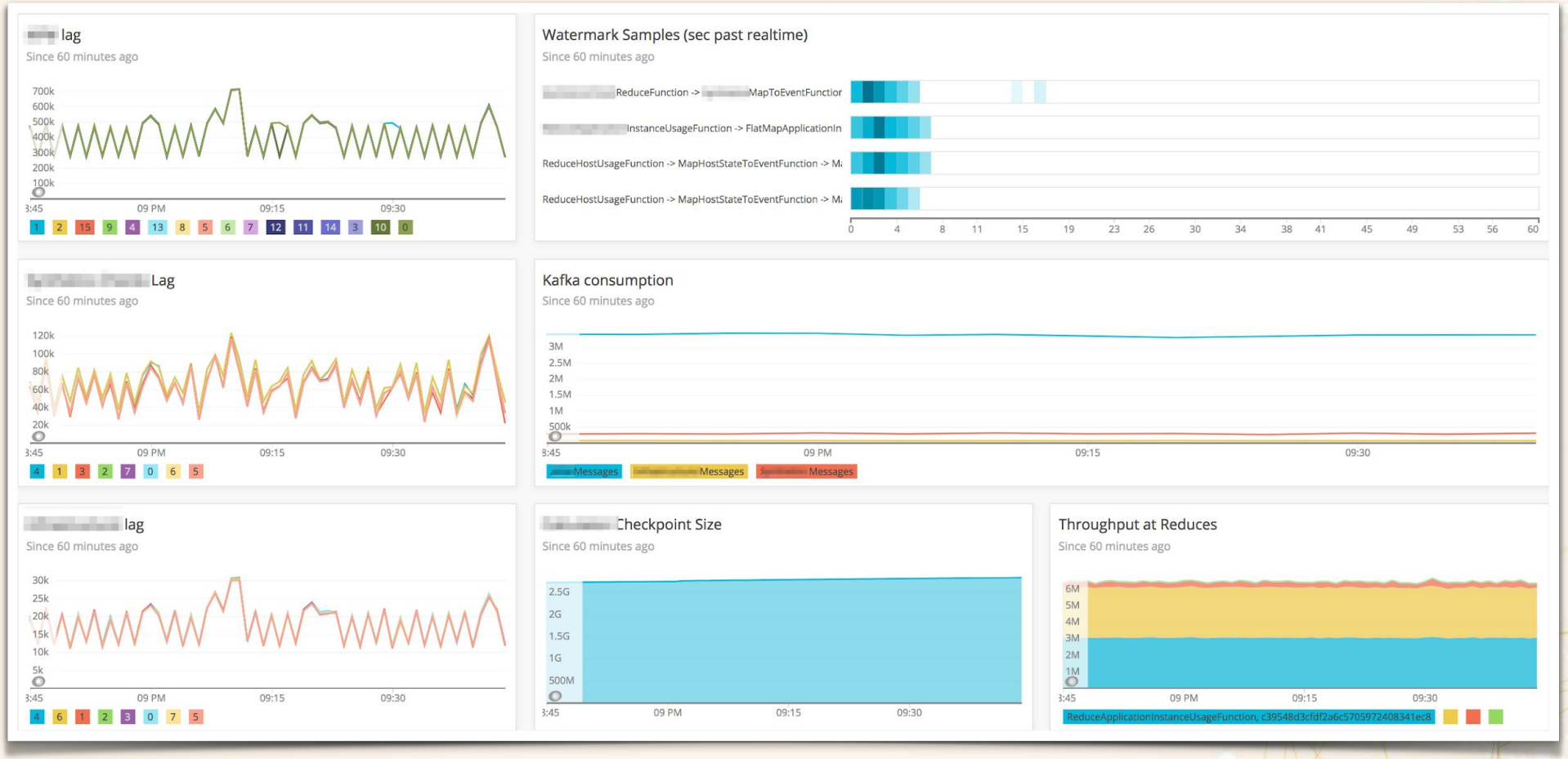
Prometheus

```
metrics.reporter.prom.class: org.apache.flink.metrics.prometheus.PrometheusReporter
```

DataDog

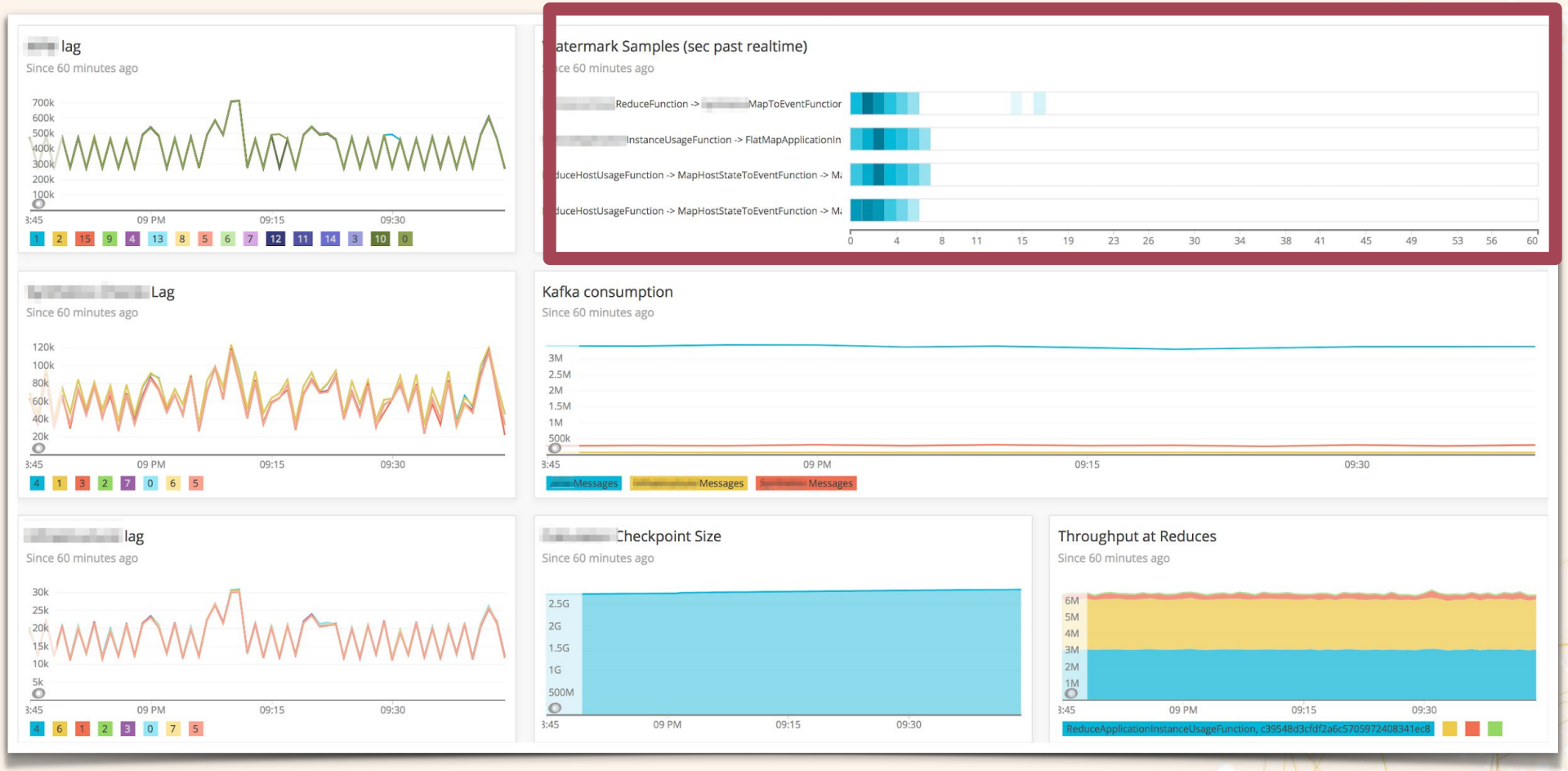
```
metrics.reporter.dghttp.factory.class: org.apache.flink.metrics.datadog.DatadogHttpRepo
rterFactory
metrics.reporter.dghttp.apikey: xxx
metrics.reporter.dghttp.tags: myflinkapp,prod
metrics.reporter.dghttp.proxyHost: my.web.proxy.com
metrics.reporter.dghttp.proxyPort: 8080
metrics.reporter.dghttp.dataCenter: US
metrics.reporter.dghttp.maxMetricsPerRequest: 2000
metrics.reporter.dghttp.interval: 60 SECONDS
```

Metrics-Driven Metrics



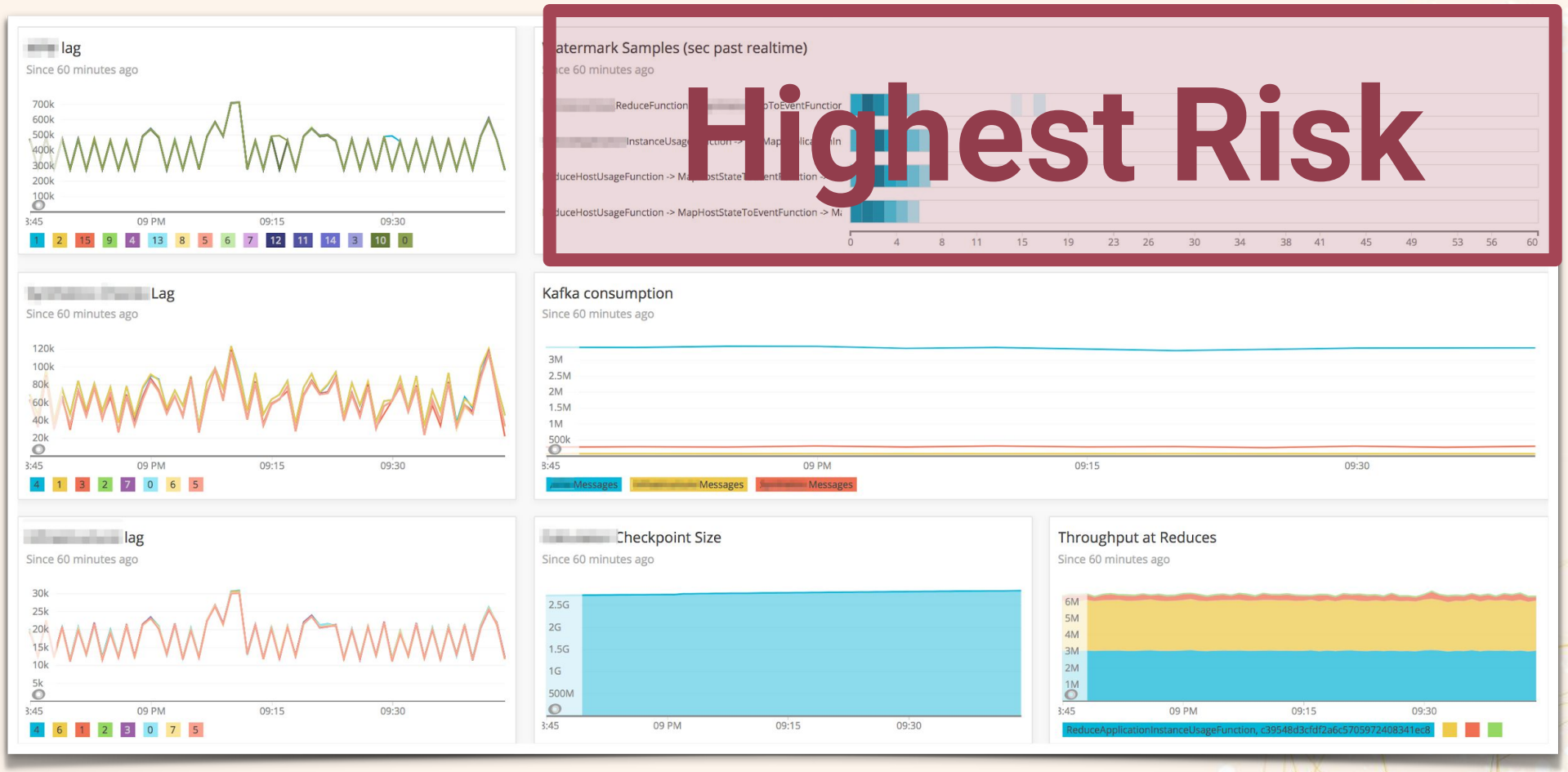
@CAITO_200_OK

Metrics-Driven Metrics



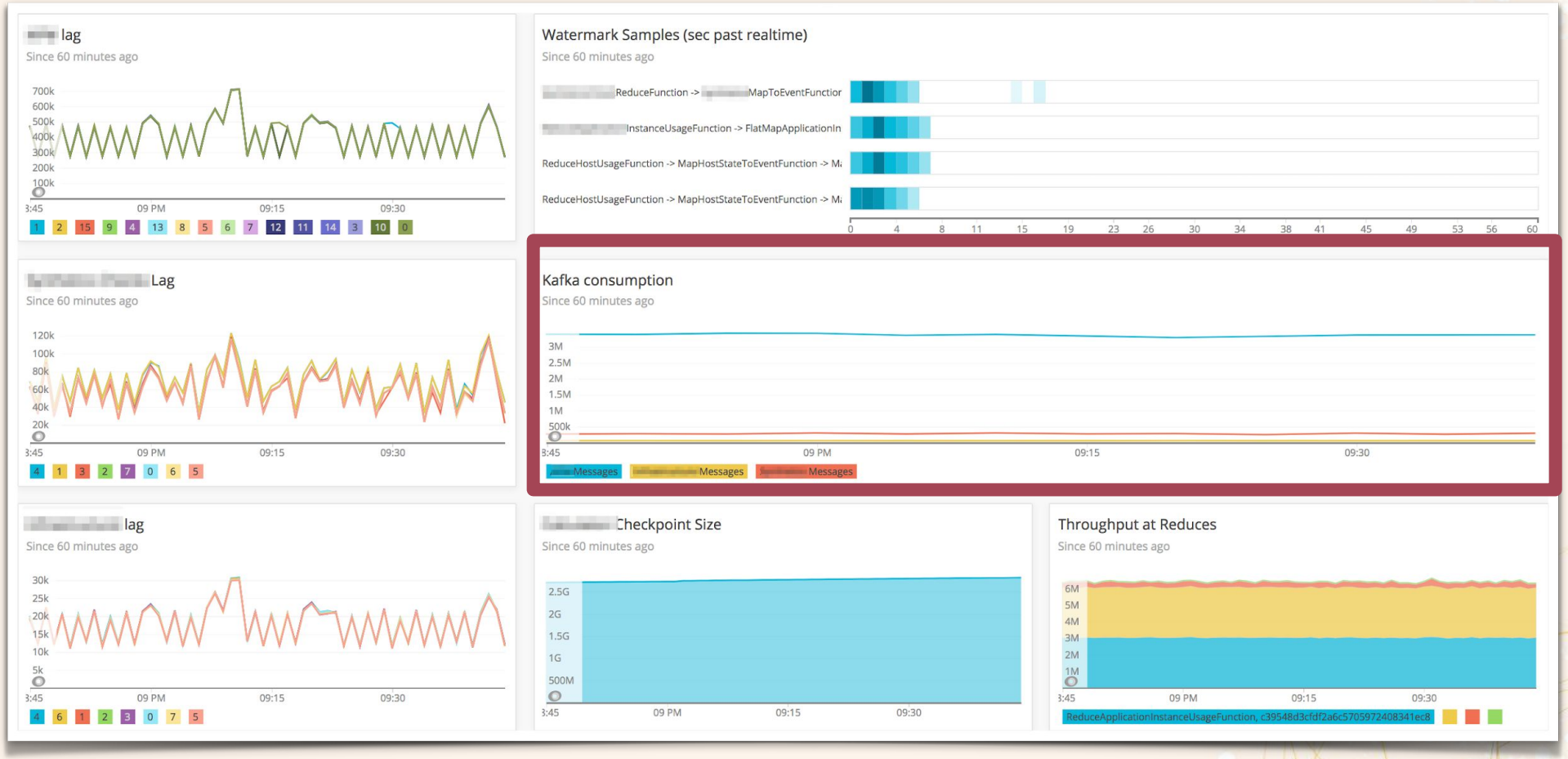
@CAITO_200_OK

Metrics-Driven Metrics



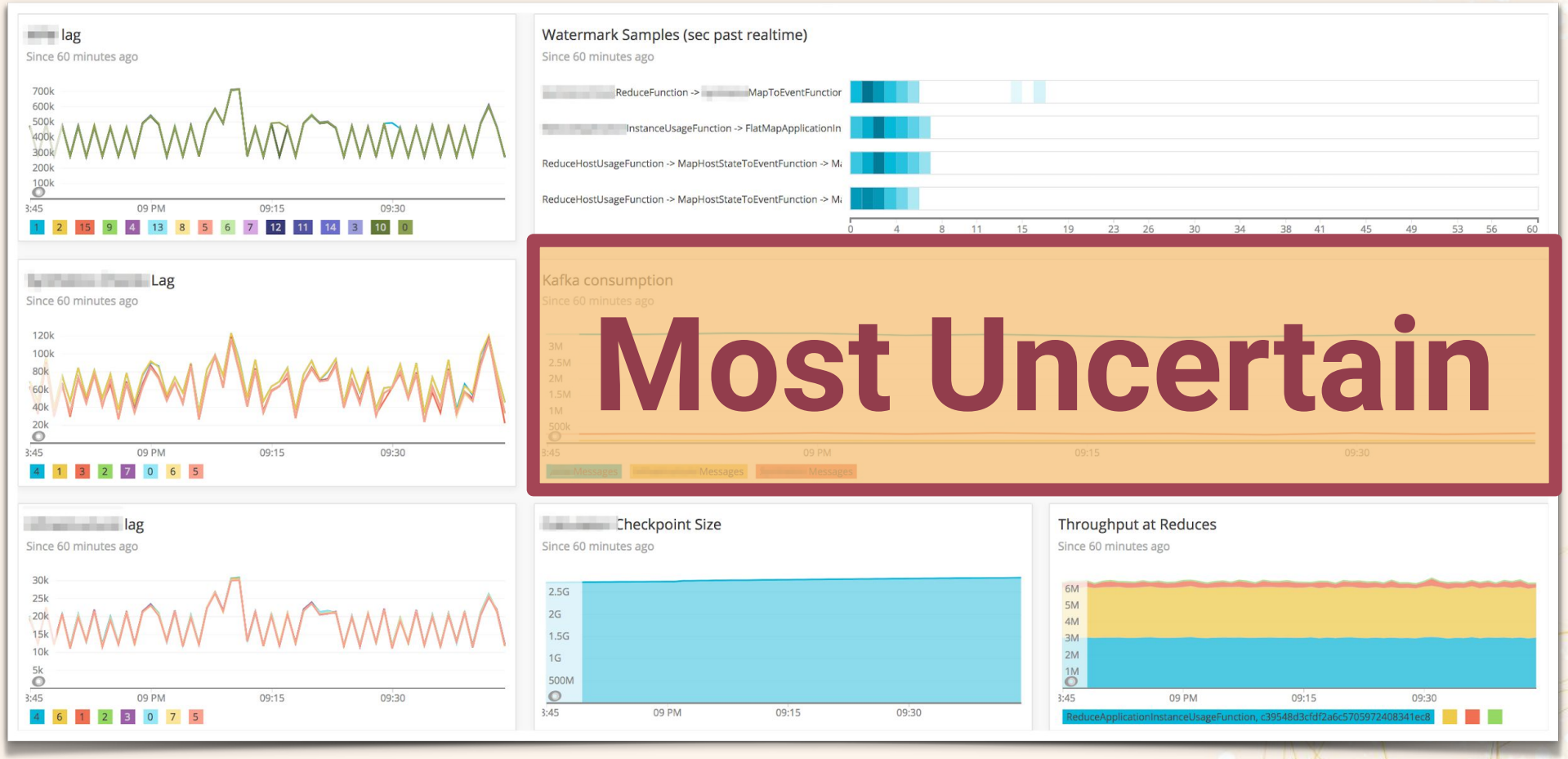
@CAITO_200_OK

Metrics-Driven Metrics



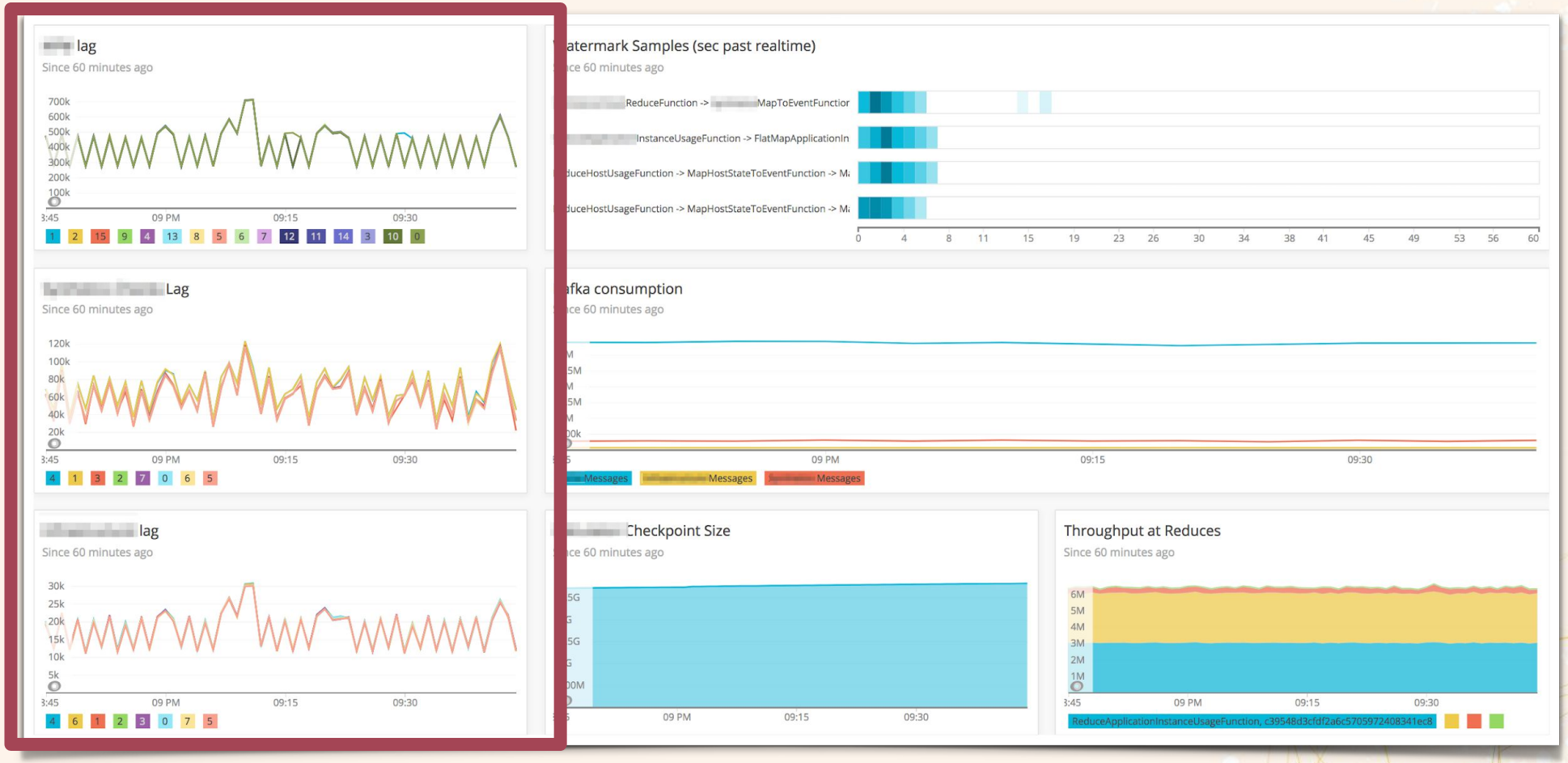
@CAITO_200_OK

Metrics-Driven Metrics



@CAITO_200_OK

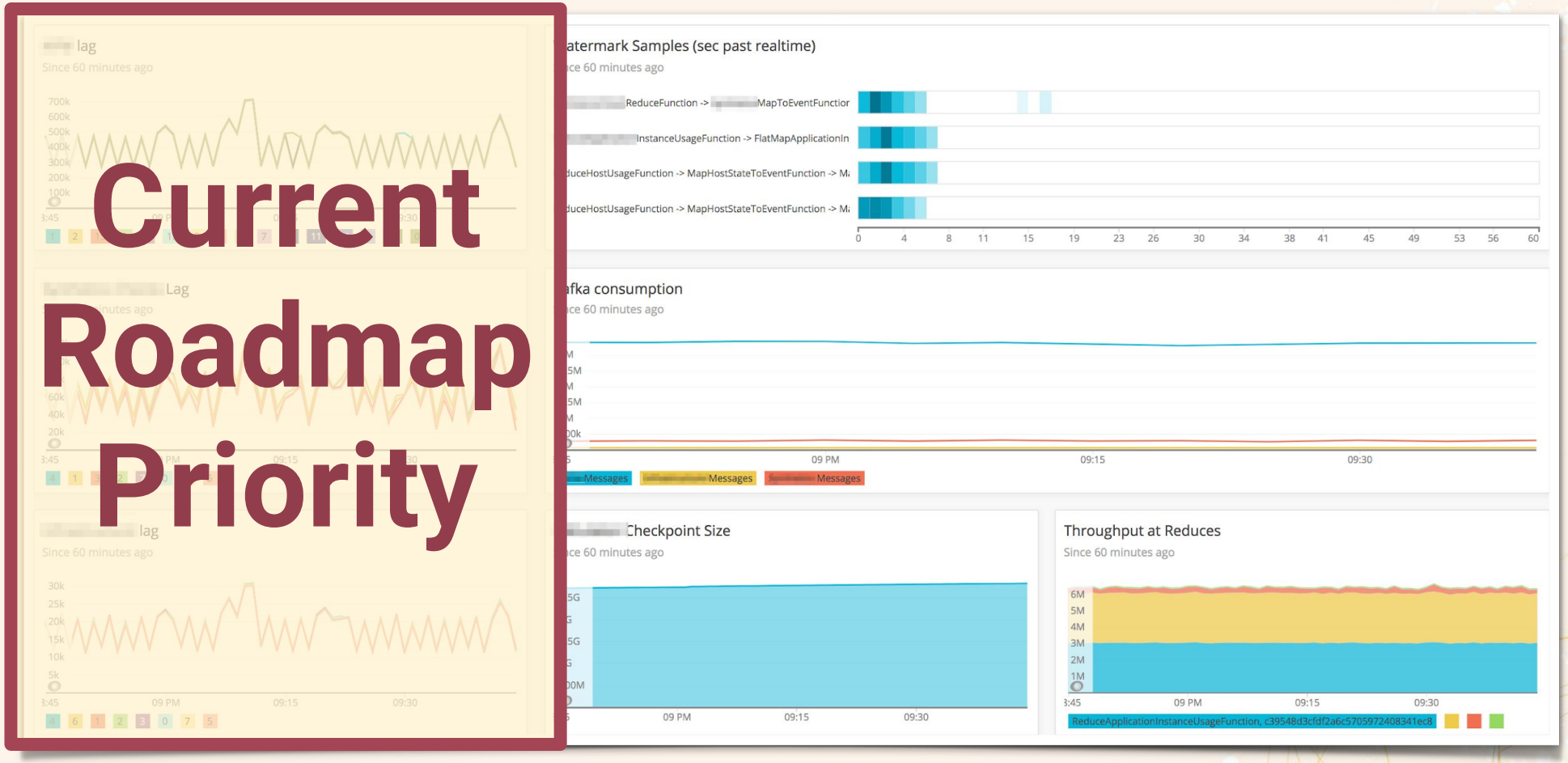
Metrics-Driven Metrics



@CAITO_200_OK

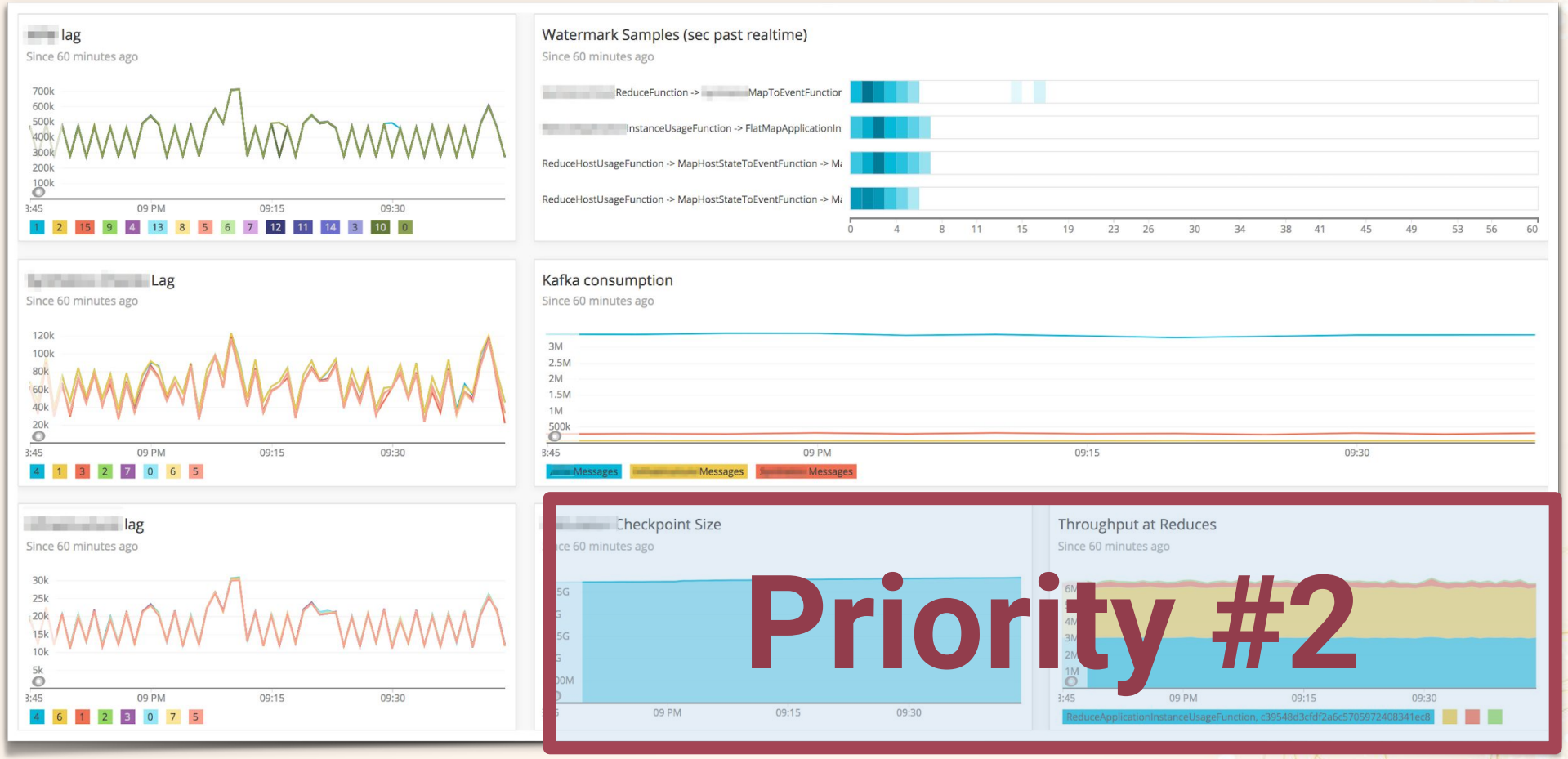
Metrics-Driven Metrics

Current Roadmap Priority



@CAITO_200_OK

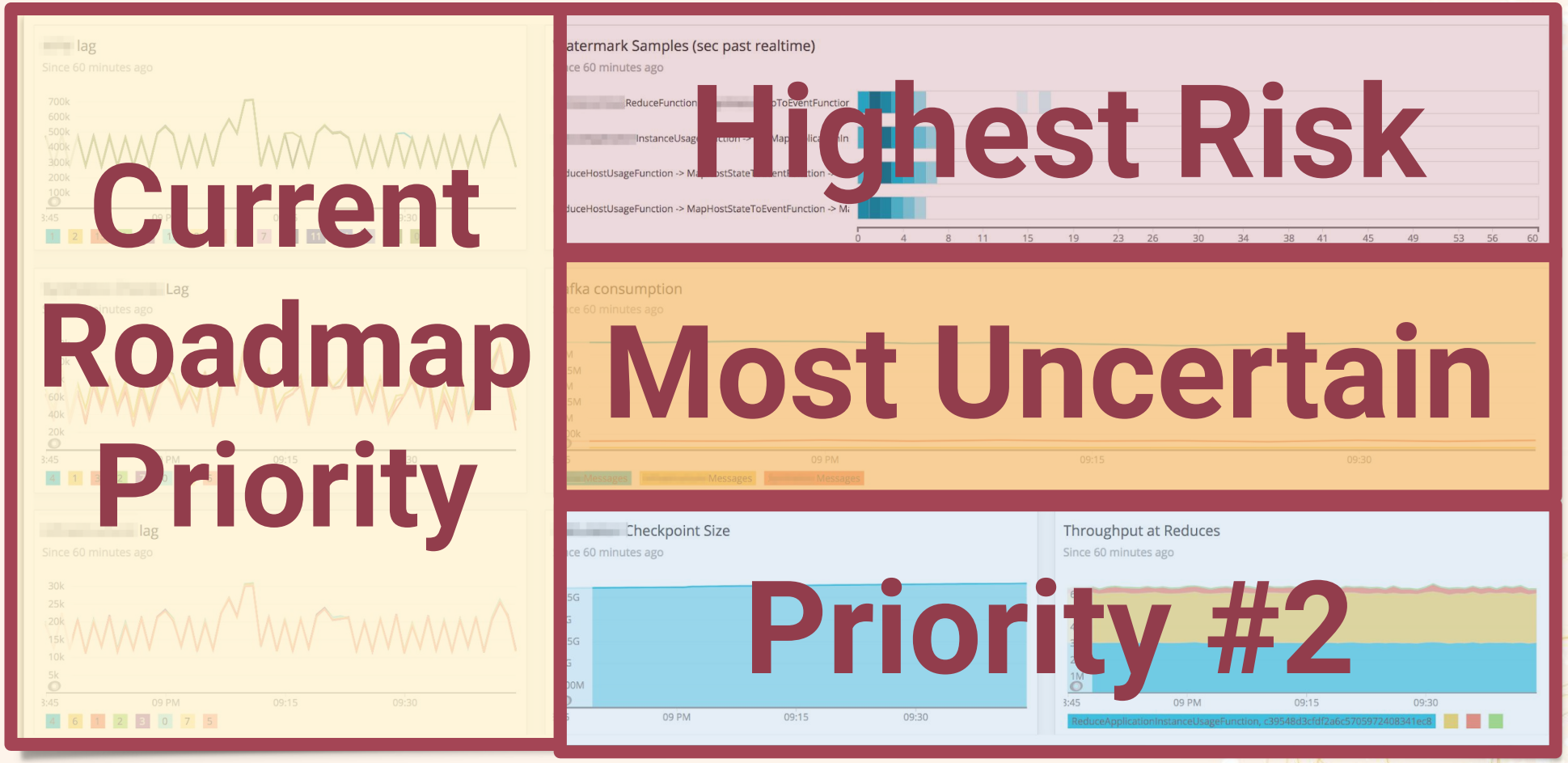
Metrics-Driven Metrics



Priority #2

@CAITO_200_OK

Metrics-Driven Metrics



@CAITO_200_OK

Metrics as a Shared Language

@CAITO_200_OK

Shared Language

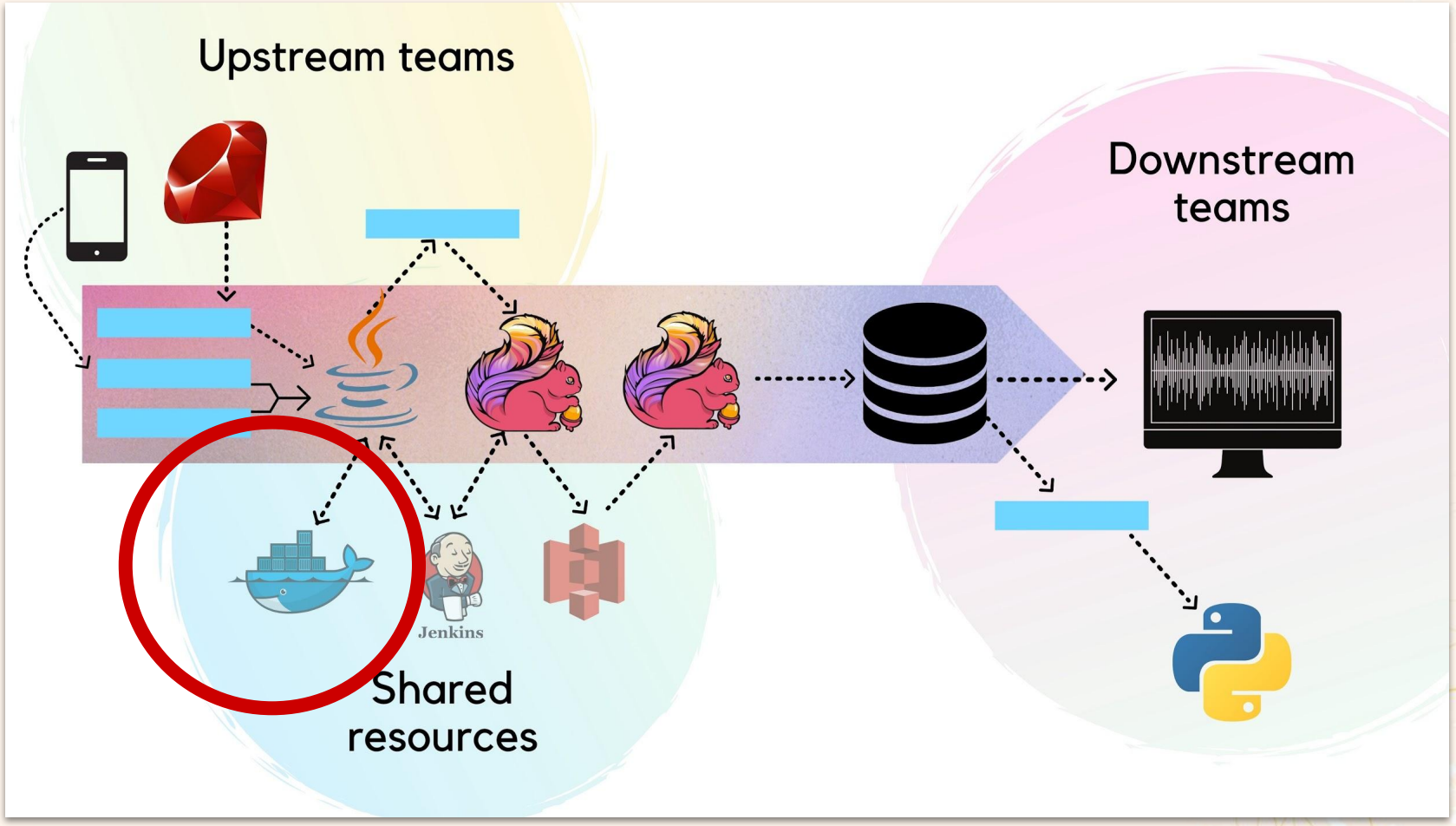
"Even the best metrics driven development fails easily, when it's only built for yourself"

Shared Language

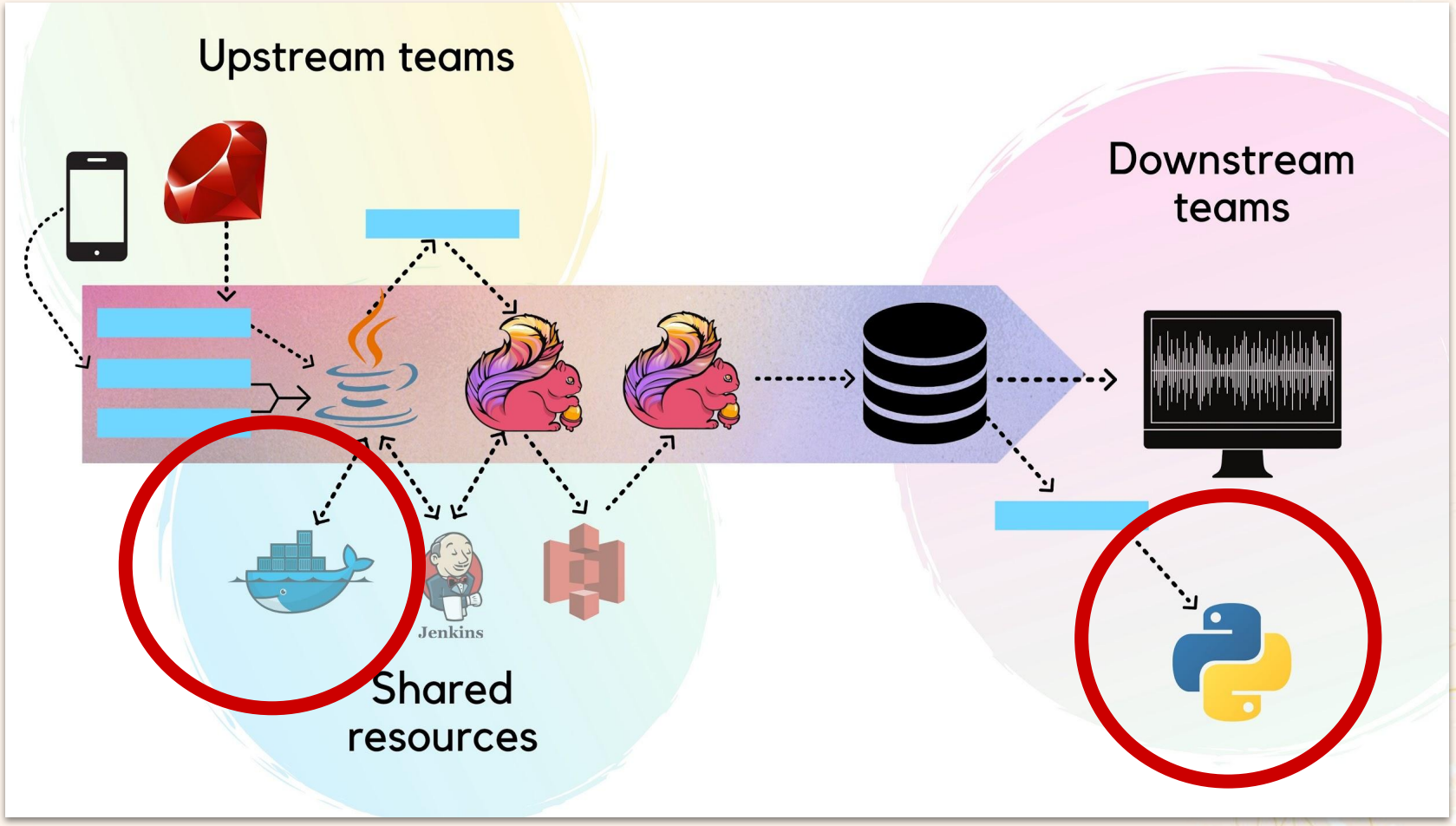
Implementation:

- Identify your impacted groups
- Identify the most effective tools for each group
- Enable automation

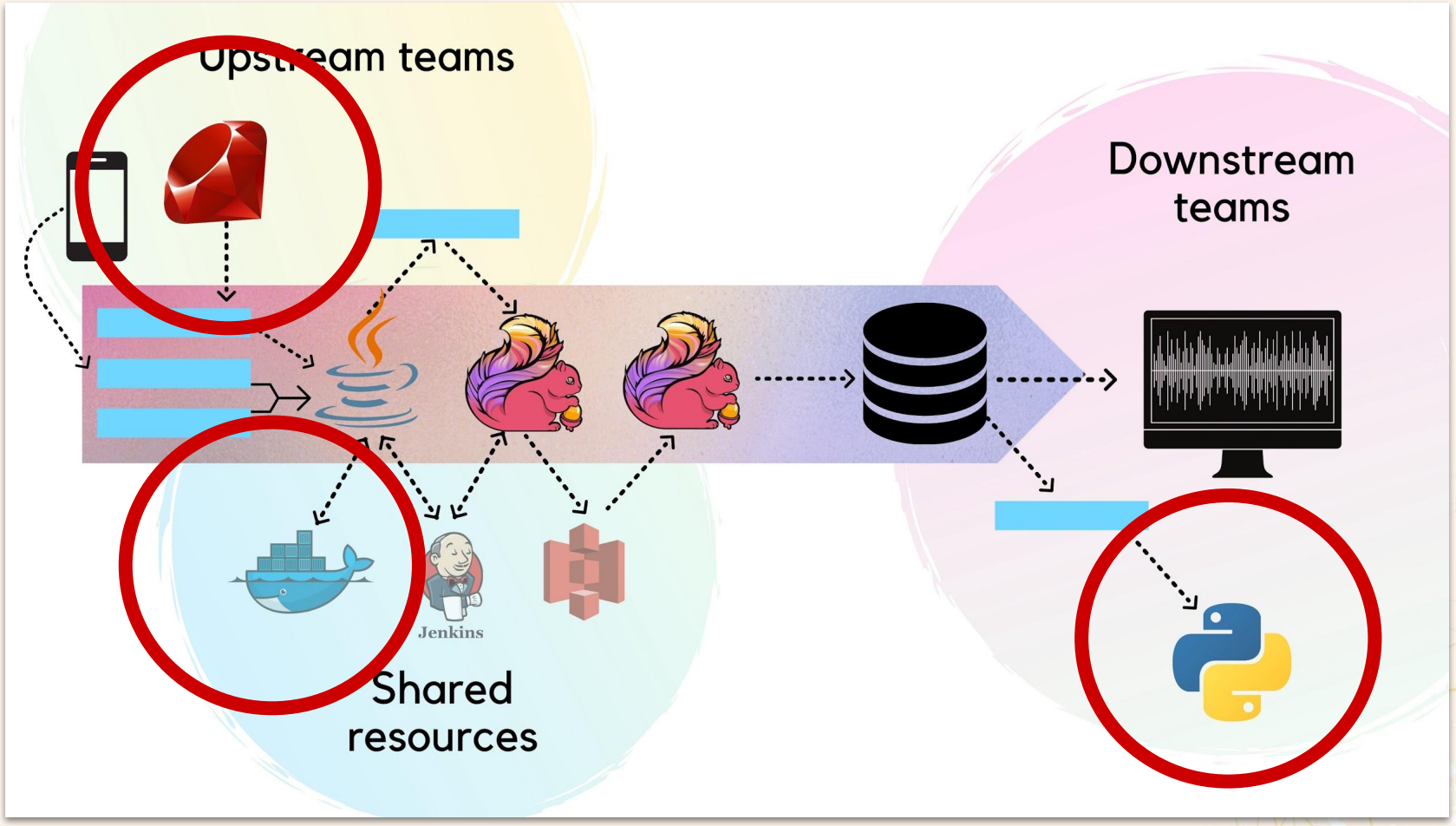
Shared Language



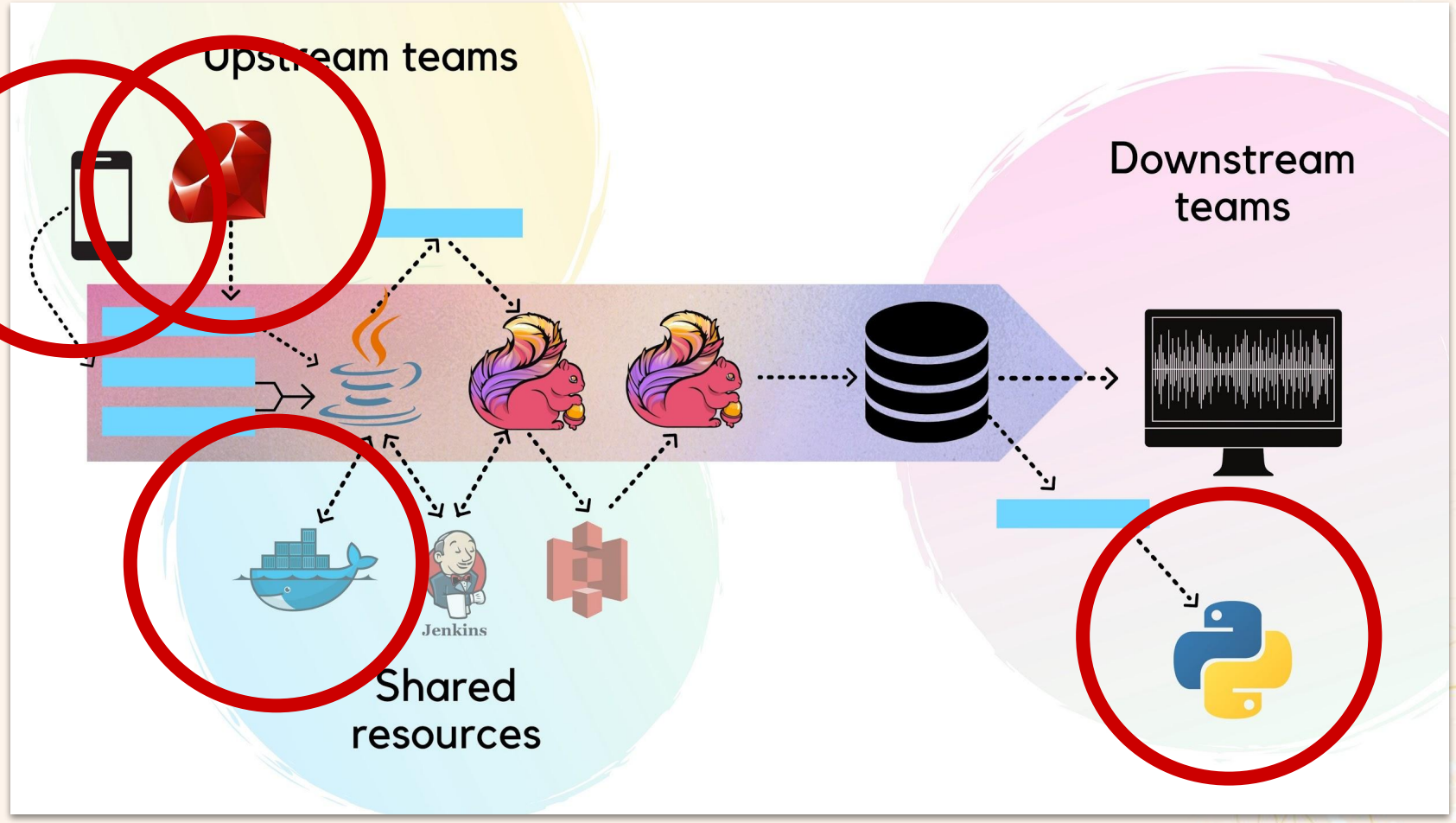
Shared Language



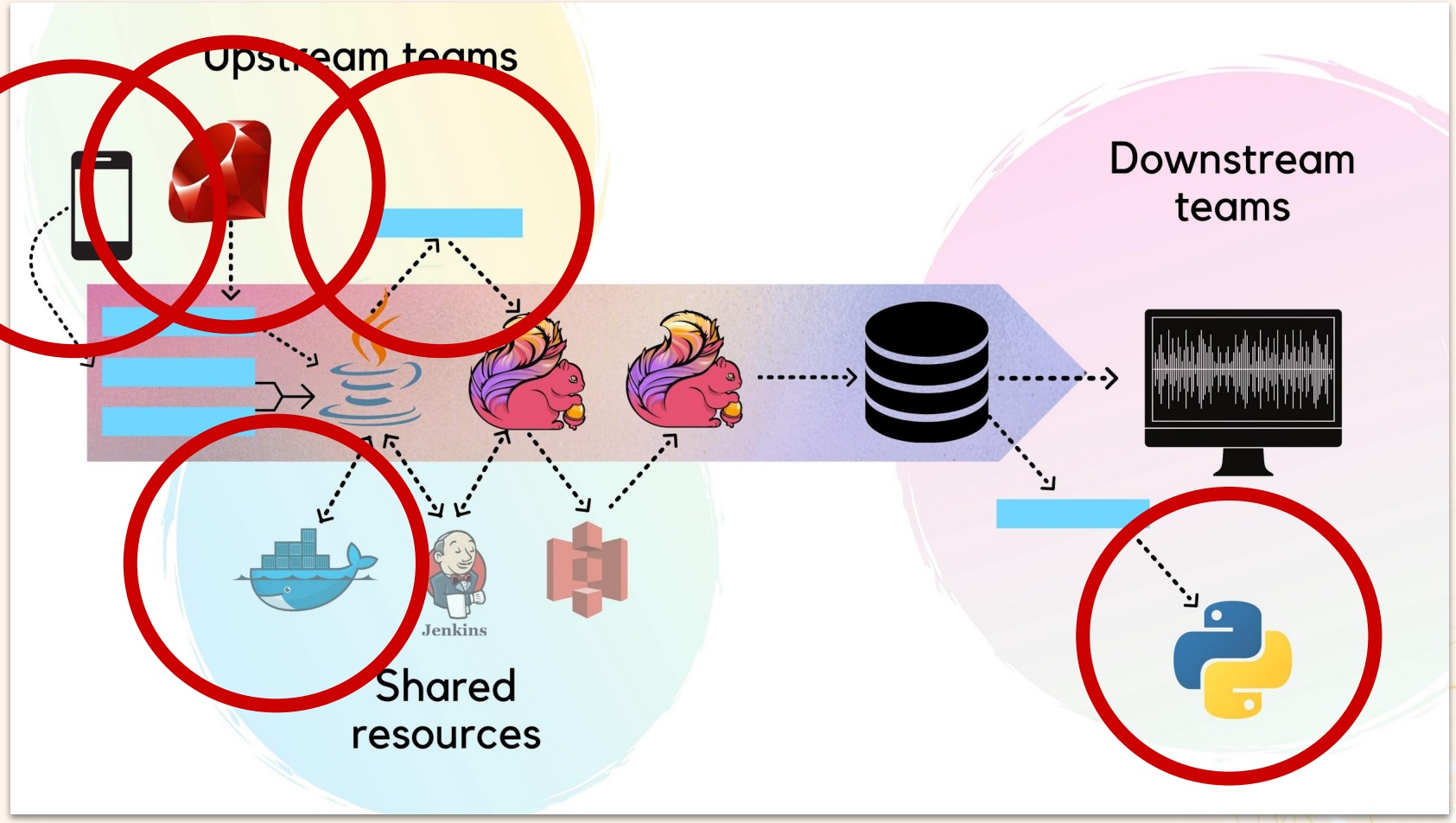
Shared Language



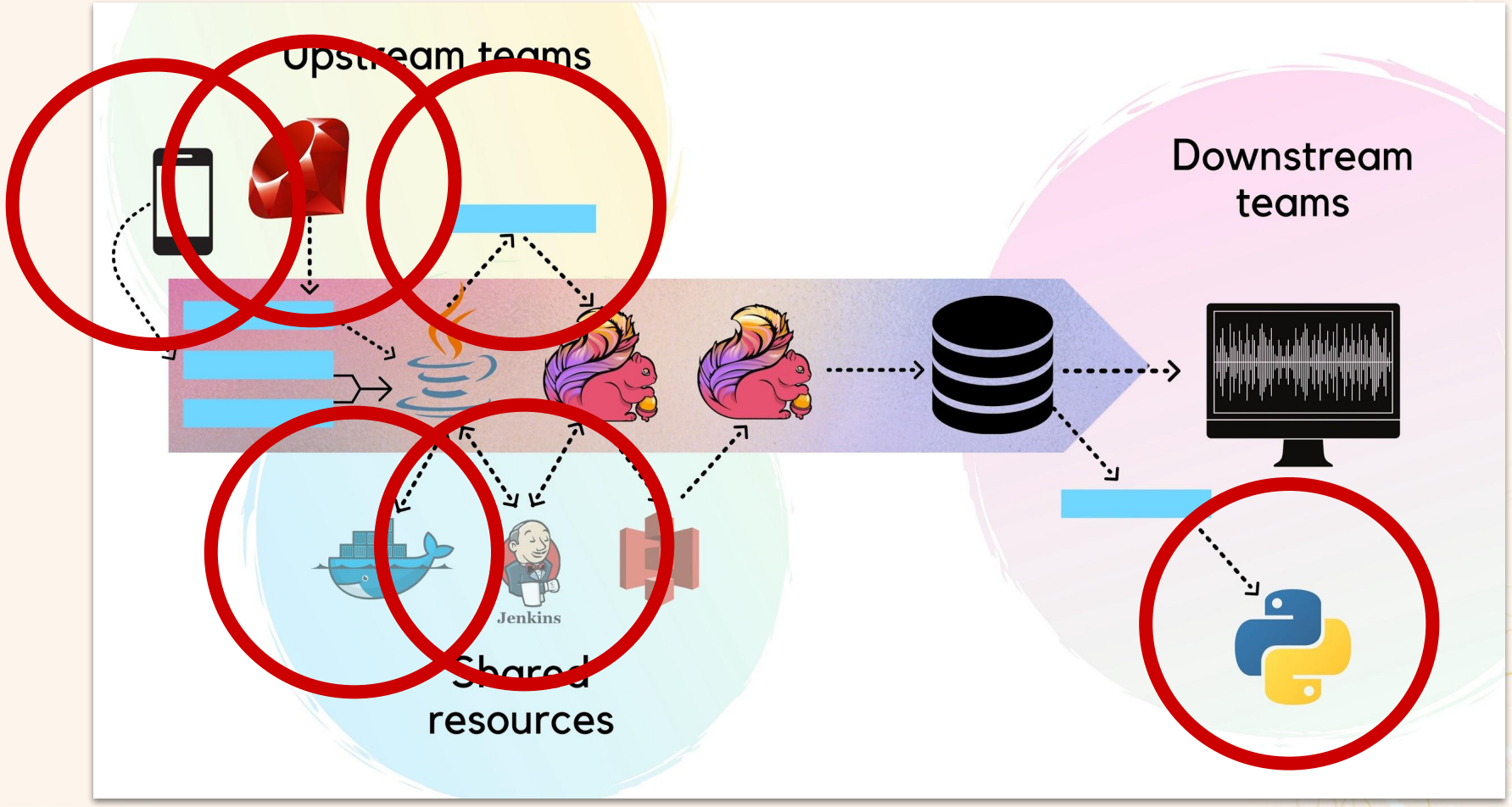
Shared Language



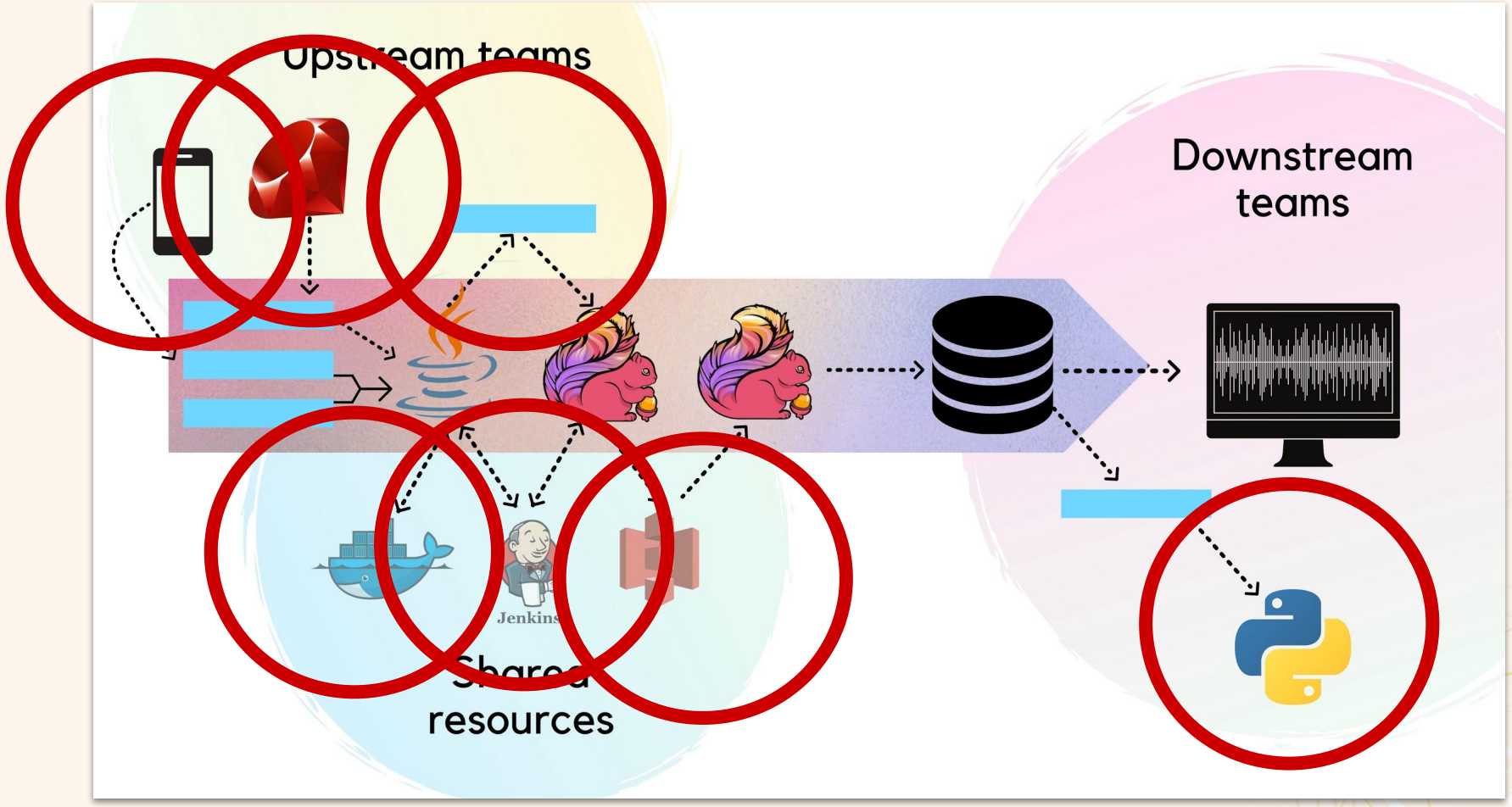
Shared Language



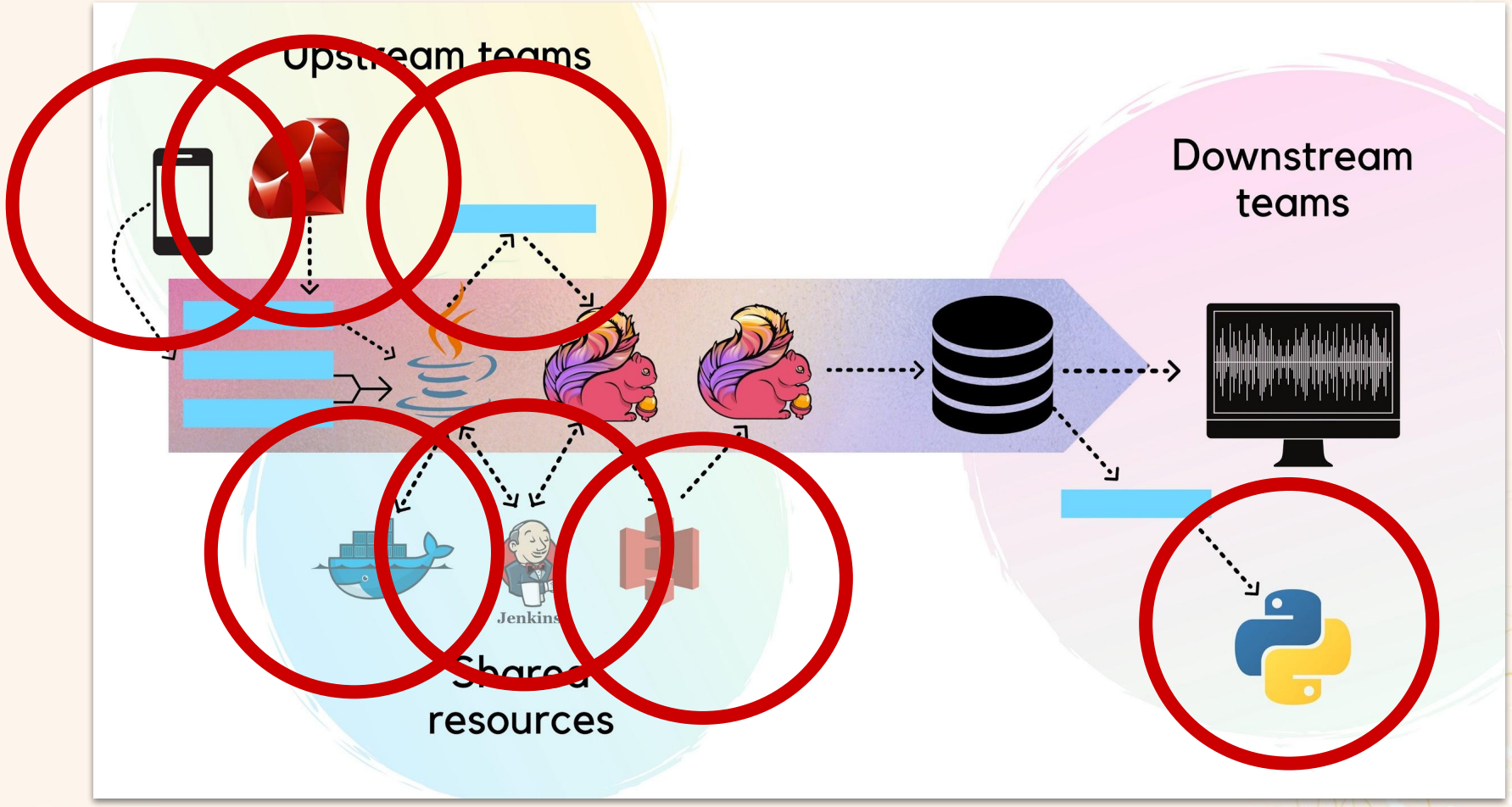
Shared Language



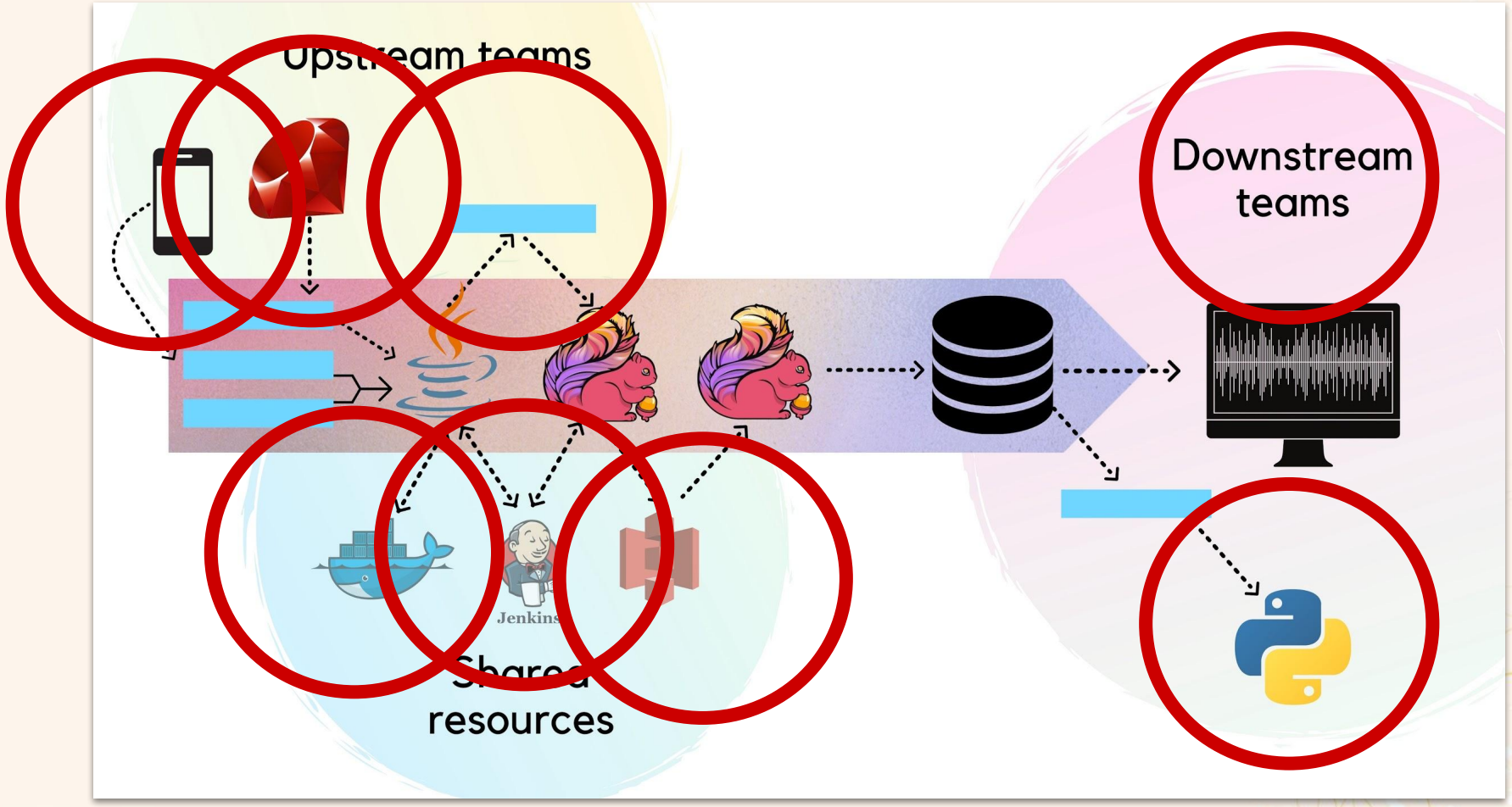
Shared Language



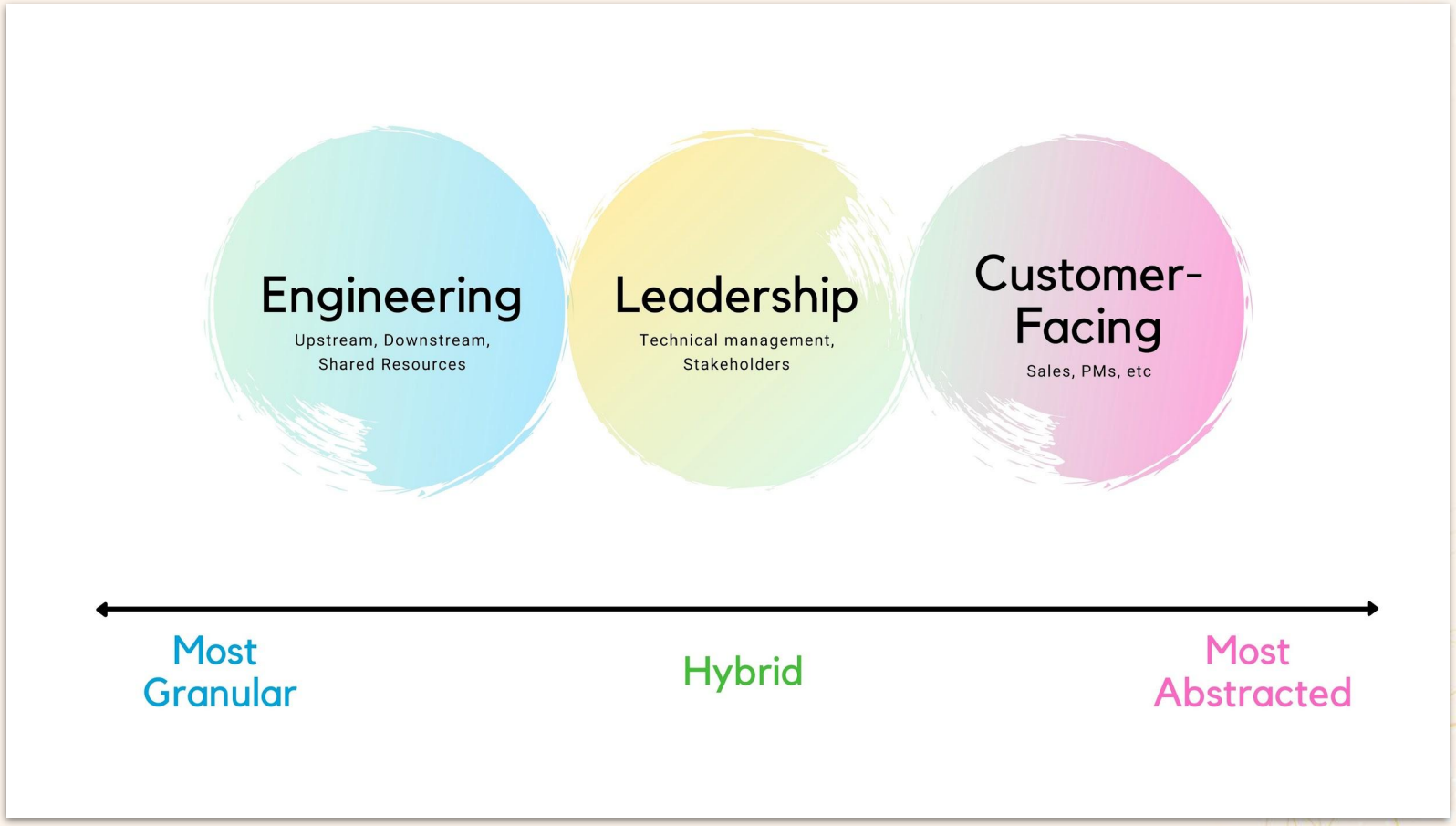
Shared Language



Shared Language

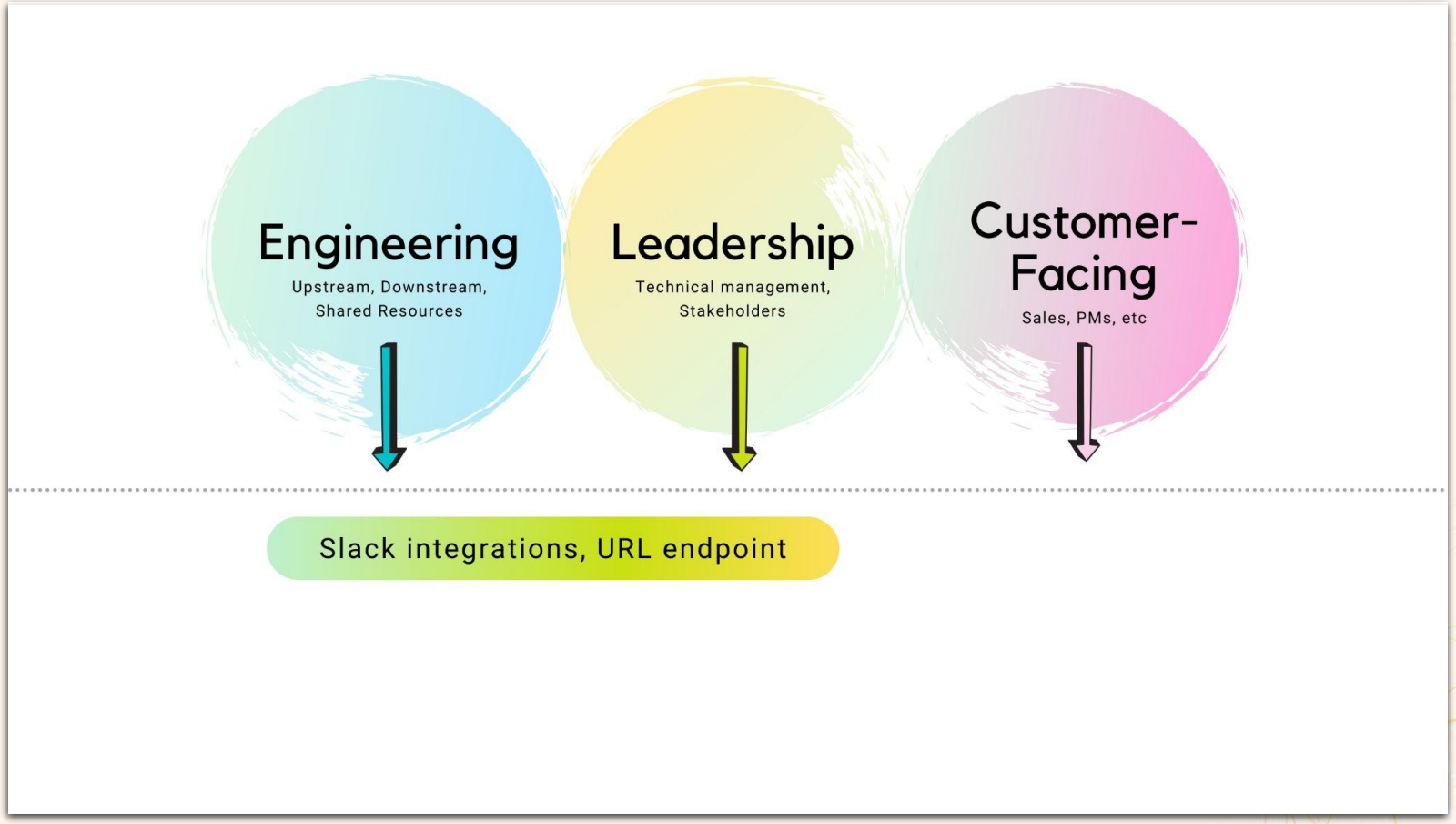


Shared Language

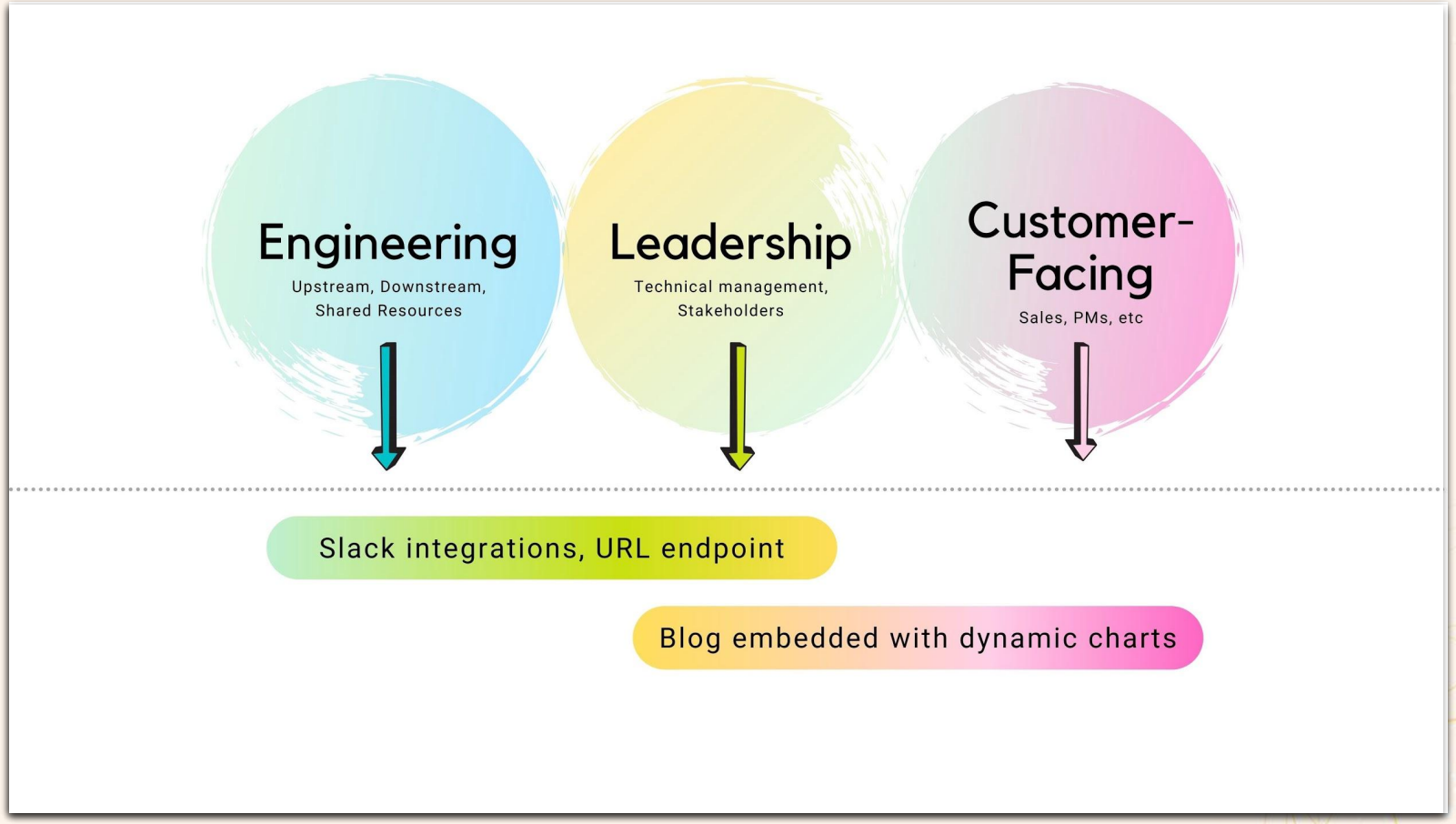


@CAITO_200_OK

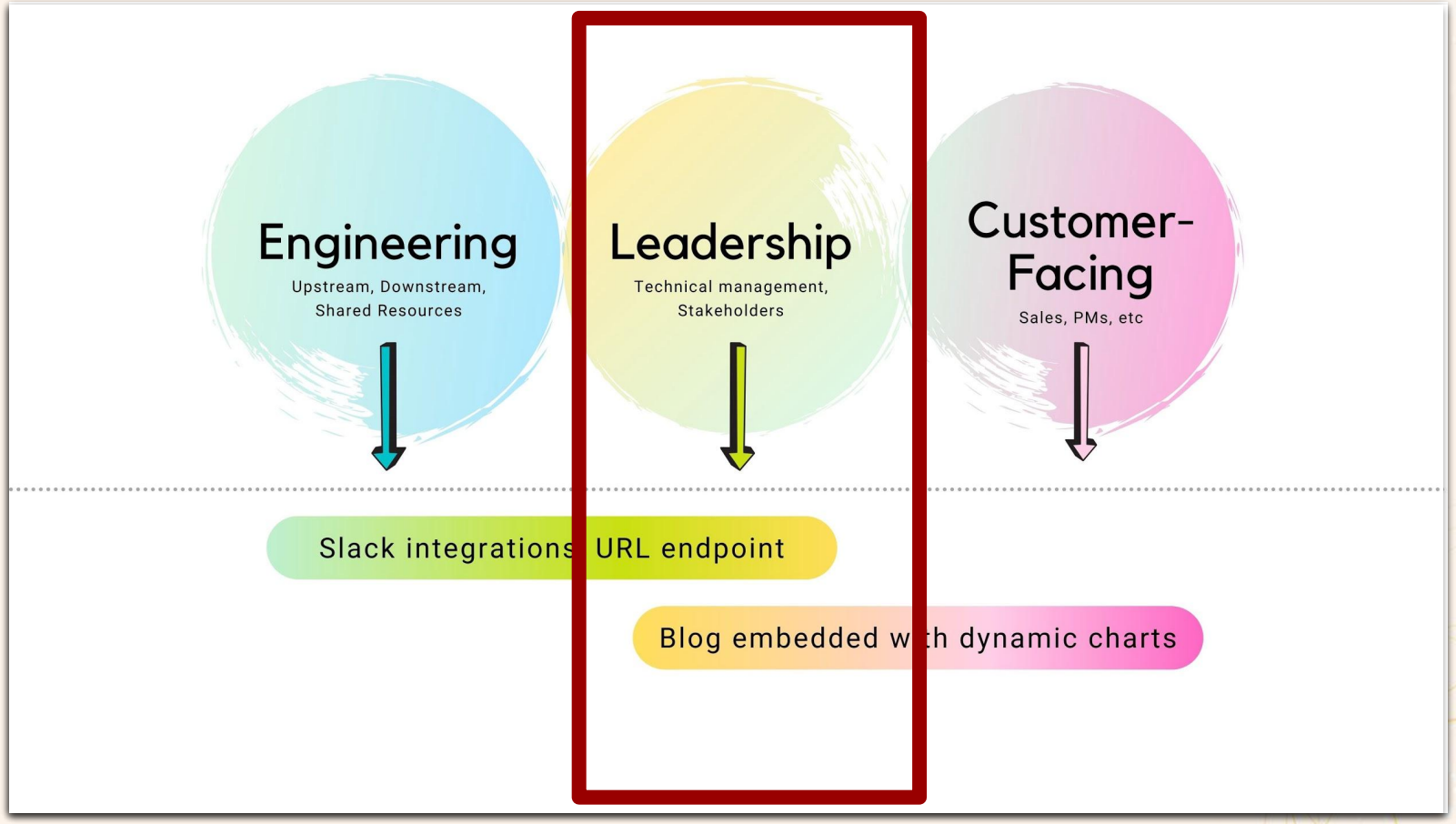
Shared Language



Shared Language



Shared Language



Conclusion

@CAITO_200_OK

Conclusion

For systems with complex integration points:

- **Metrics-Driven Metrics**: streamline technical complexity
- **Metrics as a Shared Language**: streamline interpersonal complications

Introduction

Twitter

Caito_200_OK

Content

https://medium.com/@caito

http://caito-200-ok.com/

Email

Caito@ververica.com

Thank You!

- **YOW! Conference staff!!**

- “Shared Language” photo models:
  - **Zack Hobson**, CTO
  - **Cory Johannsen**, Senior Software Engineer

Quoted Contributors & Reviewers:

- **Mandy Riso** - DevOps Engineer
- **Eric Shamow** - Engineering Manager
- **Mike Hix** - Senior Software Engineer
- **Ben Ford** - Product Manager
- **Noçnica Fee** - Developer Advocate
- **Lucy Wyman** - Senior Software Engineer
- **Logan Ballard** - Site Reliability Engineer

Resources

- Prometheus push gateway - when to use it/docuemntation: https://prometheus.io/docs/practices/pushing/
- Prometheus push gateway on GitHub: https://github.com/prometheus/pushgateway
- Prometheus Flink example: https://github.com/mbode/flink-prometheus-example
- Flink monitoring: https://flink.apache.org/news/2019/02/25/monitoring-best-practices.html

Credits - Images

- Flink logos:
  https://wints.github.io/flink-web//community.html
- Data-Driven/Aware/Informed Design:
  https://uxdesign.cc/becoming-a-data-aware-designer-1d7614ebc3ed
- Stream processing diagram:
  https://www.ververica.com/what-is-stream-processing
- 747 Airplane cockpit:
  https://www.reddit.com/r/pics/comments/5vv8qt/the_pilots_seat_and_cockpit_of_a_boeing_747/
- DataDog Flink dashboard:
  https://www.datadoghq.com/blog/monitor-apache-flink-with-datadog/
- All other photos & Images:
  Caito Scherr

Credits

- Basic Data-Driven Development principles https://www.portable.com.au/reports/principles-of-data-driven-design
- Data-driven design: https://www.springboard.com/blog/data-driven-design/
- Becoming A Data-Aware Designer, + definition image - Illustration from “Designing with Data” by King, Churchill, & Tan https://uxdesign.cc/becoming-a-data-aware-designer-1d7614ebc3ed
- https://digitalprinciples.org/wp-content/uploads/PDD_Principle-BeDataDriven_v2.pdf
- Harvard Business Review - Data Driven Culture: https://hbr.org/2020/02/10-steps-to-creating-a-data-driven-culture

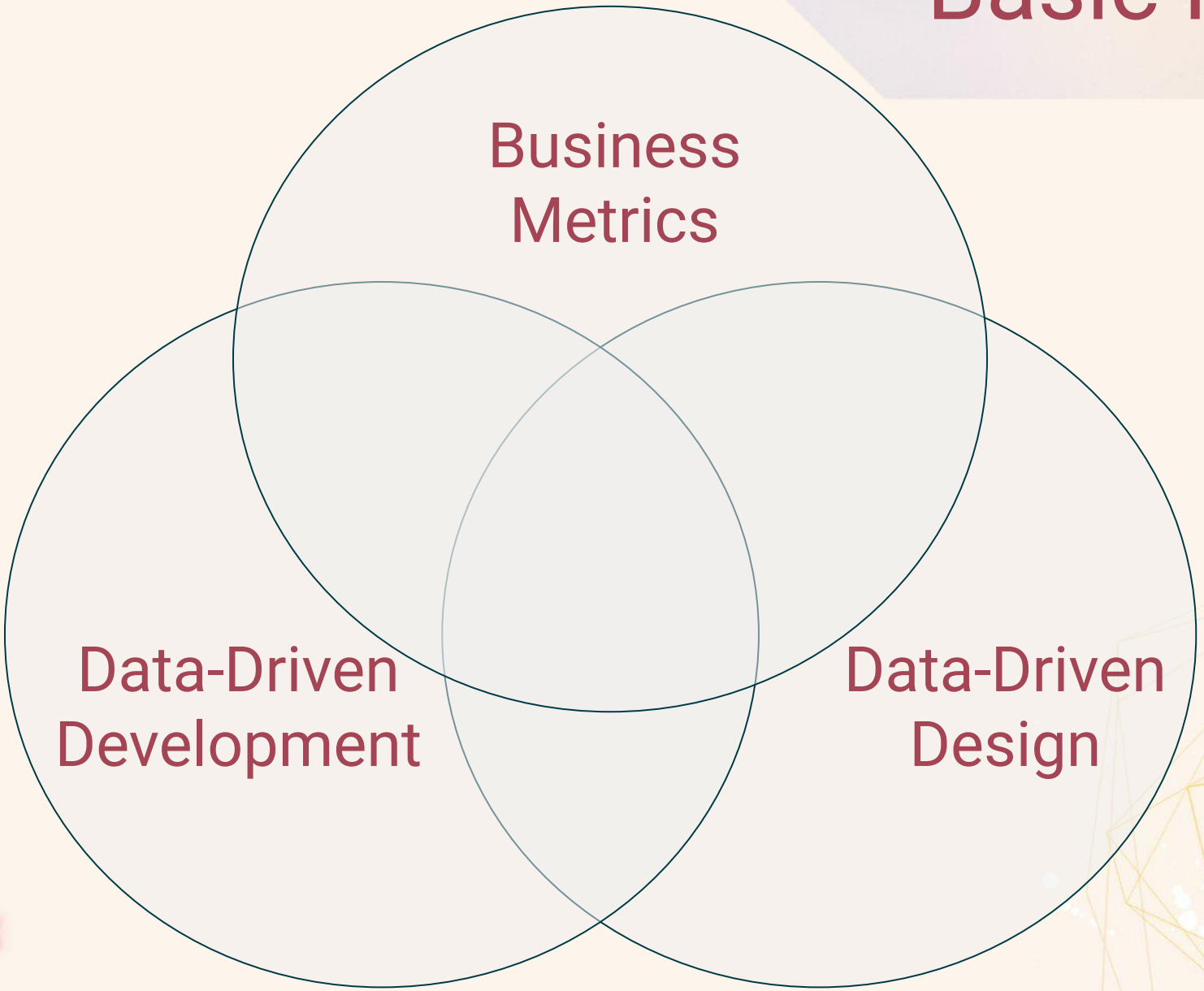
ververica

@CAITO-200_OK

Basic Principles

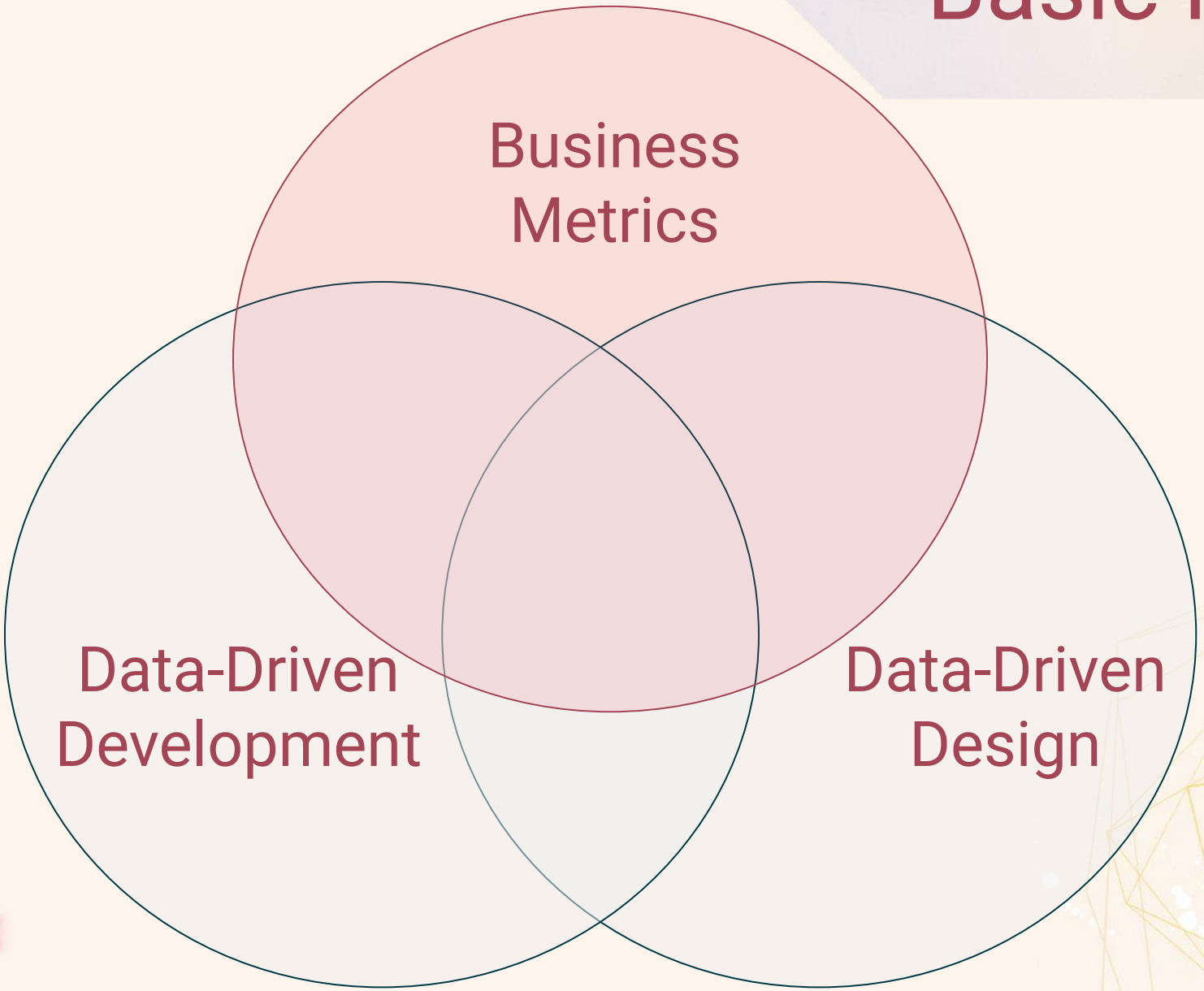
@CAITO_200_OK

Basic Principles

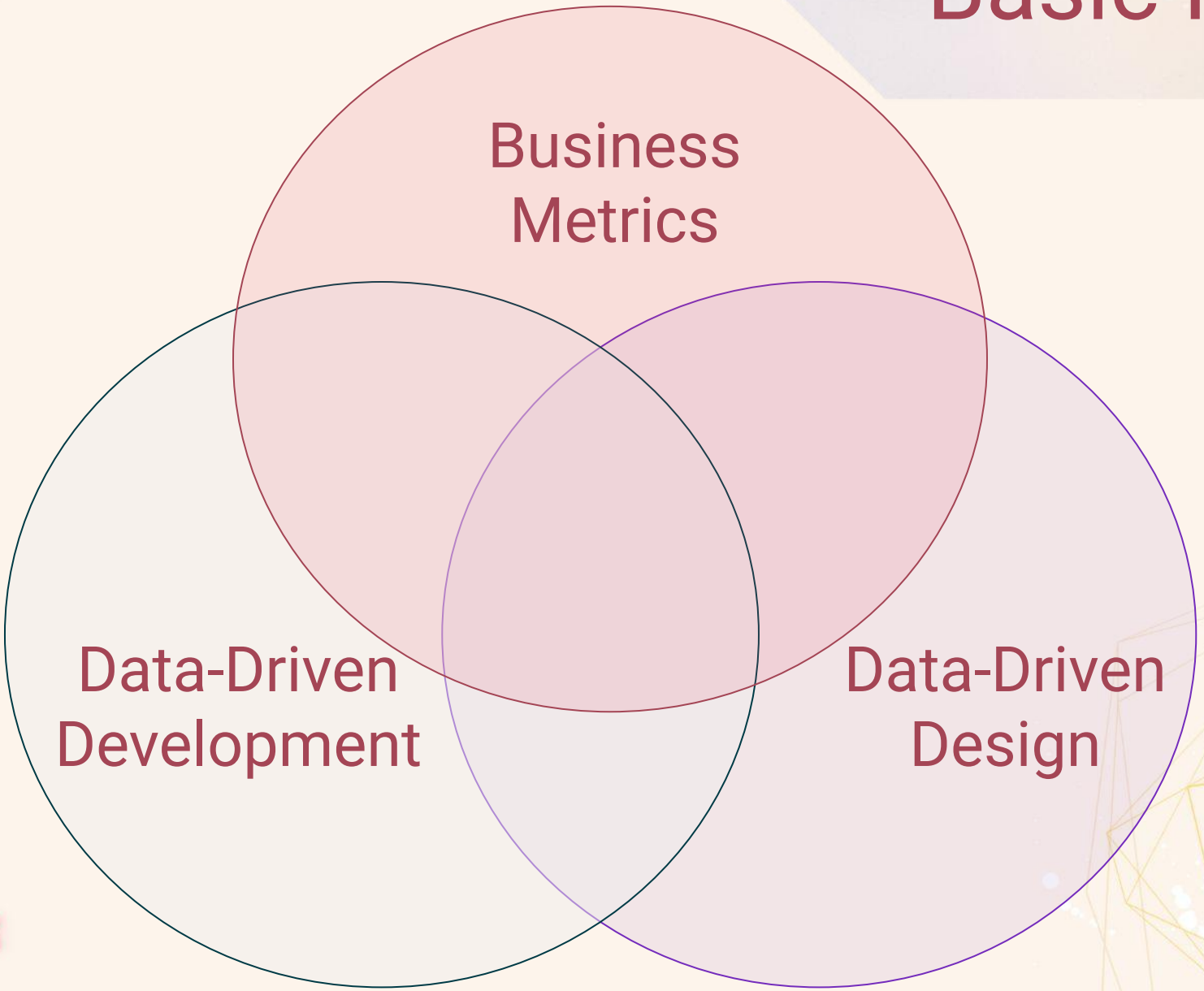


@CAITO_200_OK

Basic Principles

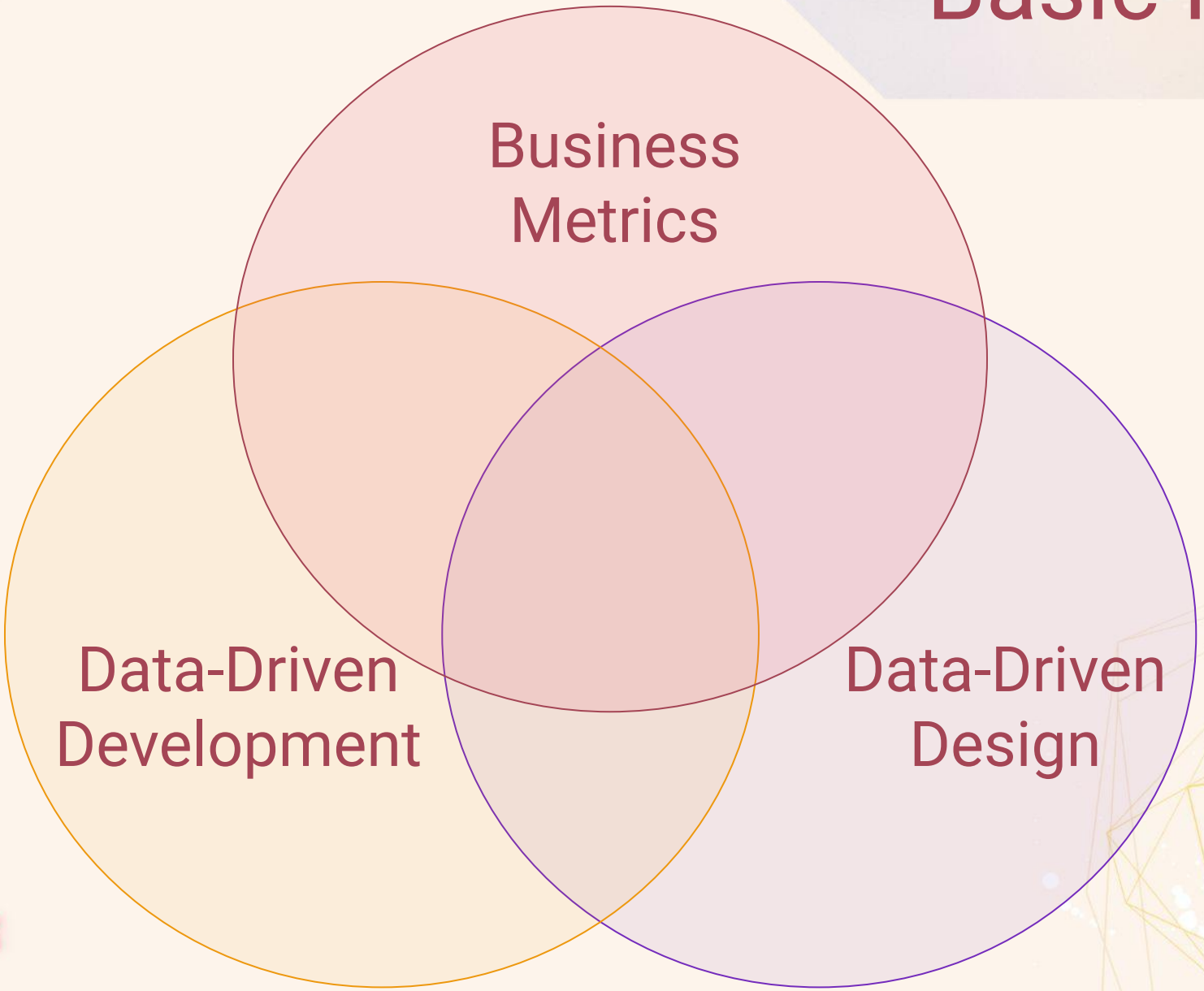


Basic Principles



@CAITO_200_OK

Basic Principles



CONCLUSION

Audience	Tactic	Tools
You/Your Team	• Dashboard should represent your priorities & values • Bring Development + Operations together	**Analytics platforms** • Flink UI, Ververica Platform • Data Dog • New Relic • Grafana/Prometheus...
Upstream + Downstream Teams	• Understand each other's "normal" • Clear ownership for integration points	**Analytics + chat integrations** • Slack + PagerDuty • Slack + Jenkins...
Leadership + Stakeholders	• Business metrics • Pivot between summary & depth • Use tools they're familiar with	**Automated, non-engineering spaces** • URL endpoints • Internal wiki/blog (with embedded, automated output)
Customers	• More manual approach	**Manual approach** • Your company's customer-facing teams

BASIC PRINCIPLES

Quantitative **Qualitative**

BASIC PRINCIPLES

Data-Driven Design

Quantitative **Qualitative**

BASIC PRINCIPLES

Data-Driven Design

Data-Informed Design

Quantitative

Qualitative

BASIC PRINCIPLES

Data-Driven Design

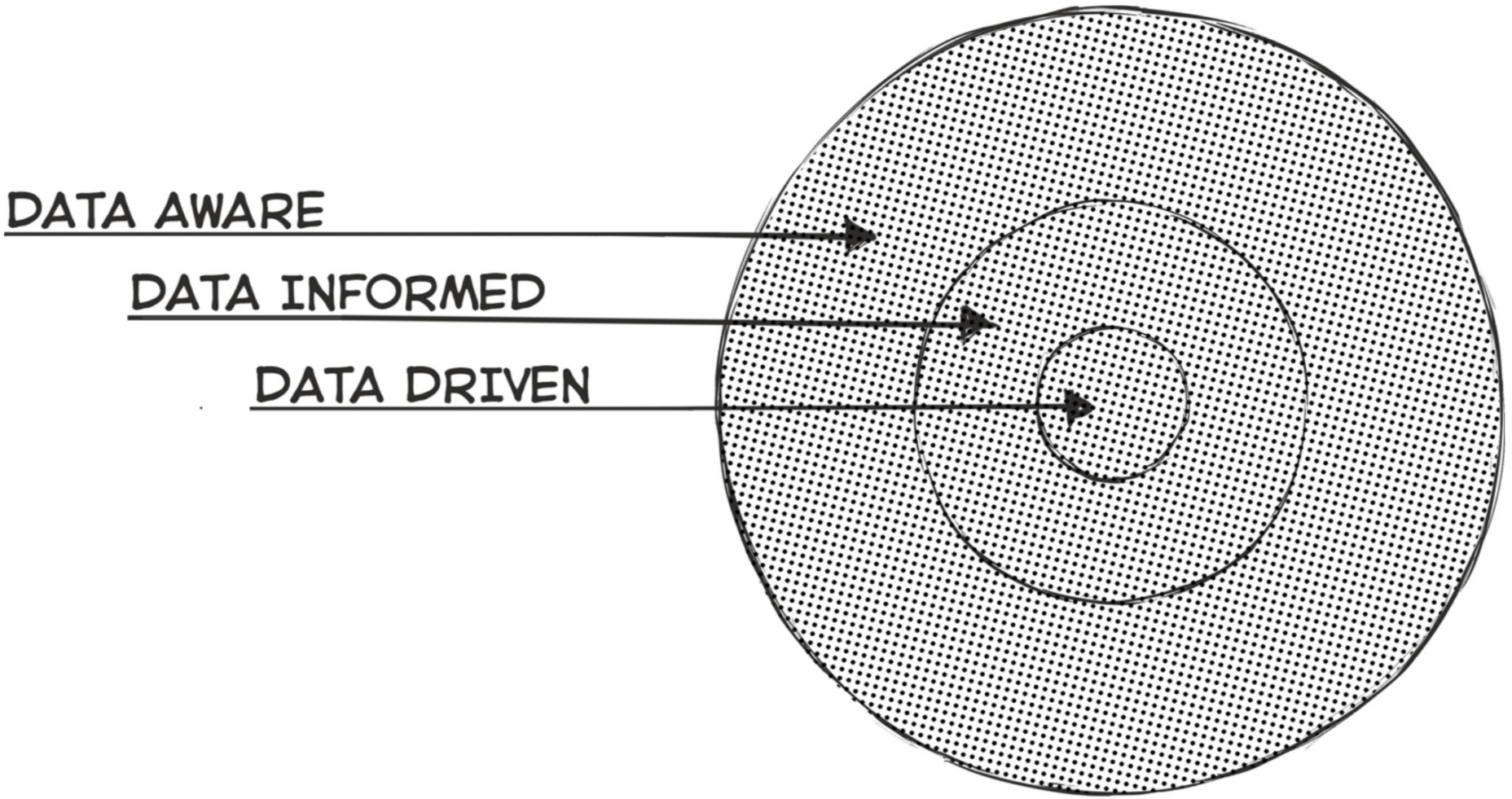
Data-Informed Design

Data-Aware Design

Quantitative

Qualitative

BASIC PRINCIPLES



THE SHARED LANGUAGE OF METRICS

A SHARED LANGUAGE

Even the best data-driven development
fails easily when it's only built for yourself.

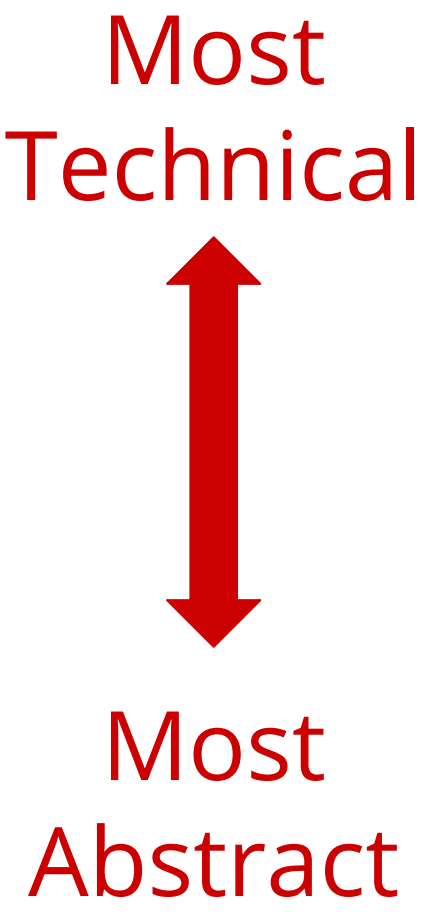
Metrics as a Shared Language

Engineering Team

Up + Downstream Teams

Leadership/Stakeholders

Customers



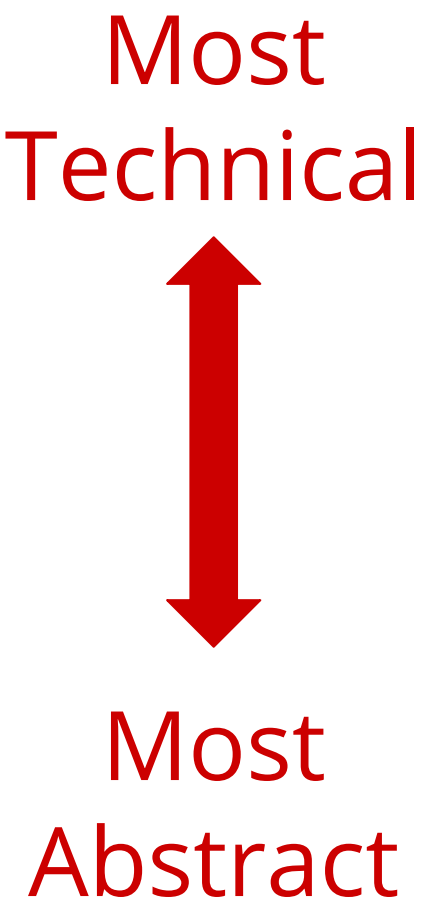
Metrics as a Shared Language

Engineering Team

Up + Downstream Teams

Leadership/Stakeholders

Customers



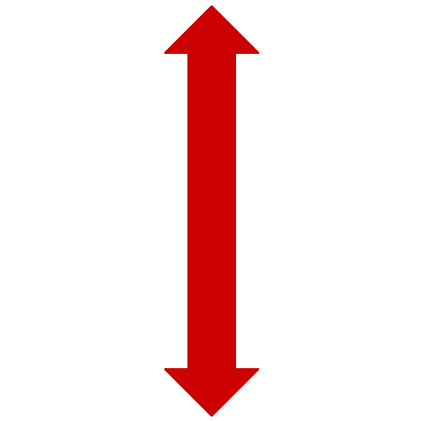
Metrics as a Shared Language

Engineering Team

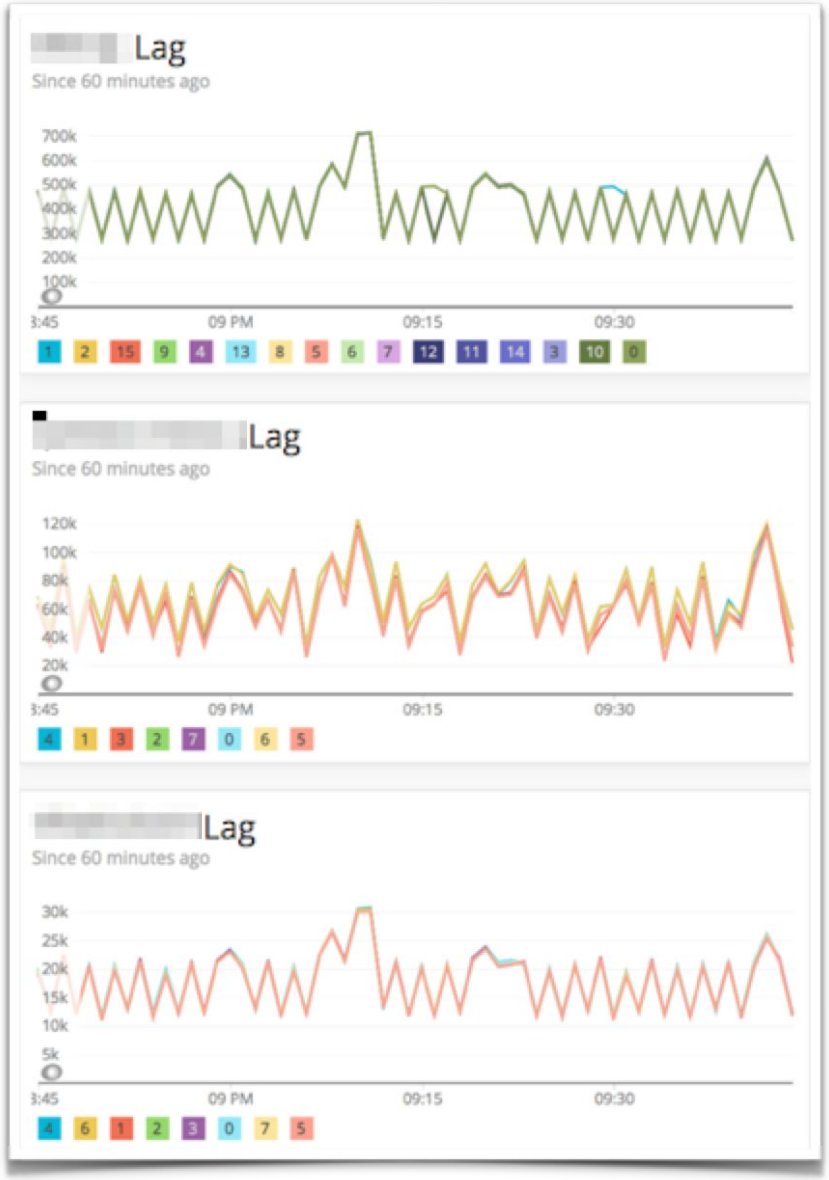
Up + Downstream Teams

Leadership/Stakeholders

Customers



A SHARED LANGUAGE



A SHARED LANGUAGE

A SHARED LANGUAGE

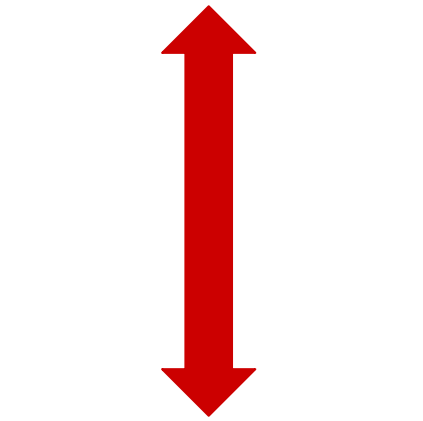
Metrics as a Shared Language

Engineering Team

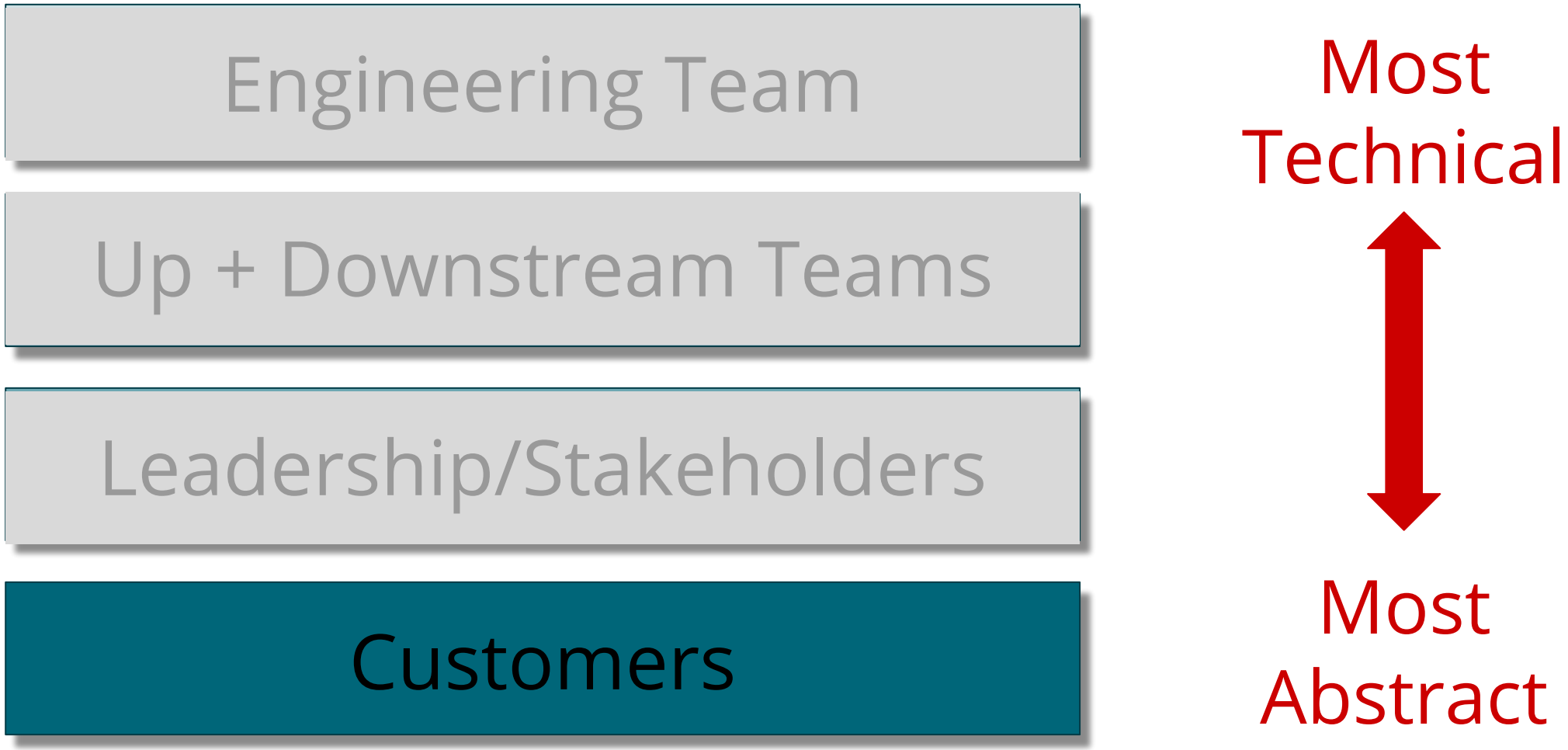
Up + Downstream Teams

Leadership/Stakeholders

Customers



Metrics as a Shared Language



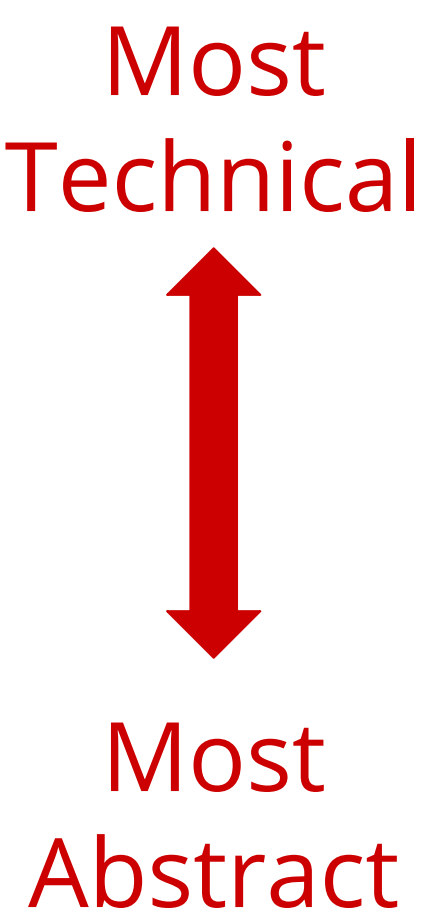
Metrics as a Shared Language

Engineering Team

Up + Downstream Teams

Leadership/Stakeholders

Customers



Summary

Concept	Principles	Implementation	Benefit
Metrics-Driven-Metrics cycle	Meaningful, iterative, accessible	High risk, most uncertain, current roadmap priority	Incident response,
Metrics as a shared language	Use tools familiar to your audience	Small modular units	Automate people & process challenges

Before We Start...

Flink's REST API integration

```
http://hostname:8081/jobmanager/metrics
                    /taskmanagers/<taskmanagerid>/metrics
                    /taskmanagers/metrics
                    /jobs/metrics?jobs=D,E,F
                    ...
```