

AKAMAS

Support your strategic initiatives with ML-powered optimization

 provance DynaTaste 2022

2022 April 27th



Agenda

- **The problem**
- **Enter Akamas**
- **A real-world case**
- **Conclusions and Q&A**



Giuseppe Nardiello

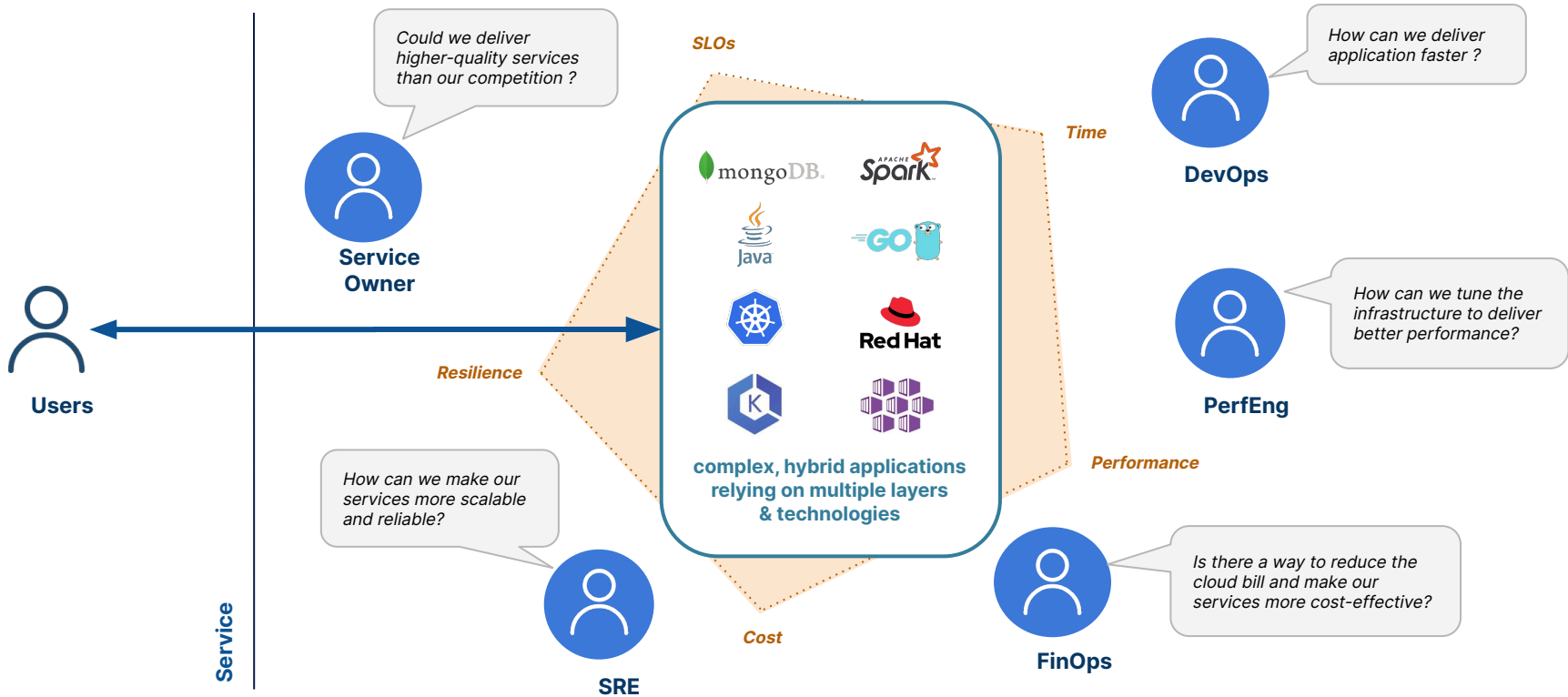
VP Product Mgmt & Business Dev

AKAMAS

The problem

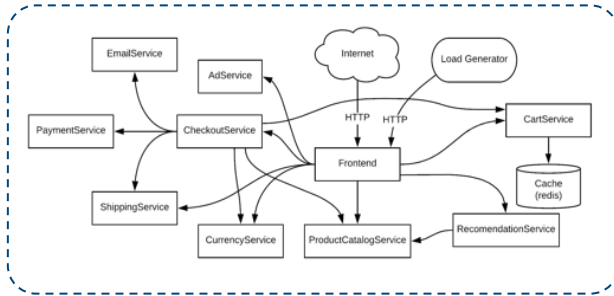


The challenge: align IT to Business



The challenge of cloud-native apps

Kubernetes microservices applications



dozens or hundreds of microservices



Tunable parameters

- Application Runtime (each microservice)
- Resource requests (each microservice)
- Resource limits (each microservice)
- Replicas & Auto-Scaling policies
- ...

hundreds or thousands of parameters

how to balance performance, resilience & cost?

Enter Akamas



Akamas autonomous optimization principles

AKAMAS

Full-Stack



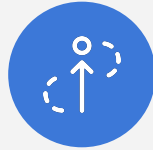
any application, any middleware, any database, any cloud - any system

Smart



patented AI quickly identifying optimal configurations beyond any manual tuning

Goal-oriented



custom-defined goals translating SLOs and business & technical constraints

Closed-loop



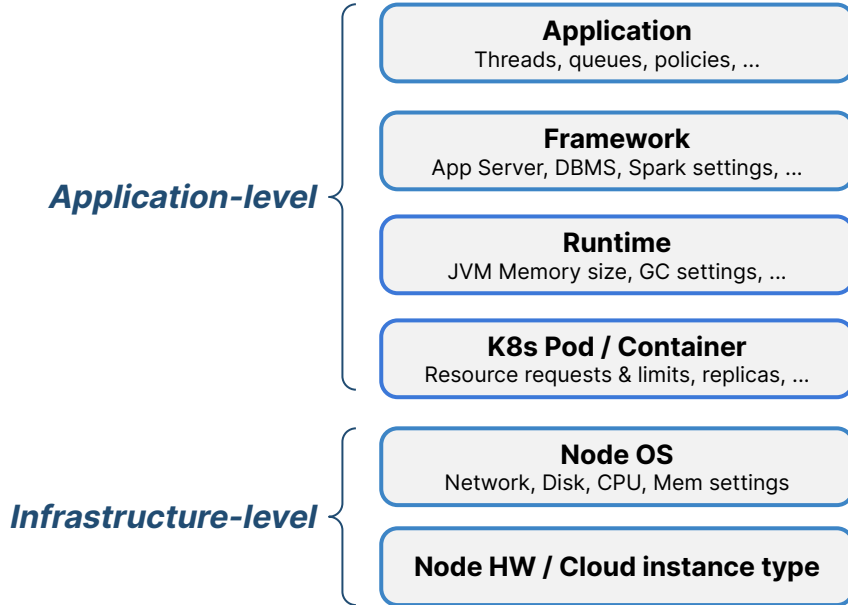
recommended changes with human-in-the-loop or fully automated within-the-pipe

Safe



parameter boundaries, incremental changes and constraint violation detection

Akamas full-stack approach



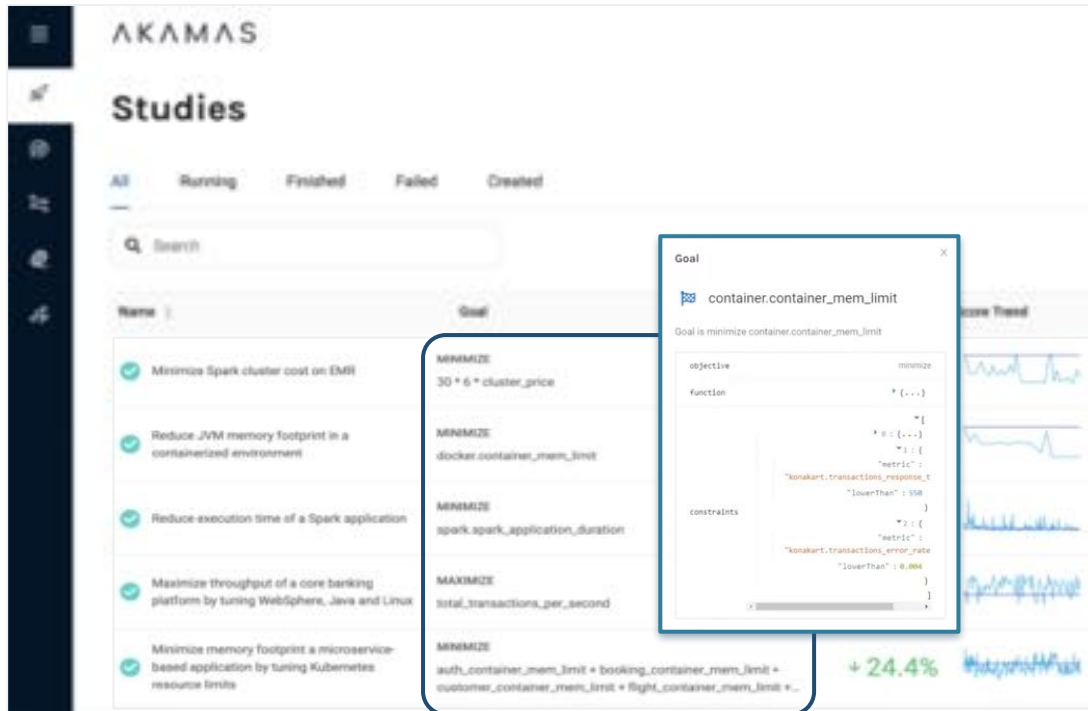
HUNDREDS/THOUSANDS OF
TUNABLE PARAMETERS AND KPIS

full-stack optimization



illustrative - sample technologies

Akamas optimization goals and constraints



declarative, custom-defined goals & constraints

MINIMIZE resource usage WITH throughput ≥ 1000

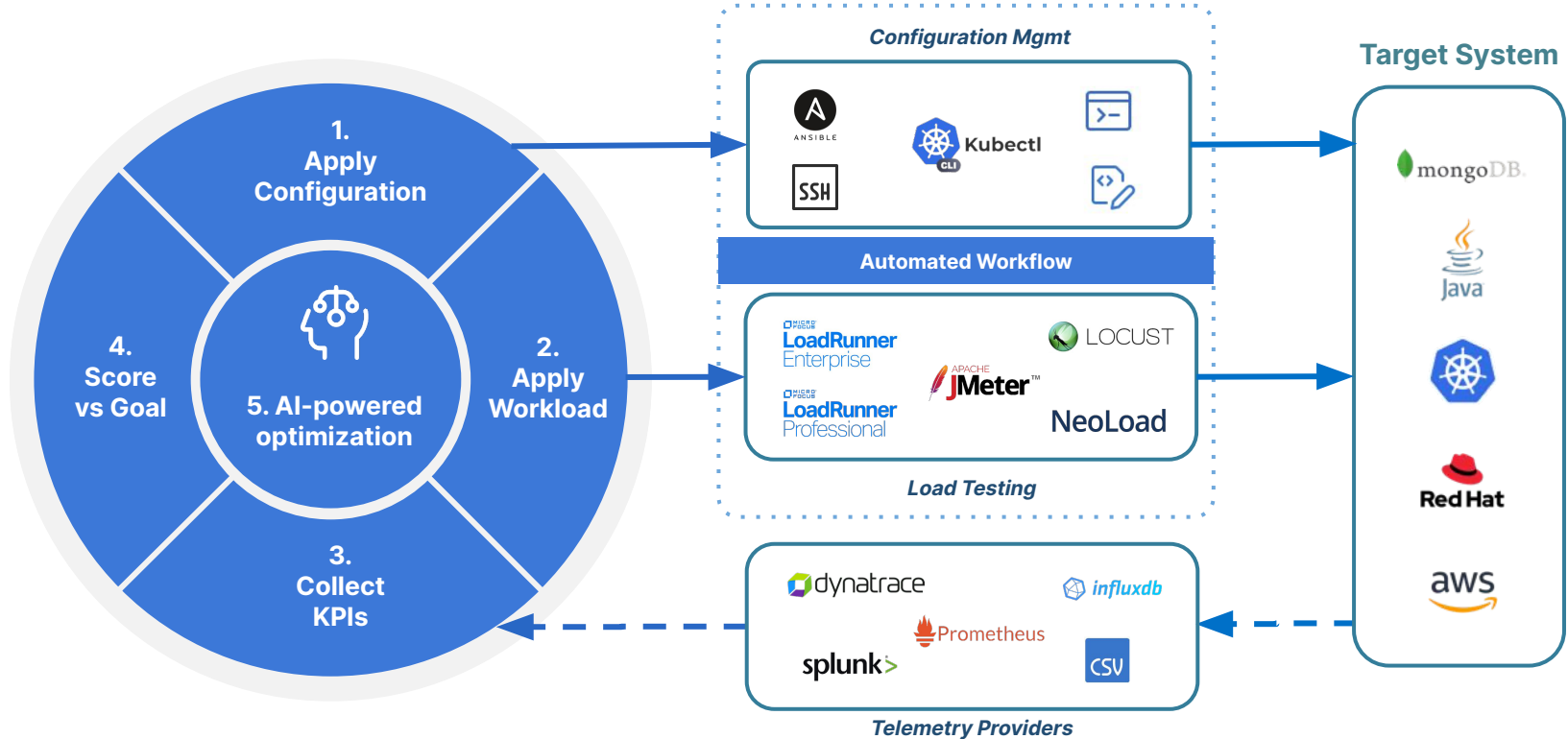
MINIMIZE response time WITH $\leq \text{error_rate (BL)} + 5\%$

MINIMIZE cloud cost WITH $\text{trans/sec} = \text{BL} + 5\%$

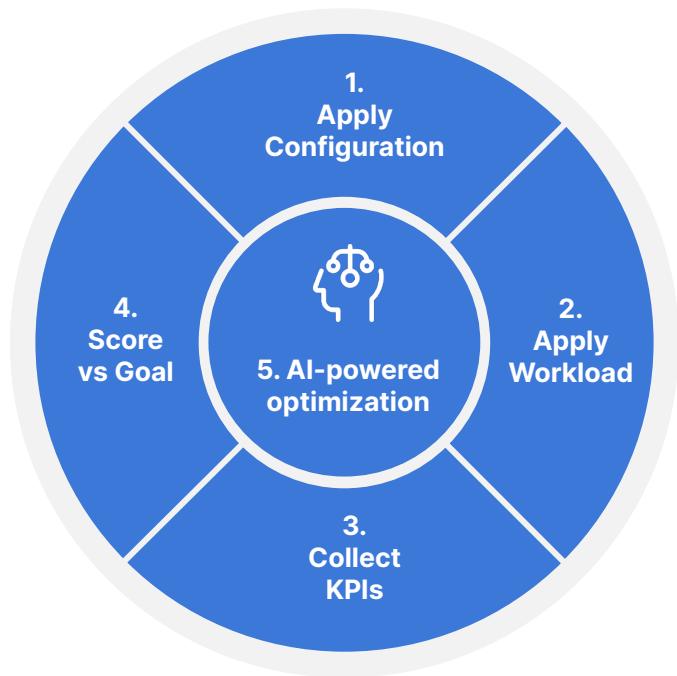
MAXIMIZE throughput / heap-size
WITH $\leq \text{err_rate (BL)}$ AND $\text{throughput} = \text{BL} + 5\%$

any formula on one or more metrics (absolute and wrt baseline)

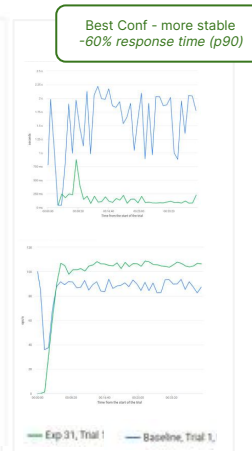
Akamas high-level architecture (offline)



Akamas smart exploration



Goal: maximize transaction / monthly cost



~35 experiments / ~24 hours elapsed to identify optimal configuration

Akamas enables a new decision-support



DevOps



PerfEng



SRE



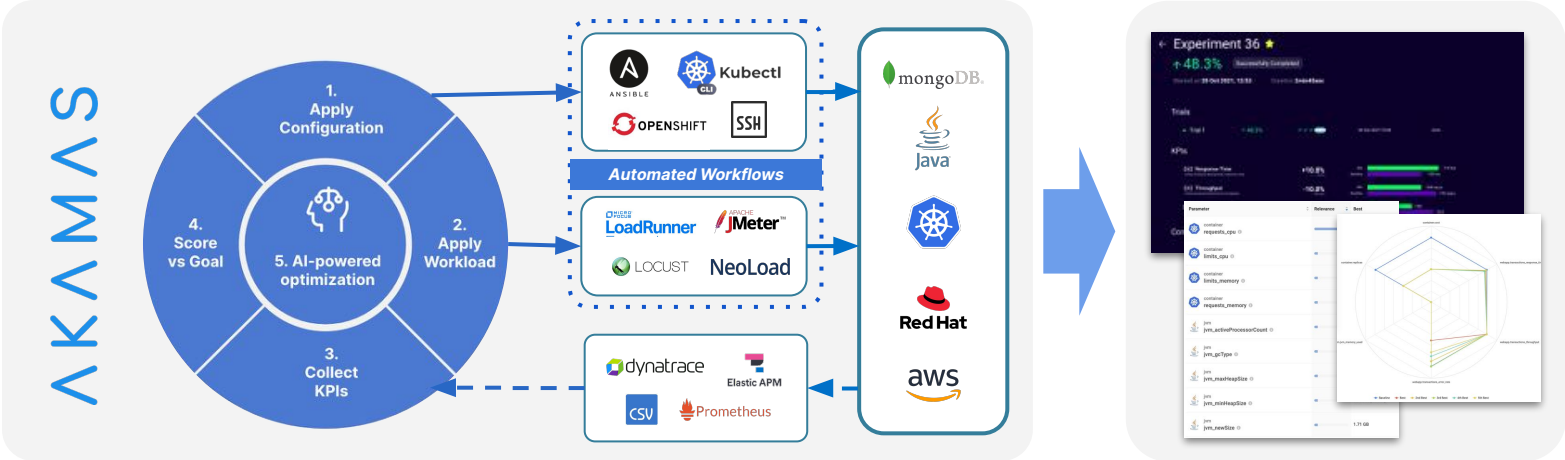
Service
Owner



FinOps

Goals & Constraints (SLOs)

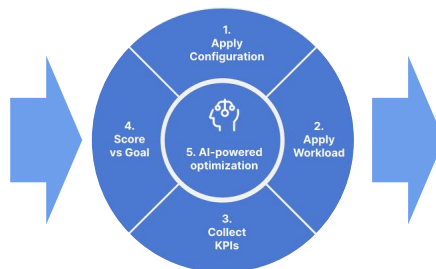
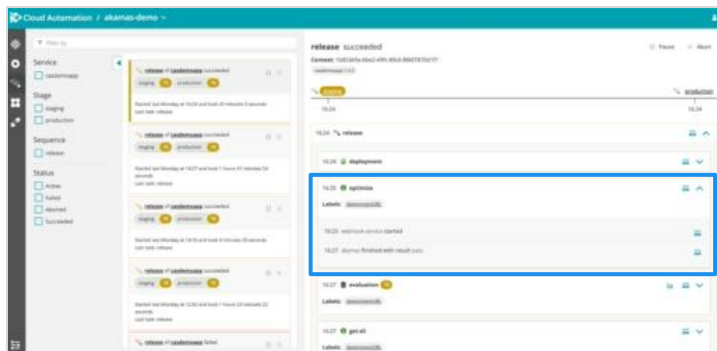
Best Configs & Insights



Akamas optimization in the CI/CD pipeline



AKAMAS



optimal configuration identified by Akamas

Akamas **Continuous Optimization (CO)** triggered in the **CI/CD pipeline** triggered during release process (under Quality Gates) or as remediation response

Key customers & metrics

Service Throughput



+30%

increase in transactions/seconds with the same infrastructure resources

Service Quality



-70%

decrease in response time with lower fluctuations and peaks

Cost Reduction



-60%

decrease in infrastructure/cloud costs with the same application performance

Operational Efficiency



-80%

savings in engineering time spent for manual tuning tasks

TripActions

lastminute.com

vodafone

Sabre

italiaonline

enel

nexi  SIA

 UniCredit

RAIFFEISEN

CEDACRI
GROUP

A real-world case




Real-world case: challenges

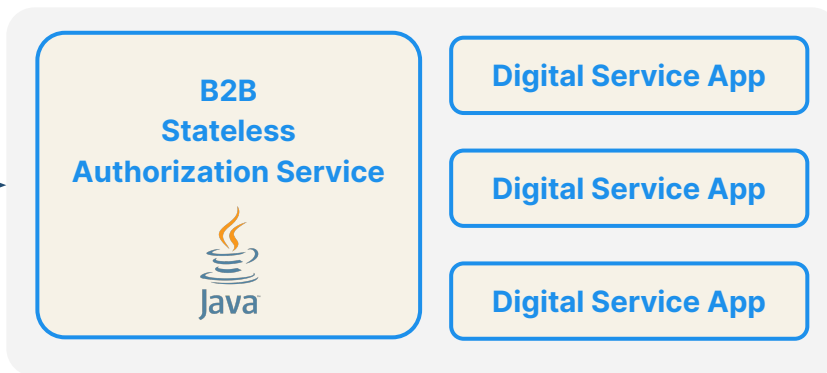


European leader in accounting,
payroll & business management software

1,7M
users



400M
invoices / year



frequent updates dictated by
business & regulatory compliance

2 months spent for tuning a
single Kubernetes microservice

need to lower cloud bills and
improve cost-efficiency



Azure Kubernetes Service
(AKS)

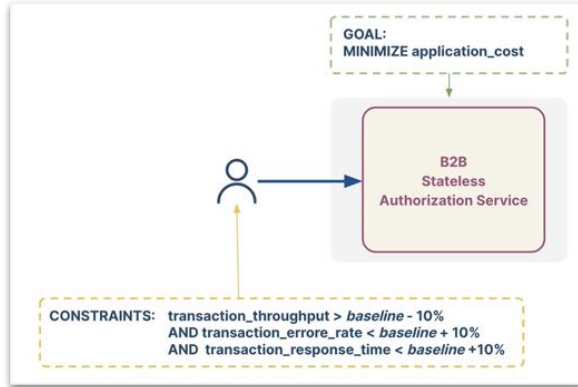


AWS Elastic Kubernetes Service
(EKS)

Akamas results



Goals & Constraints (SLOs)



Reduced Cost

CONFIGURATION #34 (after 19h)

-49.1%

Parameter	Release	BEST	BASILINE
container requests_cpu	3.77 cores (+48.3%)	2.77 cores	1.5 cores
container limits_cpu	3.87 cores (+42.3%)	2.77 cores	2.00 cores
container limits_memory	5.16 GB (+17.3%)	4.39 GB	3.72 GB
container requests_memory	5.06 GB (+48.3%)	3.42 GB	2.30 GB
jvm_jvm_activeProcessorCount	1.00%	-	-
jvm_jvm_gcType	g1	-	-
jvm_jvm_maxHeapSize	4.76 GB (+19%)	4.00 GB	3.36 GB
jvm_jvm_minHeapSize	4.37 GB (+71%)	512 MB	512 MB

CONFIGURATION #14 (after 8h)

-15.9%

Component	Parameter	HIGH RELIABILITY	BEST	BASILINE
container	limits_cpu	2.77 cores	2.87 cores	2.00 cores
container	limits_memory	5.69 GB	5.16 GB	4.39 GB
container	requests_cpu	1.17 cores	2.77 cores	1.50 cores
container	requests_memory	5.4 GB	5.06 GB	3.42 GB
jvm	jvm_activeProcessorCount	6.00%	1.00%	-
jvm	jvm_gcType	ParNew1	g1	-
jvm	jvm_maxHeapSize	3.45 GB	4.76 GB	4.00 GB
jvm	jvm_minHeapSize	1.94 GB	4.37 GB	512 MB
jvm	jvm_newSize	1.00 MB	1.71 MB	-

Improved Resilience

Conclusions and Q&A

Your next steps

1

[Download the Case Study](#)

2

[Read Gartner Peer Insights](#)

3

[Free Trial](#)

Contacts

Italy HQ

Via Schiaffino 11
Milan, 20158
+39-02-4951-7001

USA West

12130 Millennium Drive
Los Angeles, CA 90094
+1-323-524-0524

LinkedIn

[@akamaslabs](#)

Email

info@akamas.io

USA East

211 Congress Street
Boston, MA 02110
+1-617-936-0212

Singapore

5 Temasek Blvd
Singapore 038985

Twitter

[@AkamasLabs](#)

Website

[akamas.io](#)

