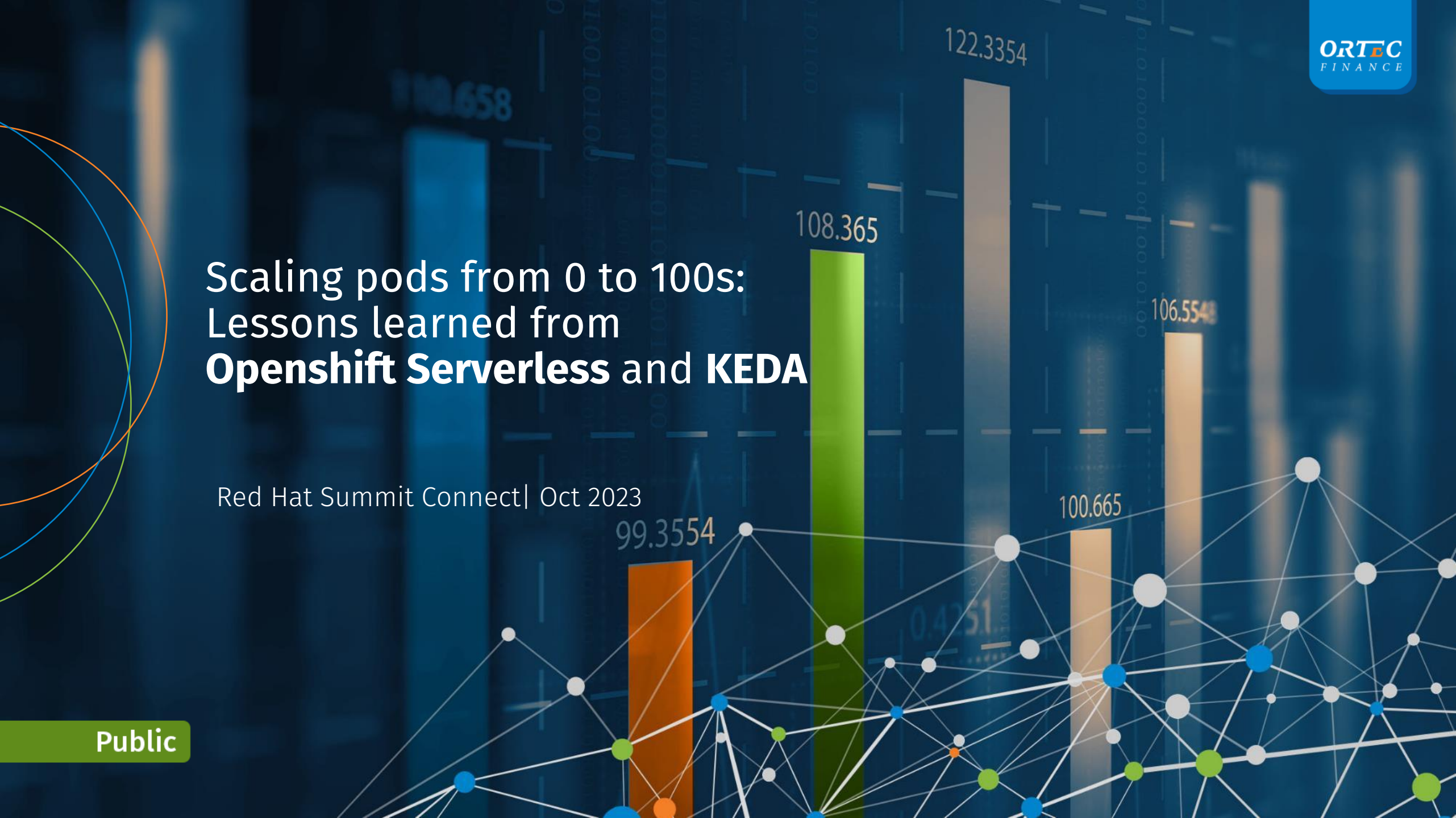


Scaling pods from 0 to 100s: Lessons learned from **Openshift Serverless** and **KEDA**

Red Hat Summit Connect| Oct 2023

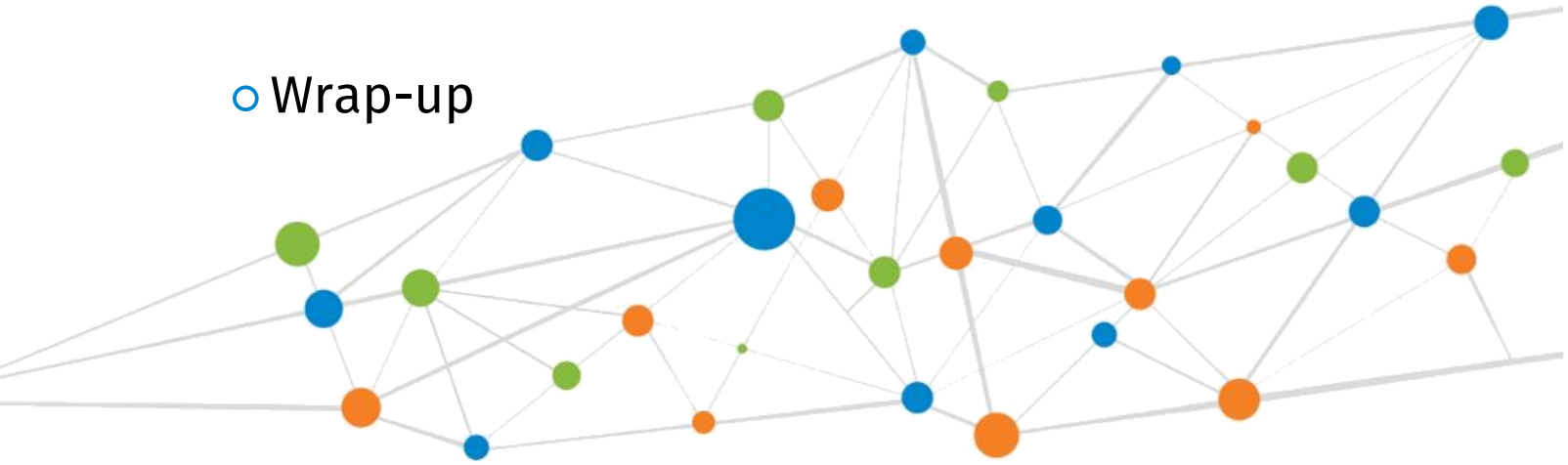
Public

Scaling pods from 0 to 100s

Lessons learned from Openshift Serverless and KEDA

Agenda:

- Context – What type of workloads are we talking about?
- Openshift Serverless + Demo
- KEDA + Demo
- Wrap-up



What we do & who we are

We enable people to manage the complexity of investment decision making

Ortec Finance is a global technology and solutions provider

Global client base

North America

Europe

Asia

Pacific

OMERS
THE WORLD BANK
OPTrust
HOOPP Healthcare of Ontario Pension Plan
UNIVERSITY OF CALIFORNIA Office of the Chief Investment Officer
BT PENSION SCHEME
USS
AP FJÄRDE AP-FONDEN
Andra AP-fonden
G.S.R. de nederlandse verzekerings maatschappij voor alle verzekeringen
SEB
apg
ST. JAMES'S PLACE WEALTH MANAGEMENT
KEVA
PGGM
ING
pensionskasse STADT ZÜRICH
Investment Data Services A company of Allianz
achmea Investment Management
ABB
ROBECO The Investment Engineers
elo
CERN
ABN·AMRO
PGB pensioendiensten
AEGON
Pensioenfonds Rail & OV
Korea Fixed Income Research Institute 한국채권연구원
income made different
QIC
ACC
cbus
futurefund Australia's Sovereign Wealth Fund
QSuper
NEW ZEALAND SUPERANNUATION FUND
AustralianSuper
ORTEC FINANCE

Public

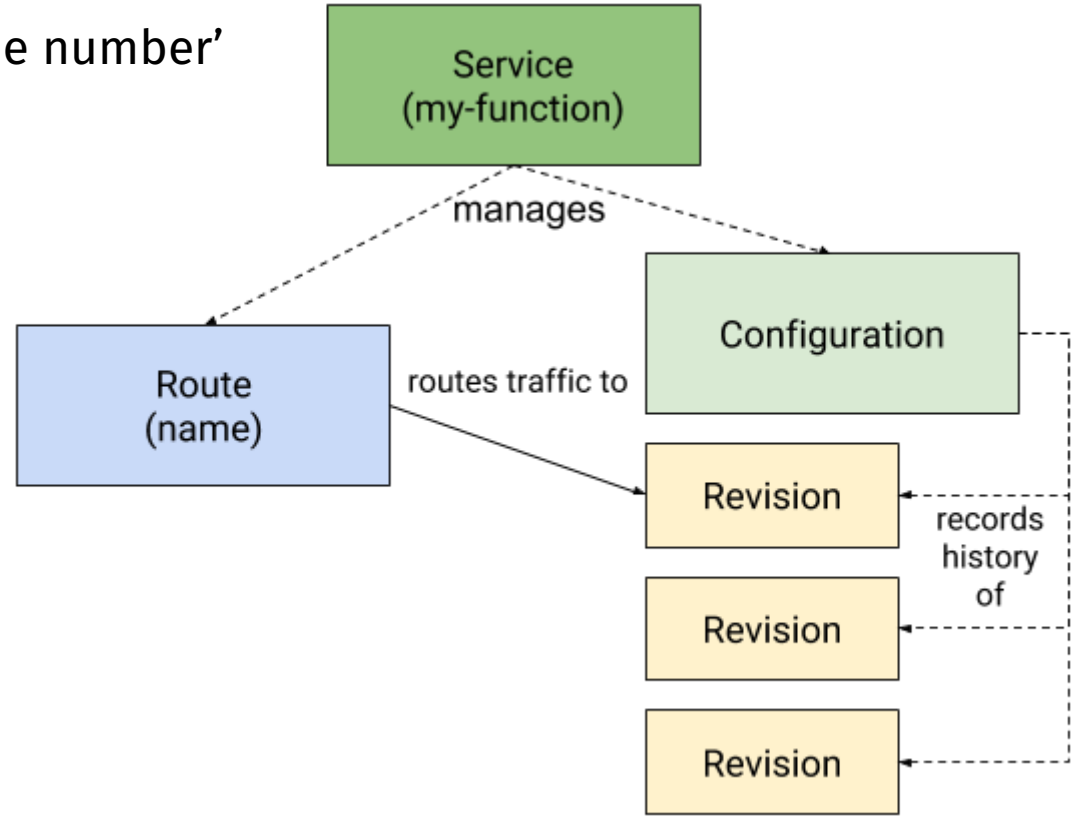
Our Cloud-Native Journey

- We like Managed over DIY:
  - OpenShift is a very mature, battle-tested, enterprise-level Kubernetes toolkit. It's a 'batteries-included' platform
  - We leverage opinionated Red Hat stack (e.g. prefer supported operators)
- We don't lift and shift. – Modernize Apps first
- Migrating Legacy Microsoft HPC workloads this year

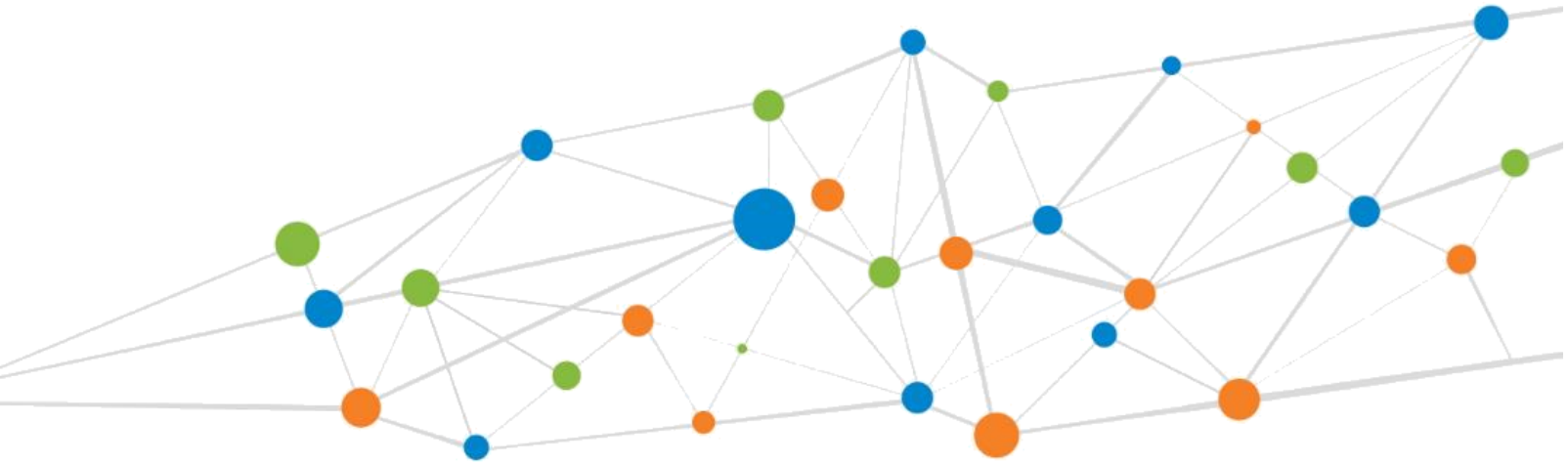
Openshift Serverless

The ability to scale to zero – “Some people call this Serverless”

- Benchmark:
  - dotnet core implementation of naive ‘highest prime number’
  - We kept it synchroneous (no Knative Eventing)
  - Node selectors enabeld



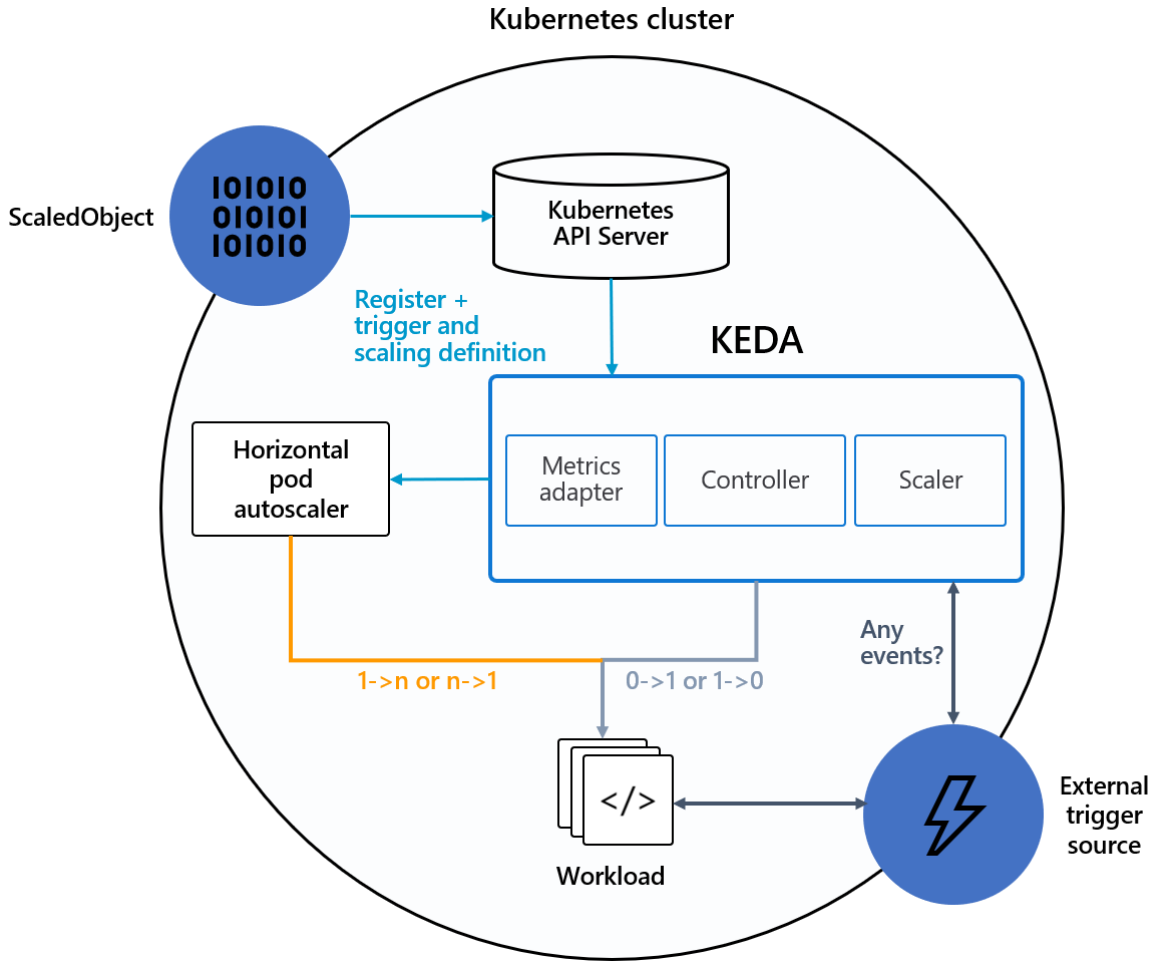
DEMO



KEDA

Kubernetes Event Driven Autoscaler

- Benchmark:
  - Same prime number algo
  - Async; Producer / Consumer set up
  - No Kafka, but AMQ Broker



DEMO

- Operates on standard k8s resources
- Can scale existing deployed apps
- Pull based approach
- Doesn't manage data delivery
- K8s Horizontal Pod Autoscaler (HPA)
- Focus is on event driven autoscaling

- Operates on Knative Service
- Existing apps must be converted
- Push based approach
- Manages data delivery (Eventing)
- Knative Autoscaler
- Demand-based autoscaling (HTTP)

https://developers.redhat.com/devnation/tech-talks/autoscaling-KEDA-knative

Wrap-up

- For (long) paralell multiproces computations we prefer the **KEDA** setup
  - \+ Operates on standard K8 resources
  - \+ Can scale exisiting apps
  - \+ Tweakable autoscaling (e.g. deal with cold starts)
  - \- Community not all KEDA scalers not production ready (EDA for prometheus is shipped in 4.11)

- For (short) multithreaded Compute we want to investigate **Openshift Serverless** further:
  - \+ Simplified deployment syntax
  - \+ Also includes traffic distribution
  - \- Async support requires rigorous design
  - \- Not designed for long running tasks (e.g. Kourier time-outs in 300s)

Contact me

Joris Cramwinckel

Head of Cloud Native Platform

+316 200 002 86

Joris.Cramwinckel@ortec-finance.com

ortecfinance.com

Rotterdam
Pfäffikon
Melbourne
Amsterdam
Toronto
London
New York