See the Full Picture:
Integrating Beam/Dataflow into Your Distributed Traces
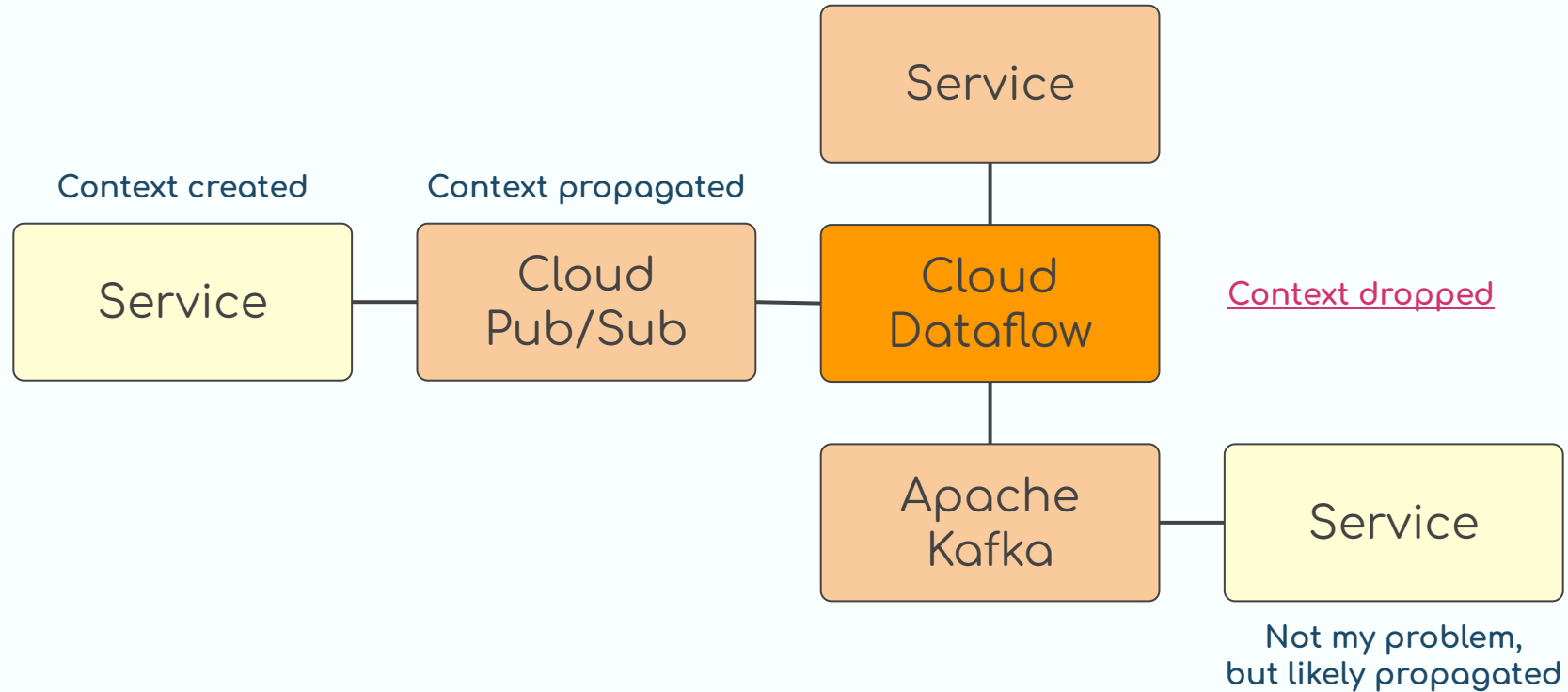
BEAM SUMMIT
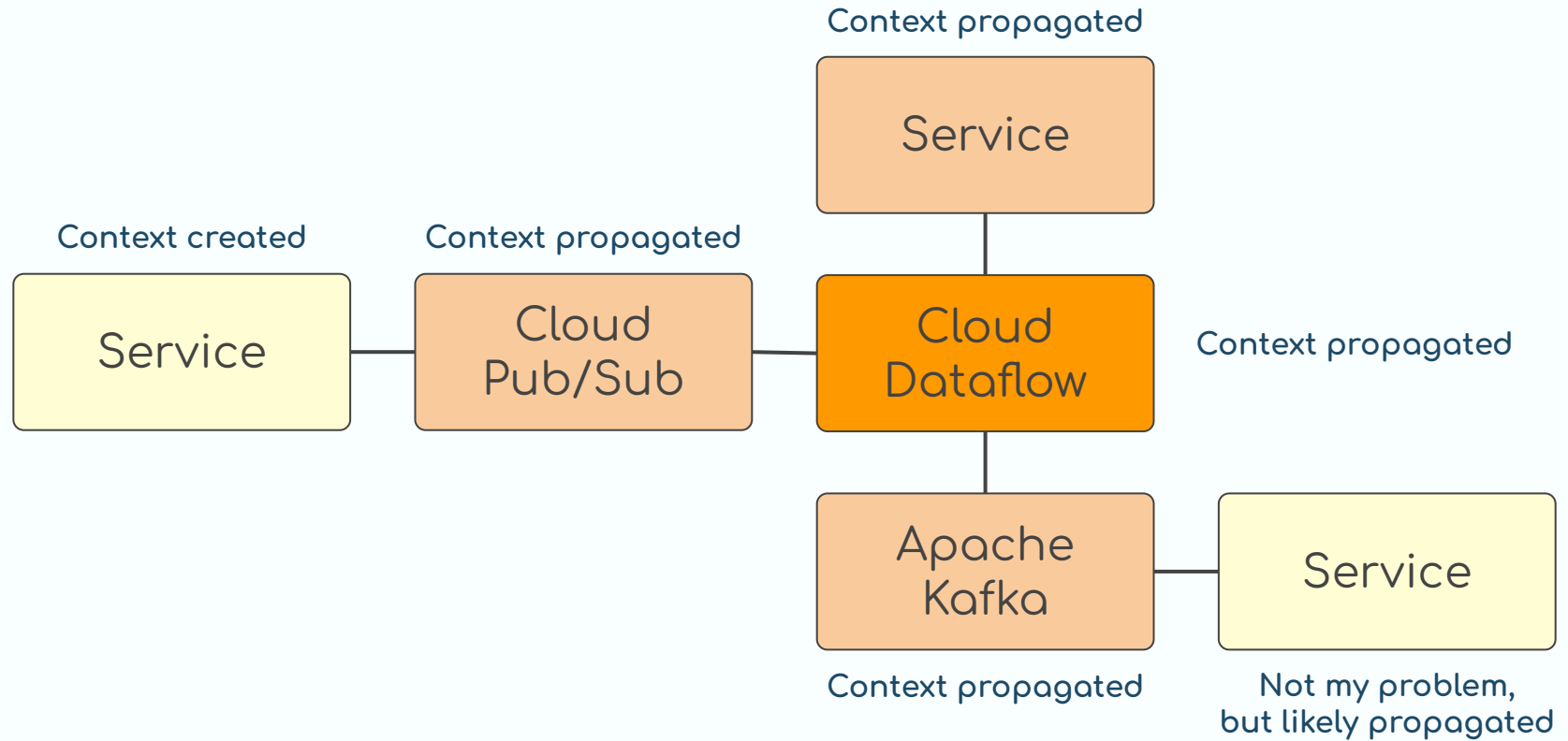NYC 2025

Steven van Rossum
Kenneth Knowles
Radek Stankiewicz

- Trace propagation in Beam runners
- Implementation status
- Demonstration

Trace propagation in Beam runners

Service

Context created

Context propagated

Service — Cloud Pub/Sub — Cloud Dataflow

Context dropped

Apache Kafka — Service

Not my problem, but likely propagated

Context propagated

Service

Context created

Context propagated

Service

Cloud Pub/Sub

Cloud Dataflow

Context propagated

Apache Kafka

Service

Context propagated

Not my problem, but likely propagated

Transforms

- Extract contexts (IO)
- Propagate contexts (IO)
- Create spans
- Add attributes

Runners

- Provide TracerProvider
- Configure instrumentation
- Send to exporter(s)
- Propagate across stages
  - Based on WindowedValue extensions

Integrating OpenTelemetry

- Portability experiment ([#35227](#))
  - enable_open_telemetry_agent
  - Adds -javaagent:… JVM arg
  - Configure via custom image
- Agent loads extensions, e.g.
  - Framework instrumentation
  - Authentication providers
- Agent injects code, e.g.
  - Bytecode for annotations
  - Hooks for frameworks
- Agent initializes global environment
  - Provides GlobalOpenTelemetry for native instrumentation
- Agent acts as exporter for the JVM
  - Forwards to external exporter

JVM in SDK container

OpenTelemetry Java agent
- Exporter
- Extensions
- Injection
- Environment

FnHarness
- DoFn
  - gRPC
  - @WithSpan
- DoFn
  - Tracer
  - Meter

Steven van Rossum

QUESTIONS?

BEAM
SUMMIT
NYC 2025