STREAMING DATA ANALYSIS WITH KUBERNETES

Great Indian Developer Summit

Galder Zamarreño Arrizabalaga
@galderz
26th April 2018
@galderz

ENGINEER
Since 2006

INFINISPAN
CO-FOUNDER (2009)
Community Lead and
Core Developer

MUDKIP ROCKS! 😎
BUILD STREAMING DATA ANALYSIS APPLICATION ON TOP OF A KUBERNETES PLATFORM
THE PROBLEM

REAL-TIME STREAMING DATA PROCESSING IS CHALLENGING

Delays can have a big impact

EXPONENTIAL DATA GROWTH YEAR ON YEAR

Smartphones, IOT devices, trillions of internet connected devices...
ZZZ... NO!
THE PLATFORM
Platform-as-a-Service (PaaS)

Platform for developing and running applications

Public or private and multi-language

OpenShift is a Kubernetes distro with extras
The Glue

Vert.x is a toolkit for building reactive apps
On JVM, event-driven and non-blocking

RxJava integrates with Vert.x
Great at event transform and coordination
Works best with many source of events (modern apps!)
DATA GRID

- Polyglot
- Extreme Performance
- Linear Scalability
- Fault Tolerant
- Event driven

Fuse "memory" across machines into a unified data store

Read-through, write-through, write-behind
THE DATA
http://transport.opendata.ch
"stop": {"station": {"id": "8500301", "name": "Rheinfelden", "score": null, "coordinate": {
"type": "WGS84", "x": 47.55121, "y": 7.792155}, "distance": null}, 
"arrival": null, 
"arrivalTimestamp": null, "departure": "2016-02-29T17:34:00+0100", "departureTimestamp": 1456763640, 
"delay": 3, "platform": 4, "prognosis": {
"platform": 4, "arrival": null, "departure": "2016-02-29T17:37:00+0100", "capacity1st": 1, 
"capacity2nd": 1}, 
"realtimeAvailability": null, "location": {"id": "8500301", 
"name": "Rheinfelden", "score": null, "coordinate": {
"type": "WGS84", 
"x": 47.55121, 
"y": 7.792155}, 
"distance": null}}, 
"name": "IR 1978", "category": "IR", "categoryCode": 2, 
"number": "1978", "operator": "SBB", "to": "Basel SBB", "capacity1st": null, 
"capacity2nd": null, "subcategory": "IR", 
"timeStamp": 1456761753983, "nextStation": {
"station": {"id": "8500301", "name": "Rheinfelden", "score": null, "coordinate": {
"type": "WGS84", 
"x": 47.55121, 
"y": 7.792155}, 
"distance": null}, 
"arrival": "2016-02-29T17:34:00+0100", "arrivalTimestamp": 1456763640, 
"departure": null, 
"departureTimestamp": null, 
"delay": null, "platform": 4}, "prognosis": {
"platform": null, "arrival": null, "departure": null, "capacity1st": null, "capacity2nd": null}, "realtimeAvailability": null, 
"location": {"id": "8500301", 
"name": "Rheinfelden", "score": null, "coordinate": {
"type": "WGS84", 
"x": 47.55121, 
"y": 7.792155}, 
"distance": null}}, #{@version": 1", @timestamp": 2016-02-29T16:02:34.781Z"}
SAMPLE DATA - SBB

{"x":"8290840","y":"47483629","name":"IR 1978","trainrefdate":"29.02.16","category":"IR","trainid":"84/25934/18/24/95","direction":"15","prodclass":"4","delay":"7","passproc":"","lstopname":"Basel SBB","poly": [{"x":"8290840","y":"47483629","passproc":"","msec":"0","direction":"15"},{"x":"8290193","y":"47483647","passproc":"","msec":"2000","direction":"15"},{"x":"8289528","y":"47483674","passproc":"","msec":"4000","direction":"15"},{"x":"8288863","y":"47483701","passproc":"","msec":"6000","direction":"15"},{"x":"8288198","y":"47483728","passproc":"","msec":"8000","direction":"15"},{"x":"8287532","y":"47483755","passproc":"","msec":"10000","direction":"15"},{"x":"8286885","y":"47483773","passproc":"","msec":"12000","direction":"15"}],"timeStamp":1456761728202,"@version":"1","@timestamp":"2016-02-29T16:02:11.321Z"}
HIGH LEVEL ARCHITECTURE
COMPONENT ARCHITECTURE
DEMO
VERSATILITY OF INFINISPAN

- Distributed Cache
- Shared Memory
- Event Broker
- Analytics

- Embedded
- Remote
- Expiration
- Eviction
- Persistence

- Security
- Code execution
- Transactions
- Listeners

- Management and monitoring
- Cloud
- Querying
- Integrations
BUILD STREAMING DATA ANALYSIS APPLICATION ON TOP OF A KUBERNETES PLATFORM
THANK YOU!

github.com/infinispan-demos/streaming-data-kubernetes
infinispan.org
openshift.com | vertx.io
GREAT INDIAN DEVELOPER SUMMIT 2019

Conference: April 23-26, Bangalore

Register early and get the best discounts!

www.developersummit.com  @greatindiandev  bit.ly/gidslinkedin