Let's code a motorsport telemetry processing pipeline

Paolo Patierno

Principal Software Engineer

Red Hat

```
apiVersion: v1
kind: PrincipalSoftwareEngineer
metadata:
  name: Paolo Patierno
  namespace: Red Hat, Messaging & Data Streaming
  labels:
    cncf/maintainer: Strimzi
    eclipse/committer: Vert.x, Hono & Paho
    microsoft/mvp: Azure
  annotations:
    family: dad of two, husband of one
    sports: running, swimming, motogp, vr46, formula1, ferrari, ssc
napoli
    community: cncf napoli, devday
spec:
  replicas: 1
  containers:
  - image: patiernohub.io/paolo:latest
```

@ppatierno

Red Hat

Building an events stream pipeline:

- How to ingest events reliably
- How to integrate with different systems for events ingestion (UDP) and providing output
- How to process events in real time
- How to show useful insights
- How to run and deploy the entire pipeline

What is Apache Kafka

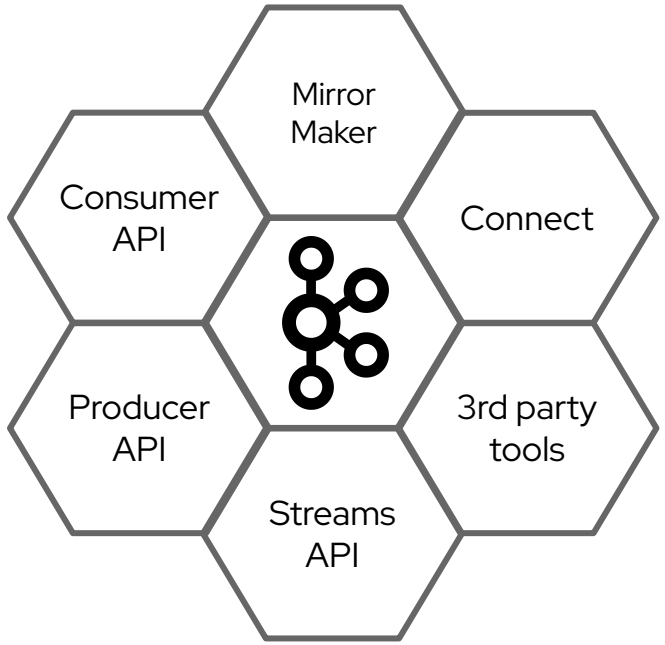
" ... a publish/subscribe messaging system ..."

" ... a streaming data platform ..."

" ... a distributed, horizontally-scalable, fault-tolerant, commit log ..."

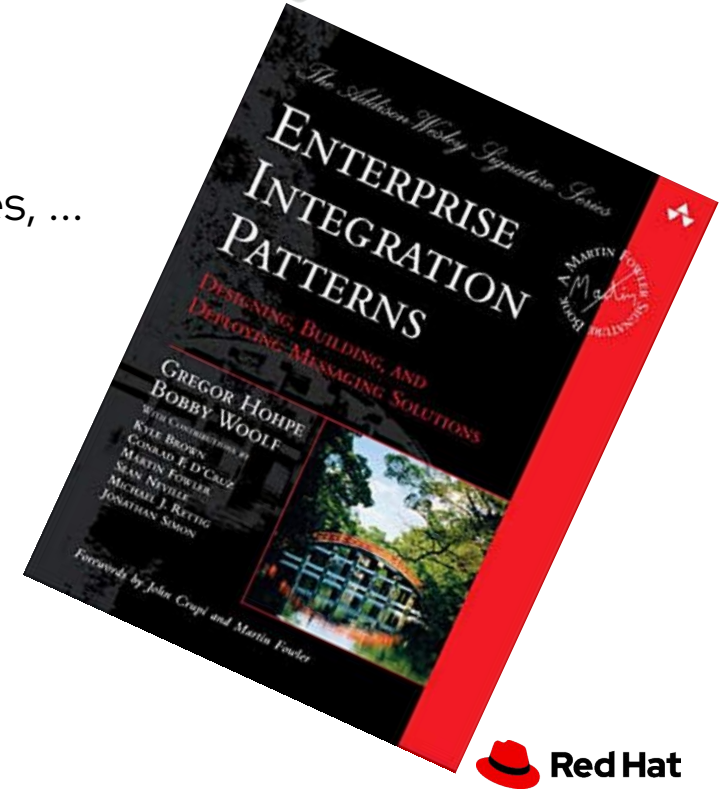
What is Apache Kafka

- Designed to be fast, scalable, durable and available
- Distributed by nature
- Data partitioning (sharding)
- High throughput / low latency
- Broader ecosystem more than just the broker
- Kafka Streams API for real-time event processing



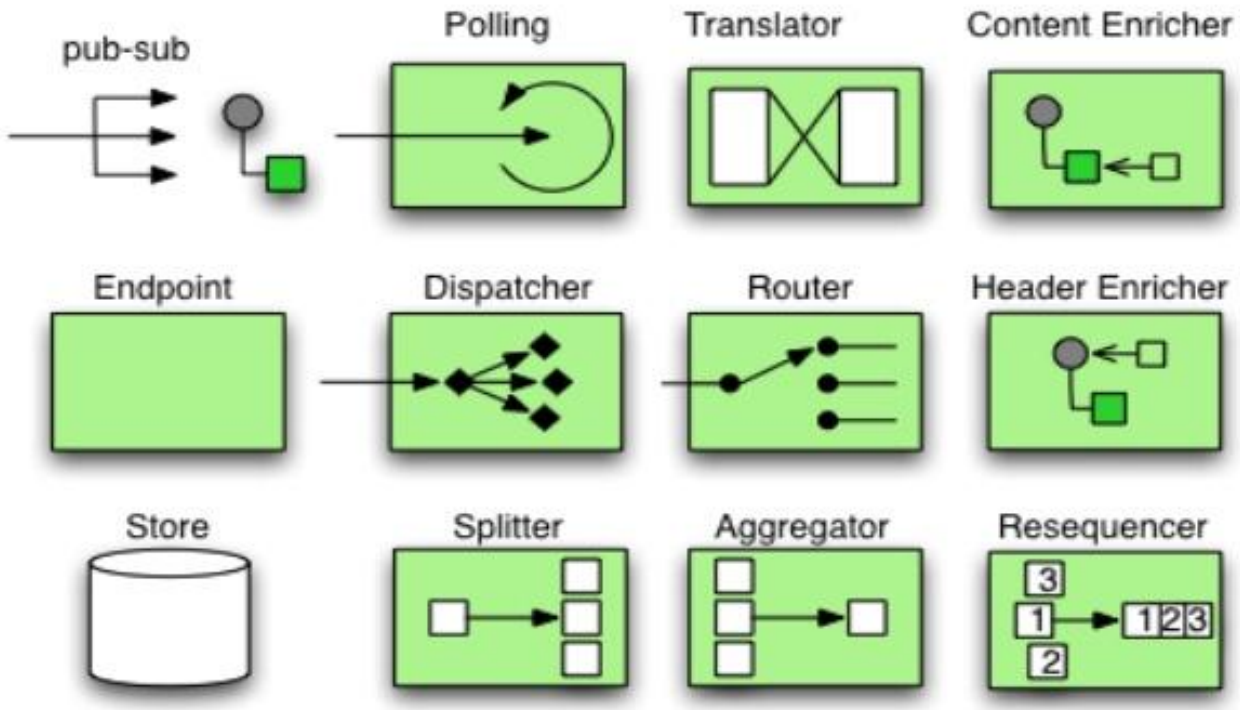
What is Apache Camel

- The swiss knife of integration
- Java based integration framework
- Based on Enterprise Integration Pattern
- Comes with 300+ components
  - UDP, HTTP, MQTT, AMQP, InfluxDB, AWS & Azure services, ...
- DSL to describe the integration flow ... aka “route”



Red Hat

Enterprise Integration Patterns



What is Strimzi

- Open source project licensed under Apache License 2.0
- Focuses on running Apache Kafka on Kubernetes
  - Container images for Apache Kafka, Apache ZooKeeper and other components
  - Operators for deploying, managing and configuring Kafka clusters
- Provides a Kubernetes-native experience
  - Not only Kafka clusters, but also users, topics and the rest of Kafka ecosystem
- CNCF sandbox project since September 2019

XBOX

F1™

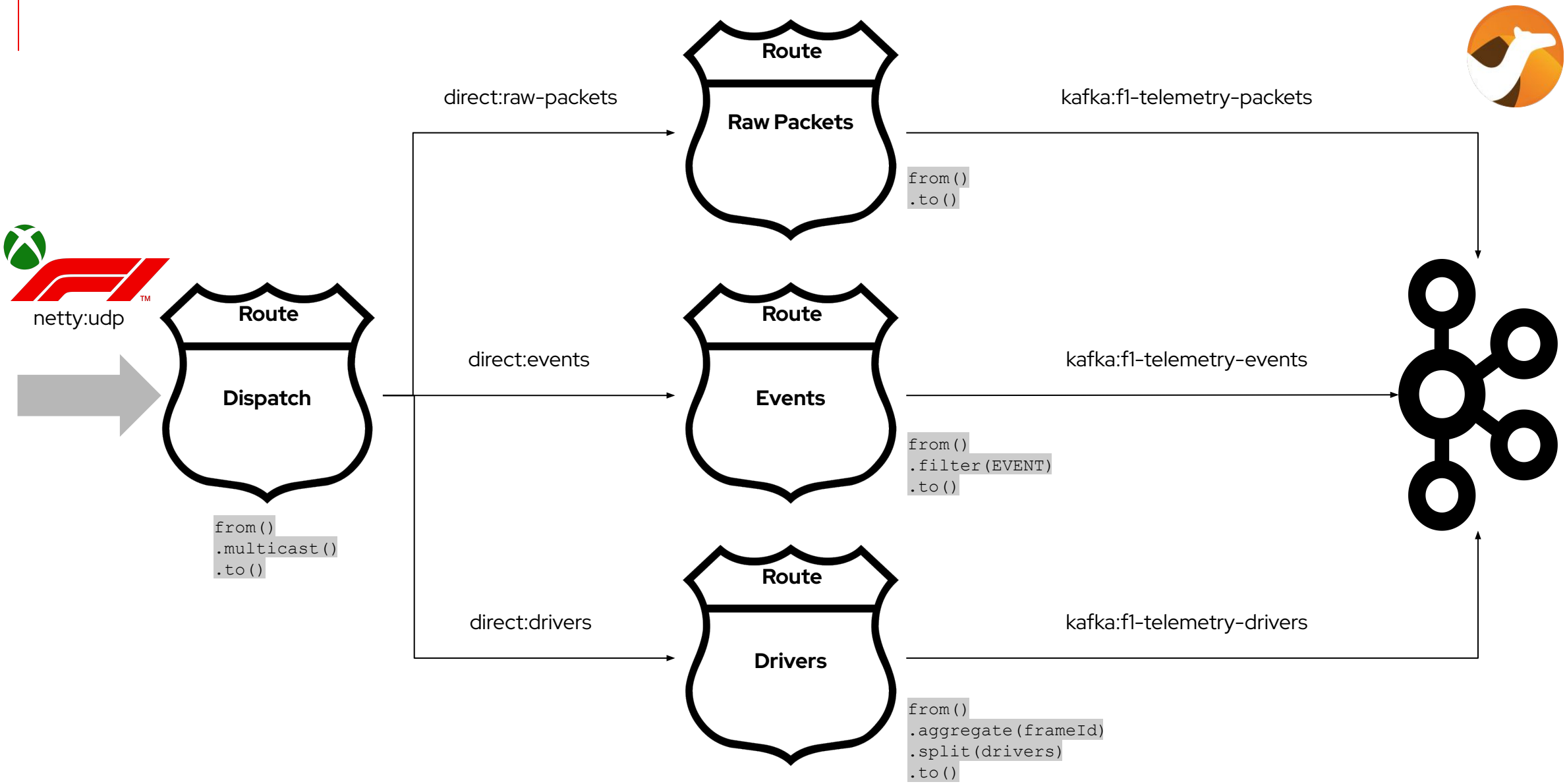
kafka

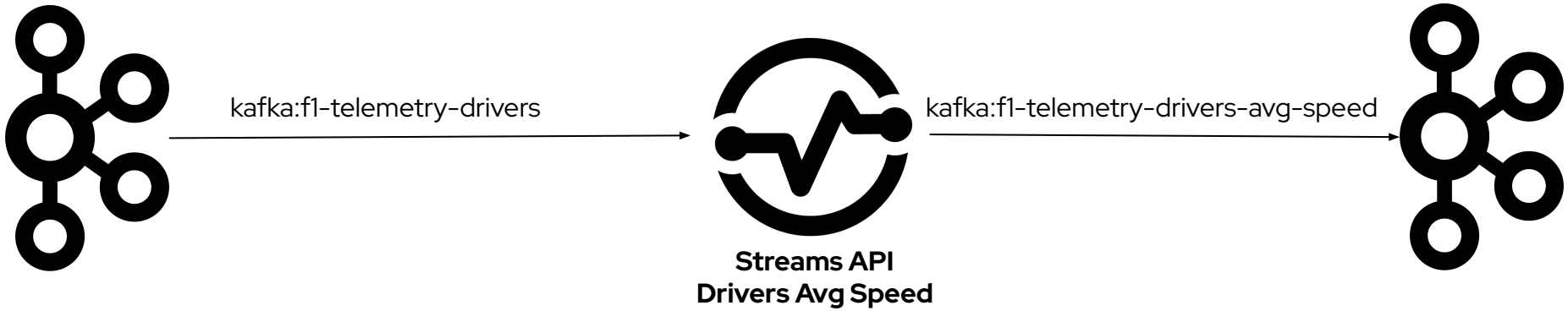
influxdb

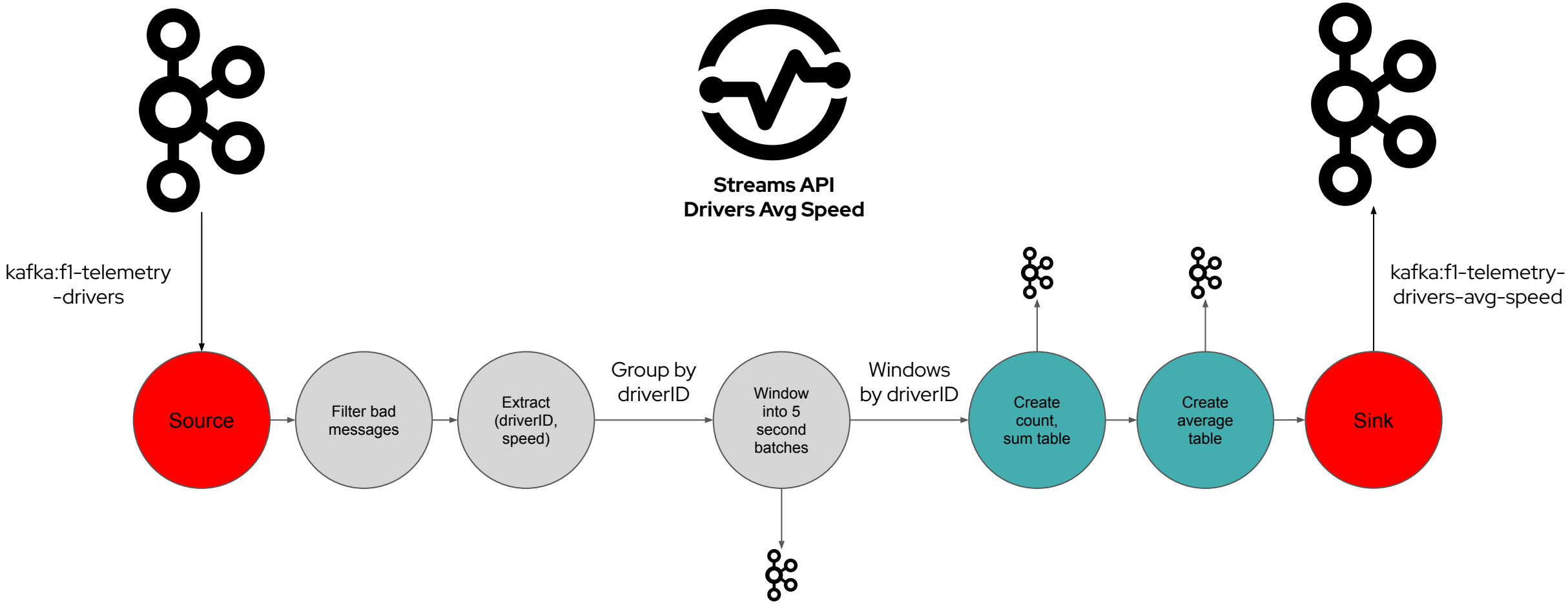
Grafana

Red Hat

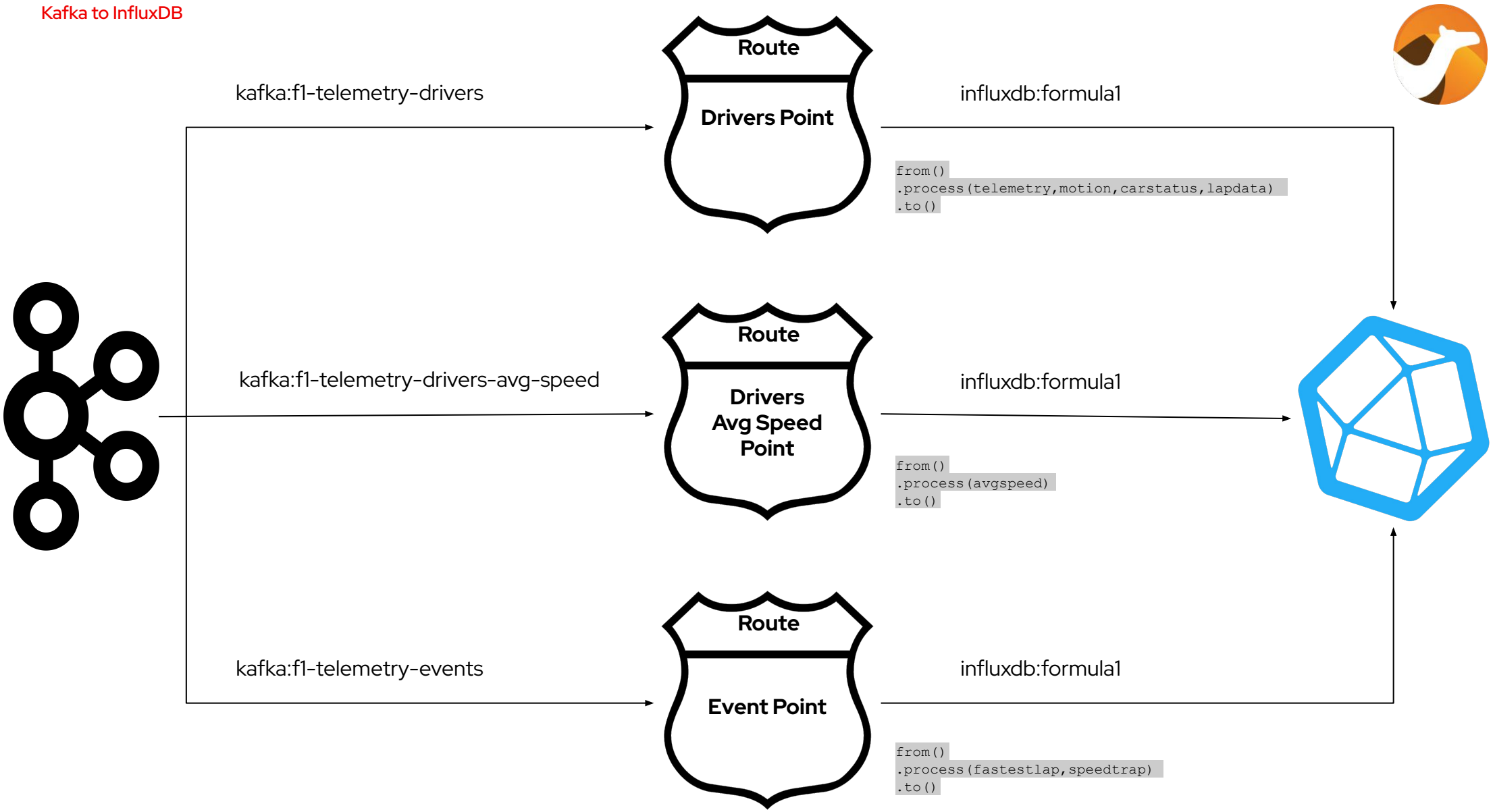
Demo


Route
Raw Packets
direct:raw-packets
kafka:f1-telemetry-packets
from()
.to()
netty:udp
Route
Dispatch
from()
.multicast()
.to()
Route
Events
direct:events
kafka:f1-telemetry-events
from()
.filter(EVENT)
.to()
Route
Drivers
direct:drivers
kafka:f1-telemetry-drivers
from()
.aggregate(frameId)
.split(drivers)
.to()


kafka:f1-telemetry-drivers
kafka:f1-telemetry-drivers-avg-speed
Streams API
Drivers Avg Speed


kafka:f1-telemetry
-drivers
Streams API
Drivers Avg Speed
kafka:f1-telemetry-
drivers-avg-speed
Source
Filter bad
messages
Extract
(driverID,
speed)
Group by
driverID
Window
into 5
second
batches
Windows
by driverID
Create
count,
sum table
Create
average
table
Sink

Kafka to InfluxDB



Red Hat

Resources:

- Blog post: https://grafana.com/blog/2021/02/02/real-time-monitoring-of-formula-1-telemetry-data-on-kubernetes-with-grafana-apache-kafka-and-strimzi/
- F1 decoding library: https://github.com/ppatierno/formula1-telemetry
- F1 Kafka project: https://github.com/ppatierno/formula1-telemetry-kafka
- Video demo: https://www.youtube.com/watch?v=Re9LOAYZi2A

- F1 2020 Codemasters game provides telemetry packets on UDP
  - Specification is available online
- Kubernetes / OpenShift
  - Deploying the Apache Kafka cluster through Strimzi project
  - Running Apache Camel applications, InfluxDB and Grafana
- Apache Camel
  - Ingesting telemetry packets to Apache Kafka
  - Gets telemetry data and race events from Apache Kafka; store into InfluxDB
- InfluxDB
  - Time-series database to provide data to Grafana
- Grafana
  - Showing all telemetry and events on specific dashboards

Thanks!

@ppatierno