IRIRI

Leveraging Streaming Data to enable AI & ML

Saulo Valenzuela | Platform Engineer

IRMRI

Agenda

- ➔ Embracing Streaming Data
- ➔ Leveraging AI & ML with Streaming data
- ➔ Use-cases

IRORI SERVICES

BENEFITS

Business Value

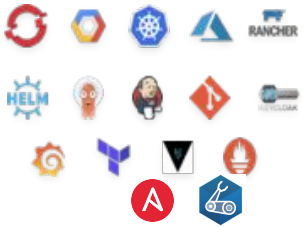
ABILITIES

Platform

Feature Implementation

Streaming Data

TECHNOLOGY



Streaming Data Software

Business value focus
Adapting solutions and automating processes for competitive digital products.

Architecture and method skillset
Apply technology, tools and patterns to achieve architectural abilities, powering the organization and accelerating feature delivery.

Tech mastery
Technology and tooling for transforming IT to modern, scalable and flexible solutions. Development pipelines automate manual steps.

Why Kafka?

Kafka is a distributed event streaming platform designed to handle real-time data feeds.

“Just one of many message brokers out there”

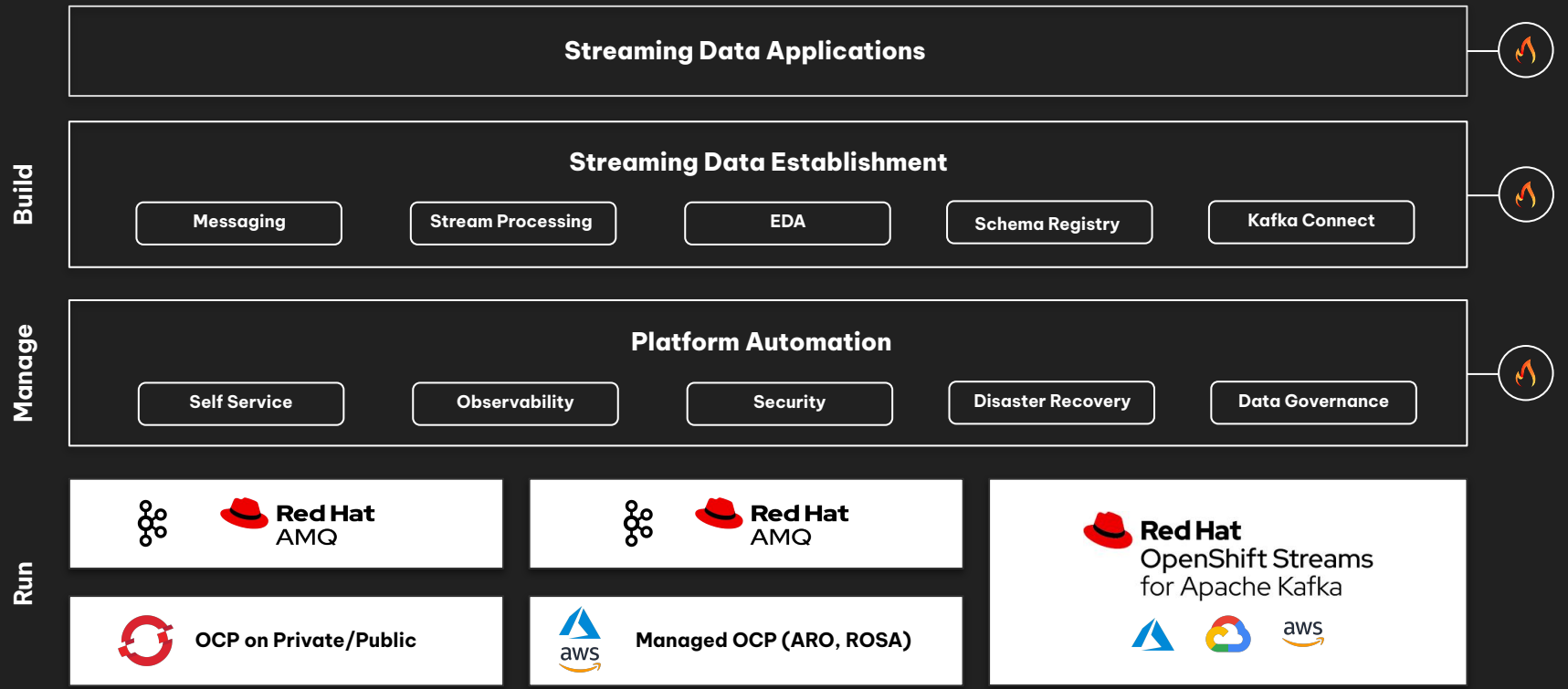
“Not the fastest / hardest / Scooter alternative”

“It’s build with spinning disks in mind…”

But…

- ➔ It’s proven
- ➔ Highly performant, scales well
- ➔ Have a rich ecosystem
- ➔ Handles unlimited retention
- ➔ Not tied to any cloud platform

Irori Data Platform



Manage and Build layer needs are similar regardless of Run-platform in our experience.

Streaming Data

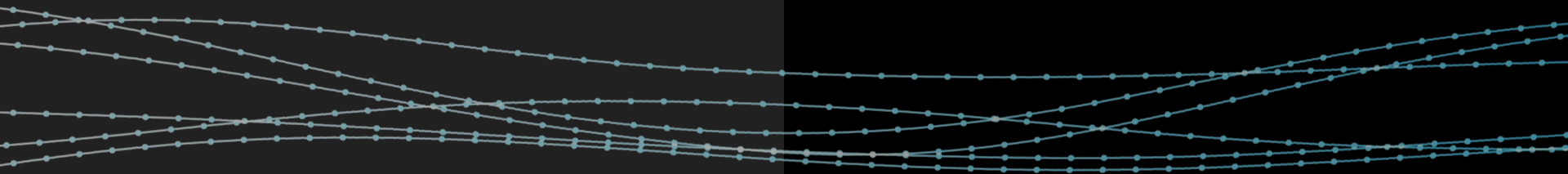
“Acquiring and processing an infinite stream of events in order to refine them and act upon them in real time”

Event driven use cases & real time apps

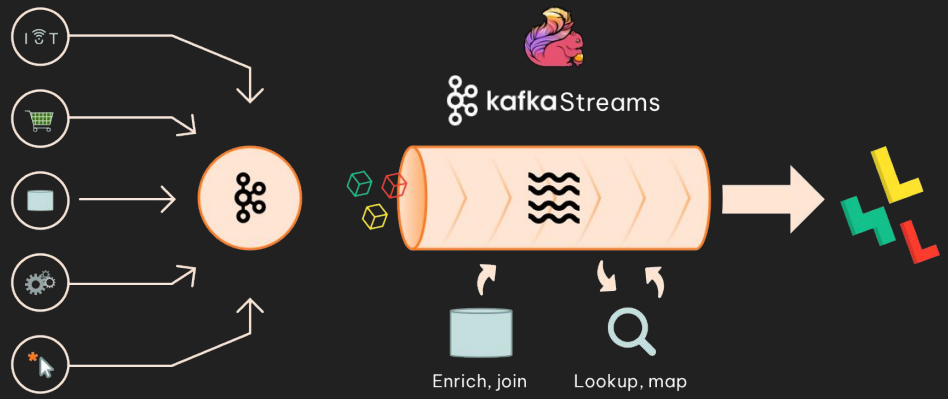
- ➔ Business processes
- ➔ External events
- ➔ Metrics
- ➔ Clickstreams

Trends:

- ➔ Applications get more connected
- ➔ Move towards real time processing
- ➔ Value of data decreases with time



Engage with Streaming Data



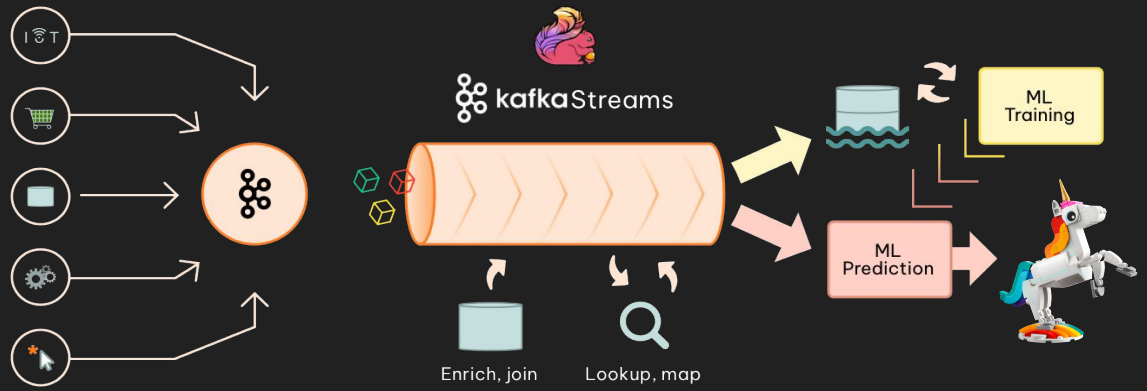
Ingestion

- ➔ Collecting raw data
- ➔ Treat as events
- ➔ Metrics
- ➔ Clickstreams

Stream processing

- ➔ Refine data into valuable information
- ➔ Process events as they arrive
- ➔ Kafka Streams / Apache Flink
- ➔ Enrichment, lookup, joins, mapping etc

Attaching the



Training

- ➔ Adapt stream for ML (normalization etc)
- ➔ Use production data (depersonalize if needed)
- ➔ Collect events in DB for training iterations
- ➔ Don’t reuse the DB, reuse the Stream Process!

Predictions

- ➔ Same Stream Process is used
- ➔ Predictions in (near) real time
- ➔ Simple to bench multiple models in parallel
- ➔ Similar setup for LLMs

Use case
Payment fraud prevention

Goal

- ➔ Detect fraudulent behaviour
- ➔ Prevent payments as they are happening

Stream process:

- ➔ Enrich with merchant type, amount, location
- ➔ Attach previous payments: amount, distance
- ➔ Perform feature normalization

Use case

Optimize offered loan interest rate

Goal

- ➔ Maximize chance of winning the customer...
- ➔ ... but without offering too low interest

Stream process:

- ➔ Join multiple sources: application form data, credit score, previous engagements, transaction history and more
- ➔ Mapping to interesting features
- ➔ Response from AI model to application process

Thank you

Saulo Valenzuela

saulo.valenzuela@irori.se

Read more: www.irori.se