WELCOME TO | WILLKOMMEN ZU

EUROPE | BERLIN 2024

Conference Keynote

Delivering the Next Wave of Innovation with OpenSearch

OpenSearchCon
EUROPE | Berlin 2024

Mukul Karnik

General Manager, OpenSearch

Search is the foundation

OpenSearch is an information retrieval system

Delivers the most relevant results

Text search, faceting, geospatial, auto-complete, fuzzy matching

Helps finds patterns in logs

Unlocks the latest AI/ML technologies

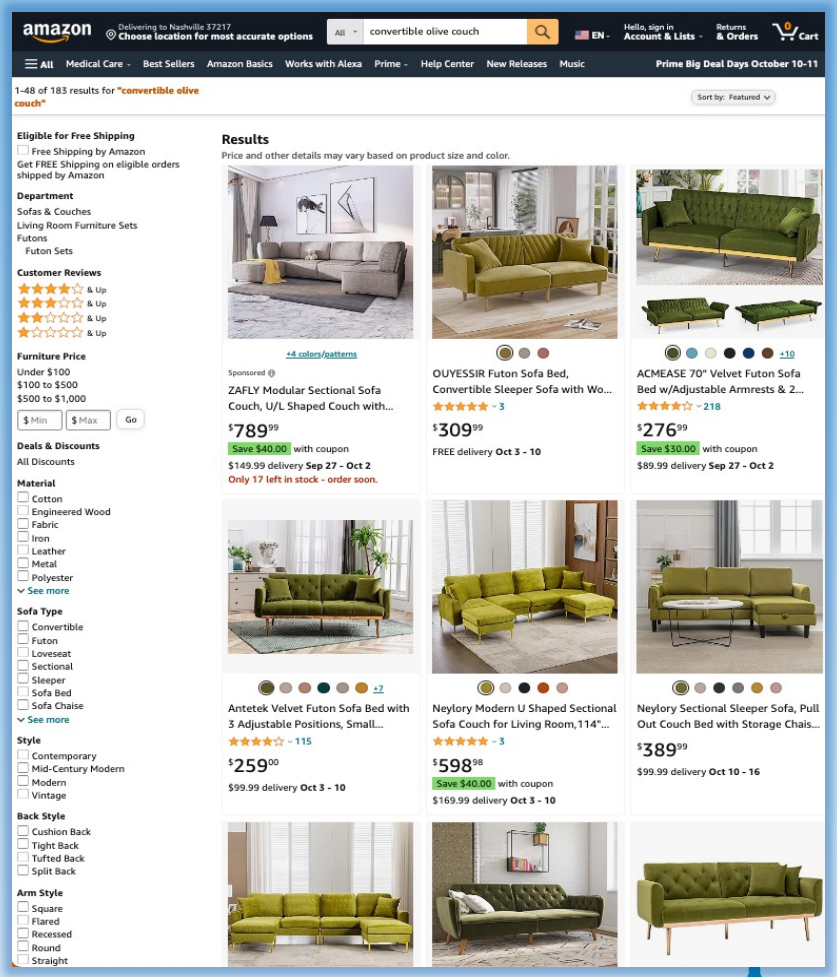
Text search

Search engines match terms to objects in a catalog

Match free text or structured data

Relevance determines sort order

Facets let users drill into the search results to narrow by attribute value



The boom in AI/ML

Large language models (LLMs) providing good baseline NLP

ChatBots hit the public's awareness

"Semantic" capabilities are directly relevant for search workloads

Other AI/ML techniques are highly relevant for search as well

OpenSearch enables working with vector embeddings

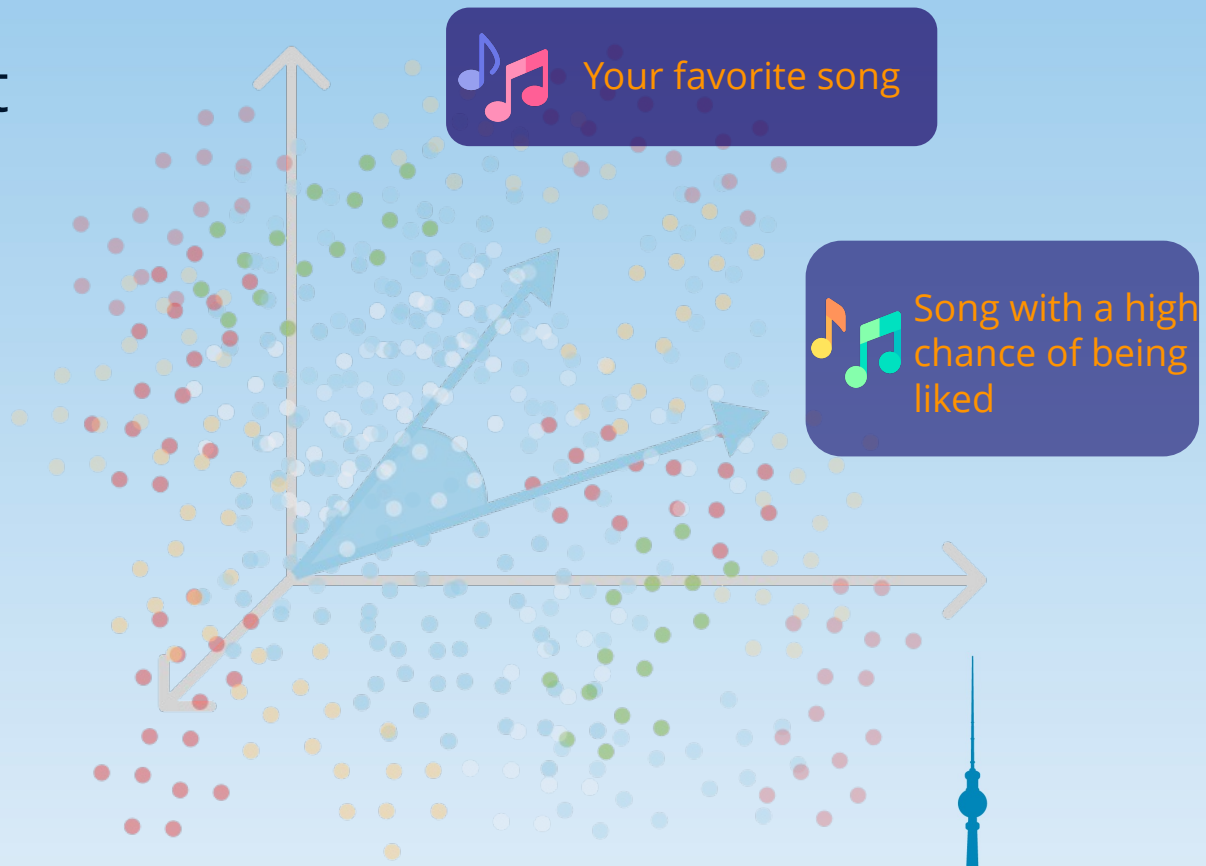
Vector similarity search

Generative AI models represent unstructured data as vector embeddings

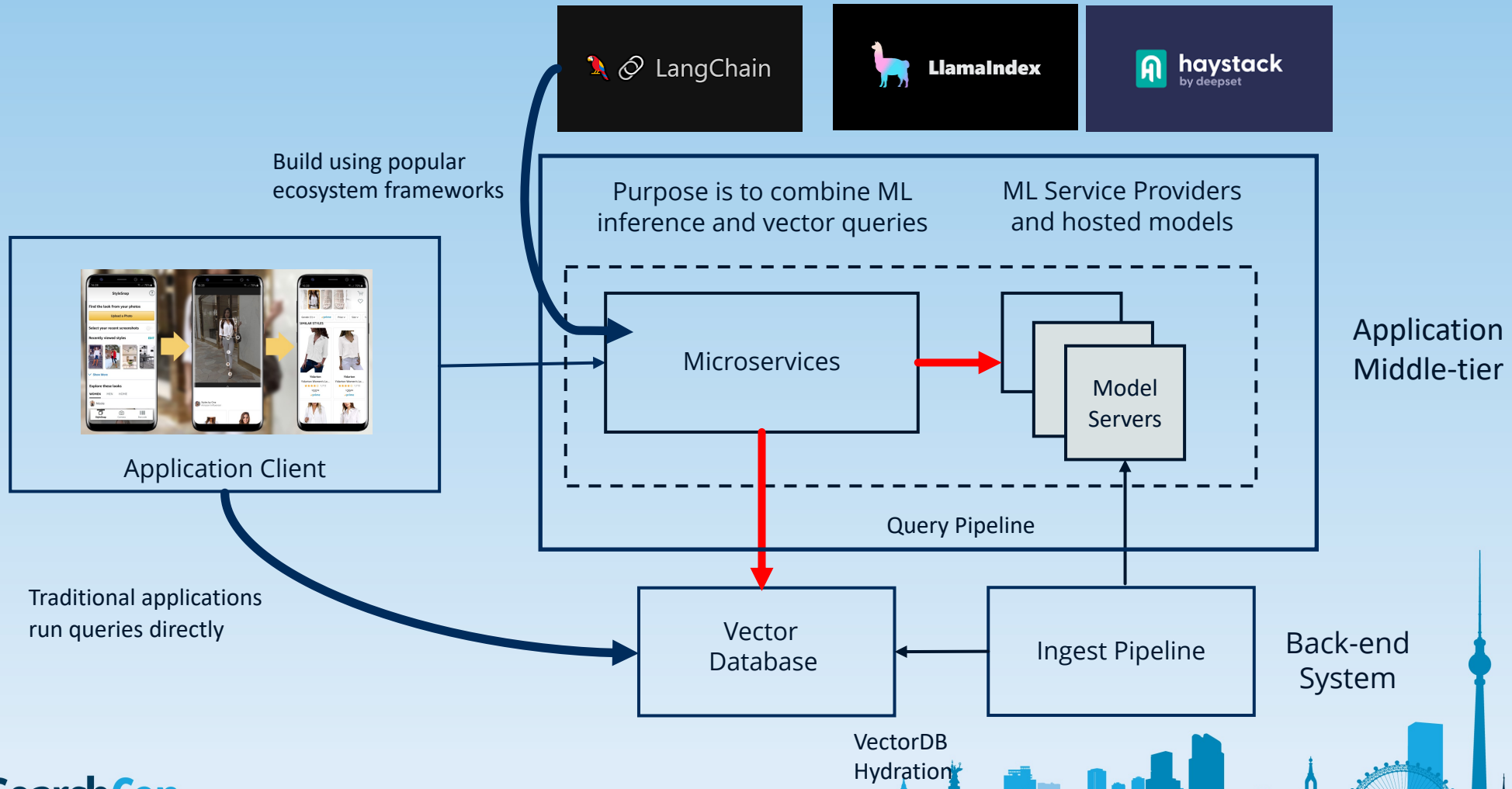
Embeddings capture semantic meaning and content relationships

Similar vectors = similar meaning and context

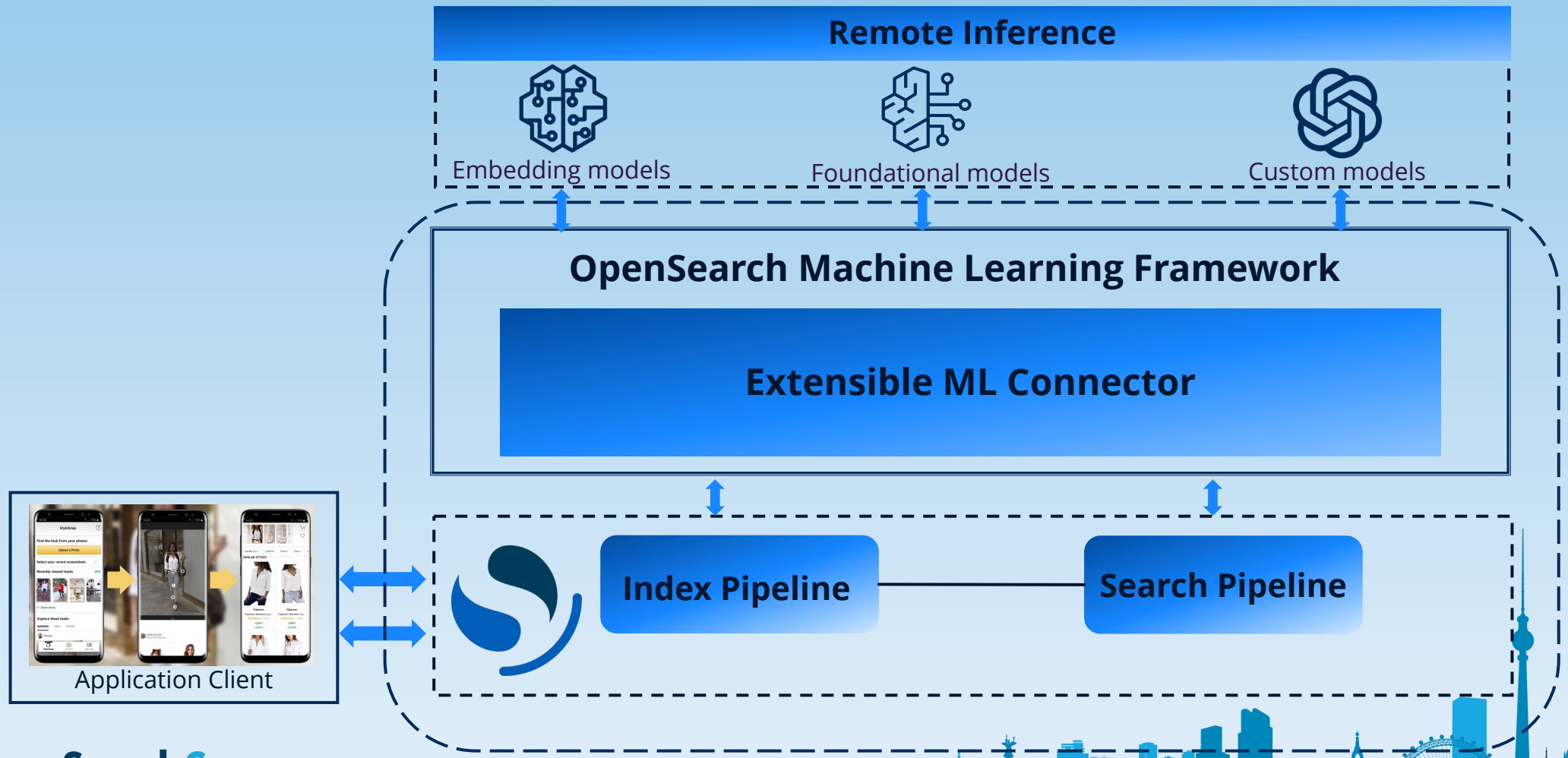
Deliver search results based on vector proximity by comparing vector distances



Vector search architecture



Build with neural search



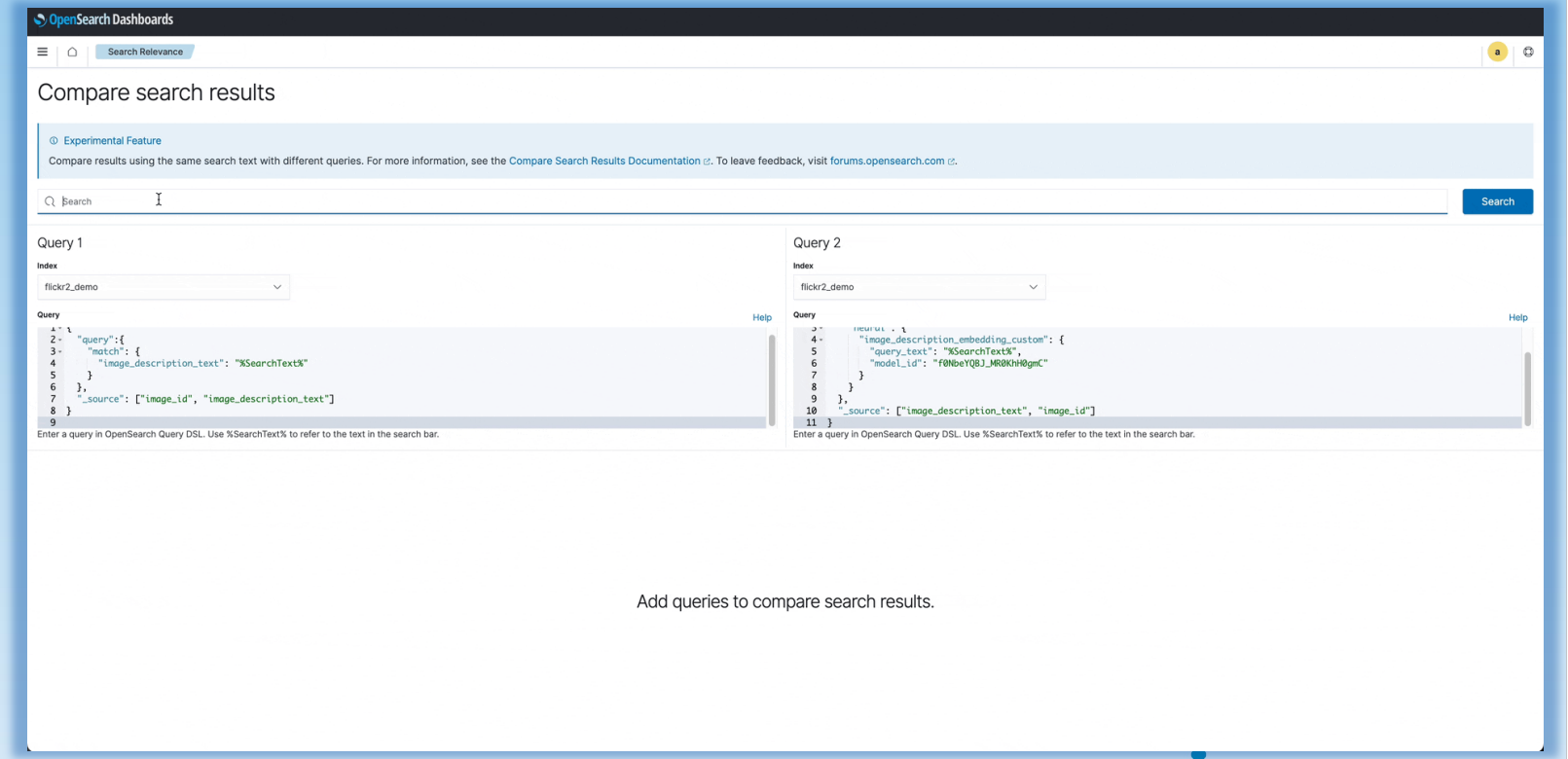
Integrated semantic and hybrid search

Hybrid search

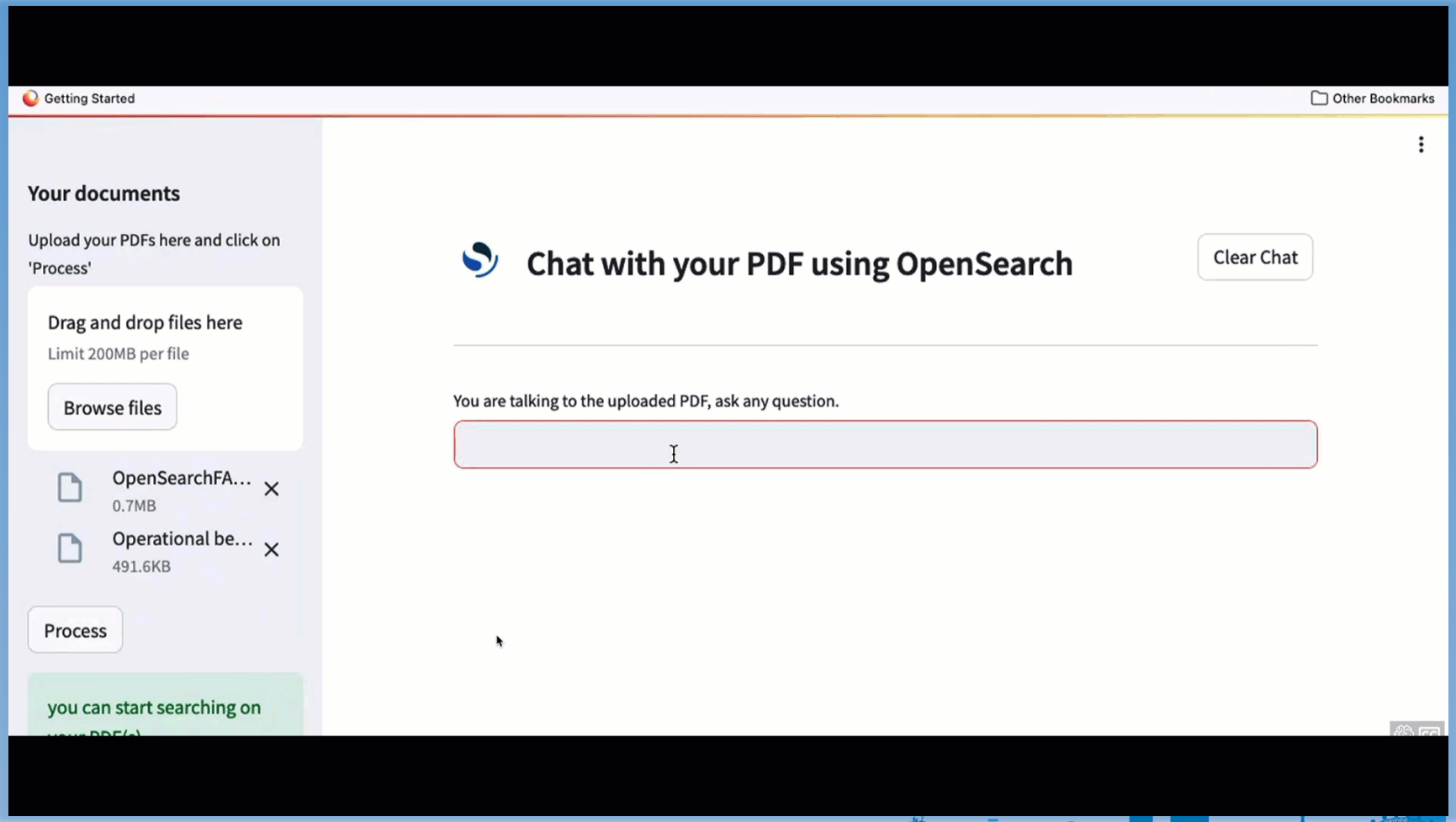
Fine-tuned models

Sparse vector retrieval

Multimodal search



Conversational search



Observability with OpenSearch

Rich log analytics capabilities

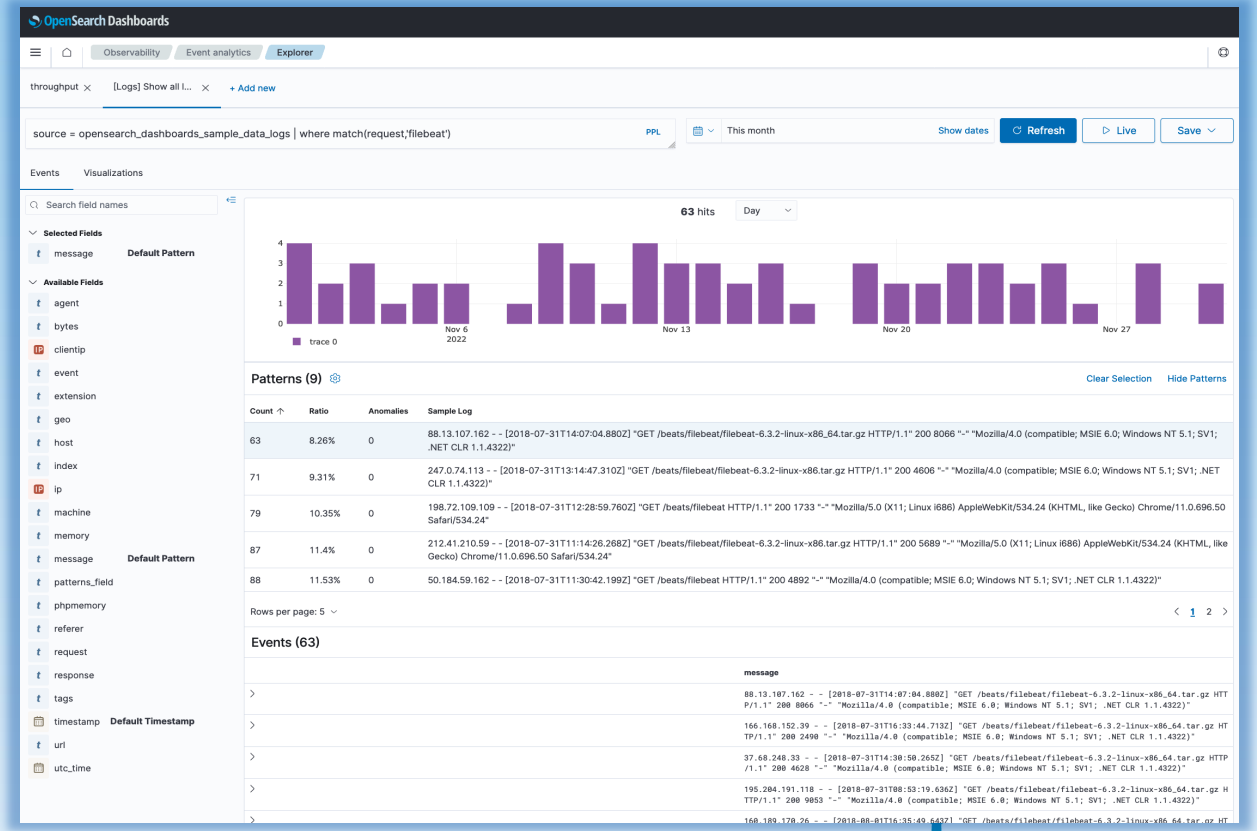
- Built-in anomaly detection & alerts
- Support for OpenTelemetry data
- Log patterns, tailing, surround, and more
- PPL query language

Extend tracing with Jaeger

- Spans, trace groups, and service maps

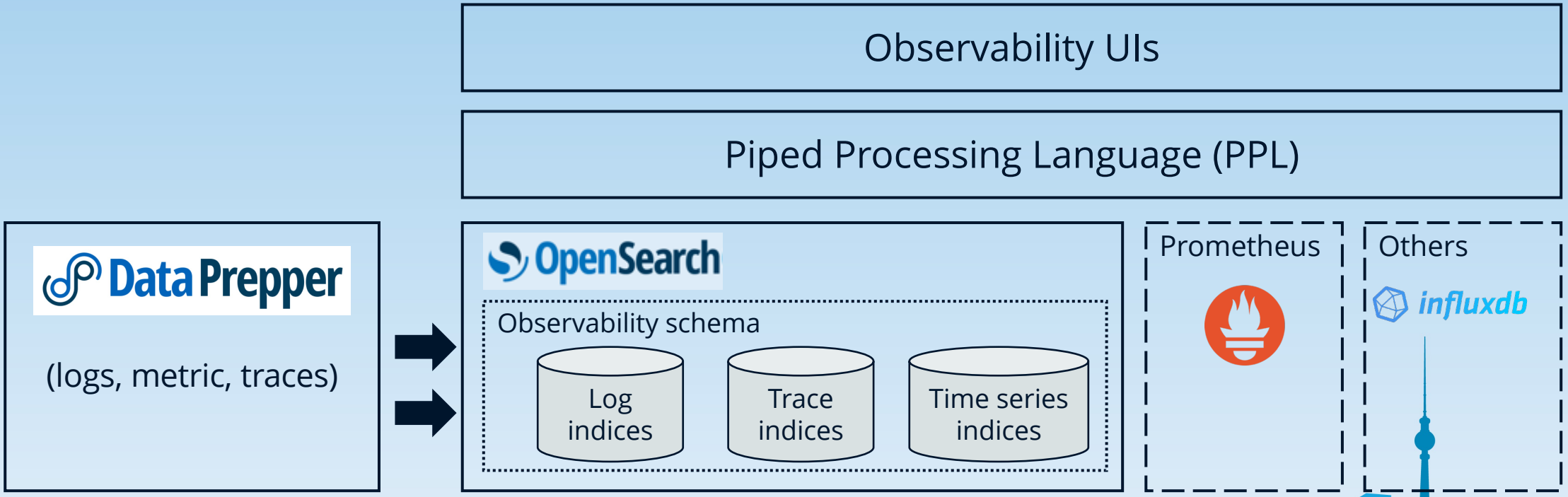
Metric Analysis

- Metric extraction from logs



Observability architecture

Open source | Open standards | Customizable



Anomaly Detection

Detect and mitigate issues faster with anomaly detection

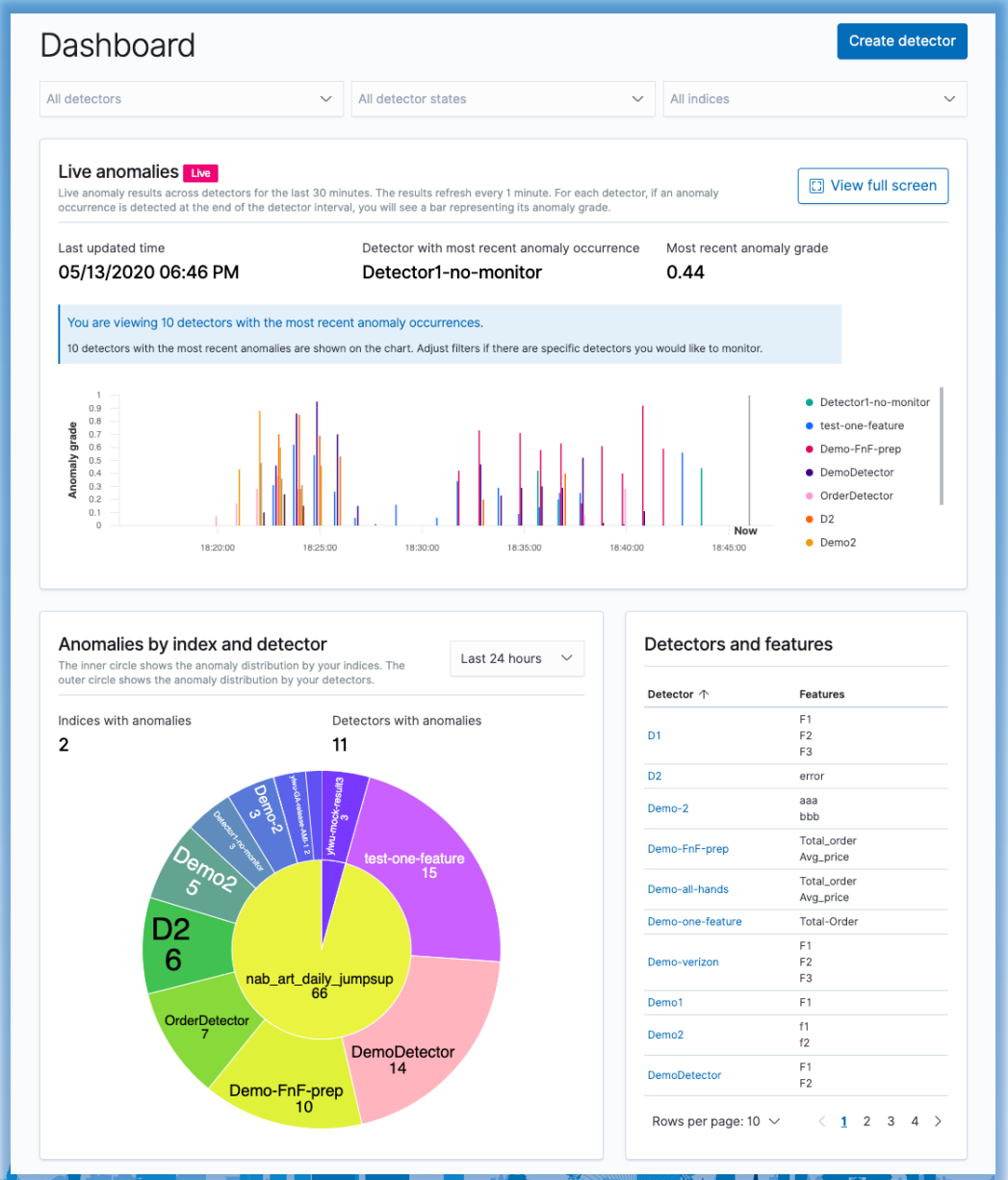
Performant at scale

- Machine learning models are distributed and processed across nodes

Easy to use

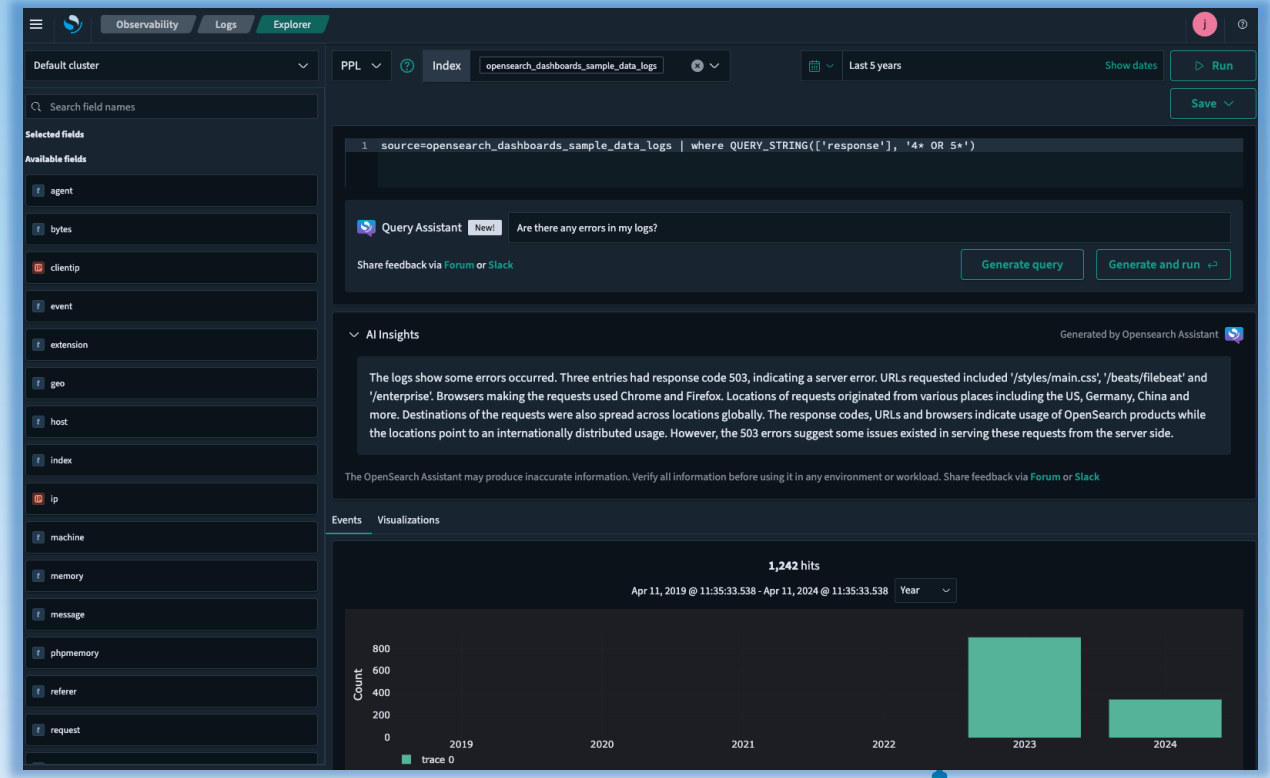
- No machine learning expertise required

Based on Random Cut Forest (RCF)



The OpenSearch Assistant

- Interact using natural language
- Open-source customizable toolkit
- Simplified and rich insights
- Implement your own skills
- Connect to your preferred model



Security Analytics with OpenSearch

Decrease time to response

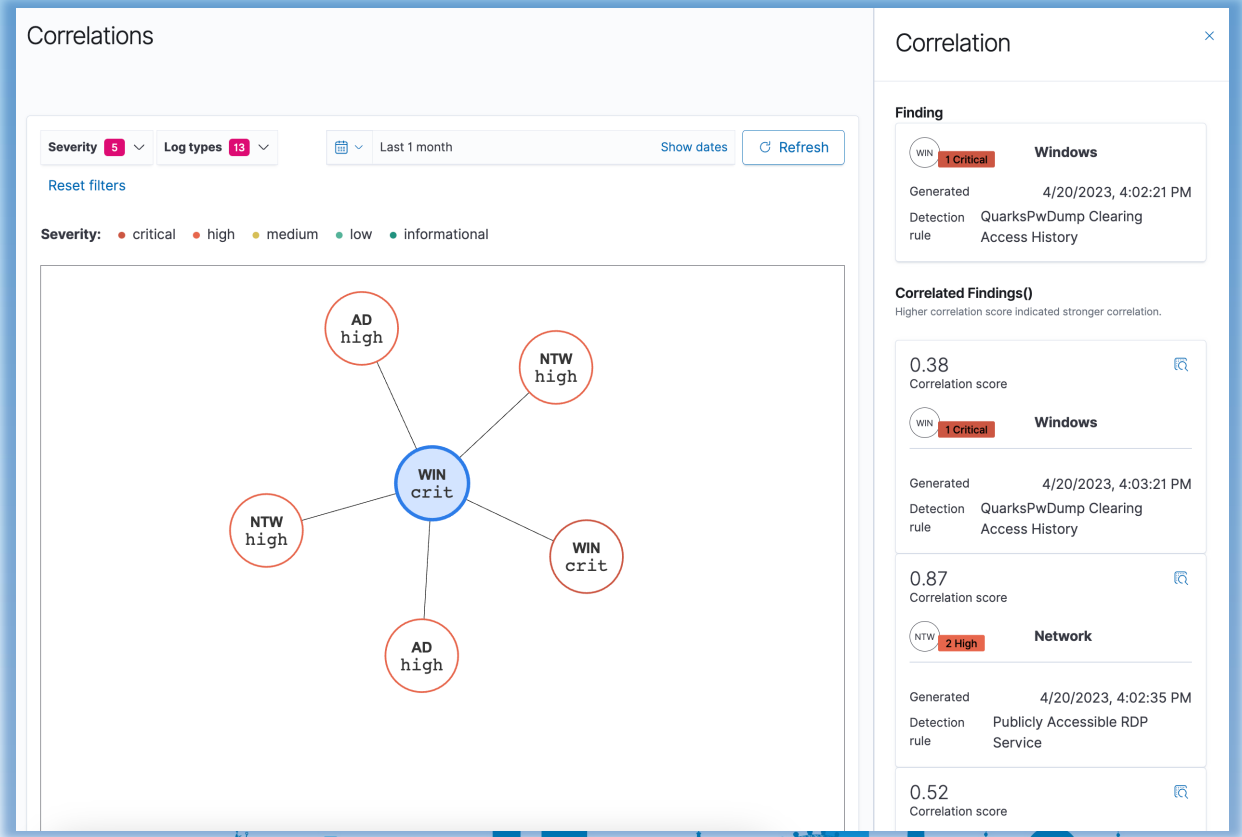
Detect potential threats quickly

2000+ security rules built-in

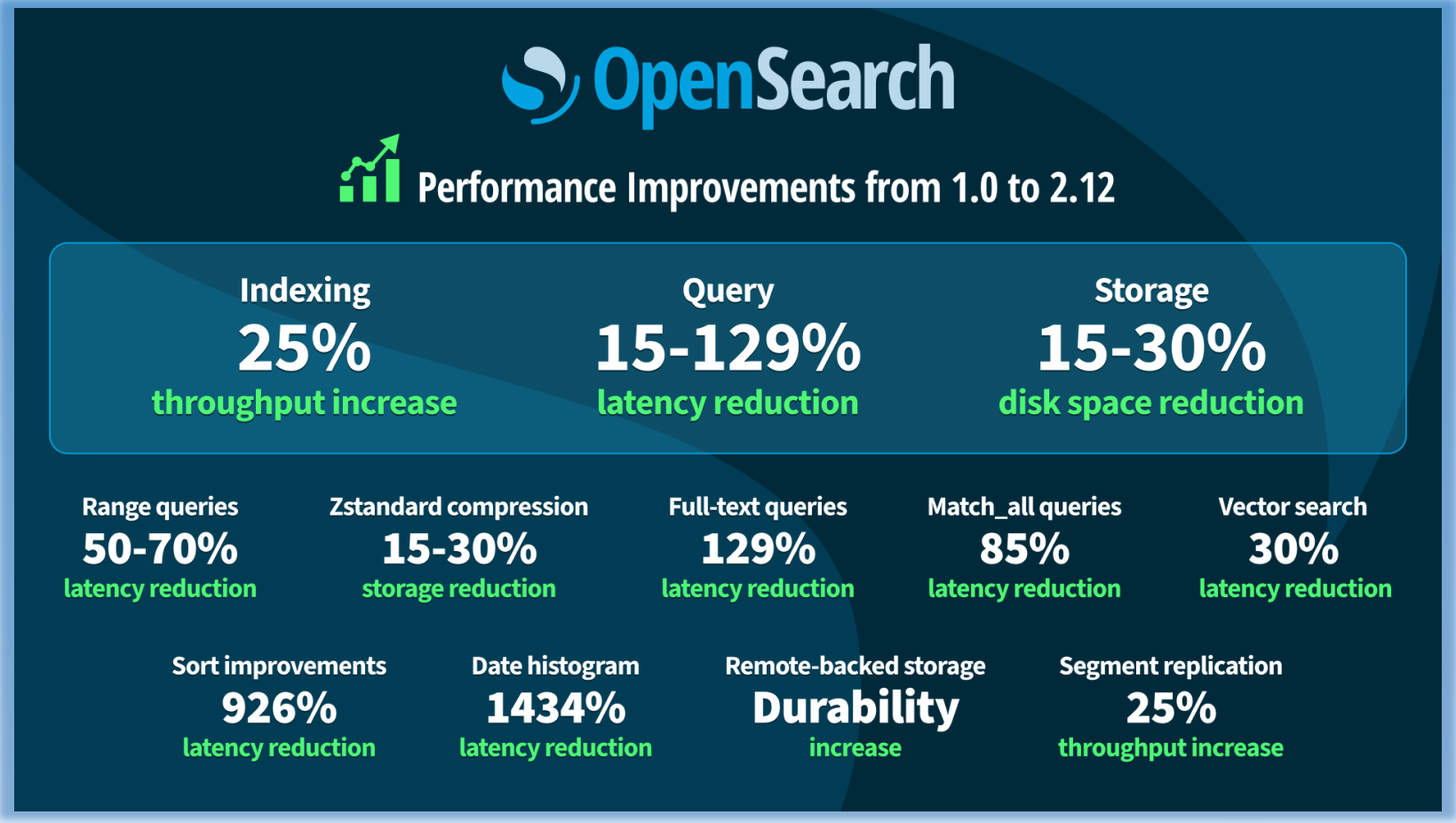
25+ log sources supported

Alert stakeholders

[New v 2.9] Correlation Engine



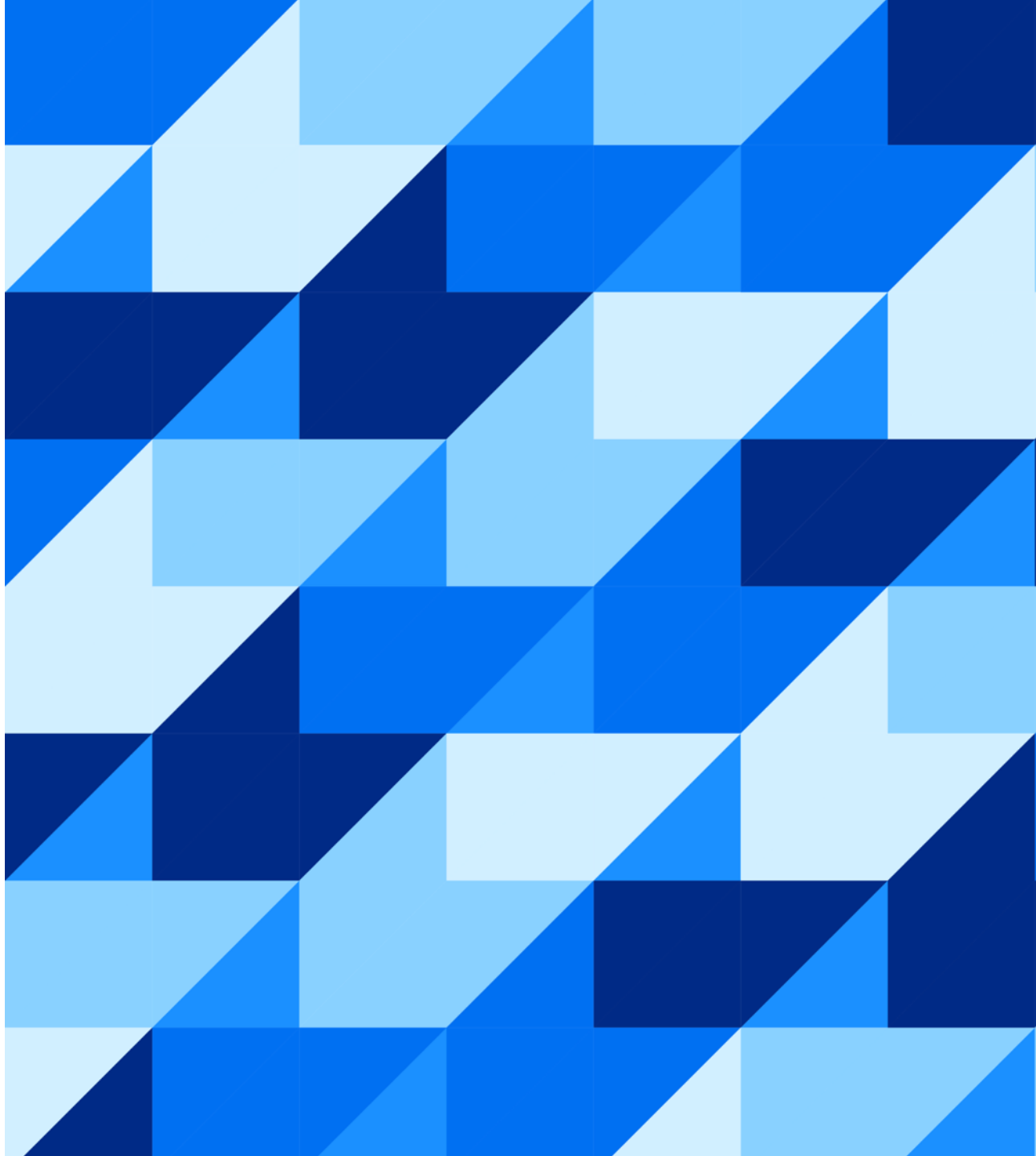
Performance Improvements



Introducing...

Hariharan Gandhi

Product Owner, SAP



OpenSearch @ SAP BTP Logging Service

Hariharan Gandhi, SAP
May 07, 2024

Public

OpenSearch @SAP BTP Logging service

SAP's Business Technology Platform (BTP)

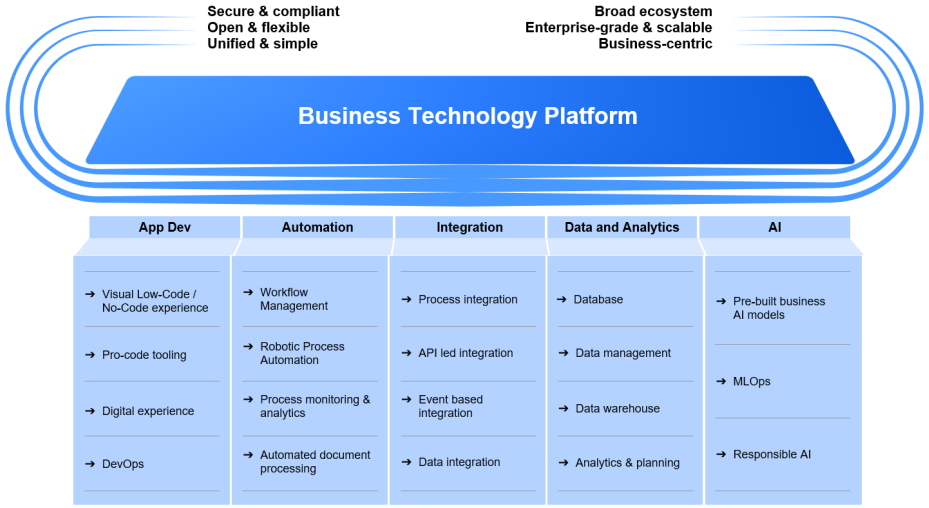
- PaaS Runtimes – CLOUD FOUNDRY Kyma
- Services
- Customers building Business apps

SAP Service Teams

External Customer

SAP Platform Teams

SAP BTP Cloud Capabilities



Need: Hosted Observability solution

- Reduced Operational Overhead
- Pre-defined contents tailored for SAP BTP runtimes – CF & Kyma
- Compliance needs: Retention, encryption, privacy etc.

Our team: **BTP Logging service**

– Offers hosted observability as a services
– Consumed by the platform's runtimes and services
– Consumed by customers of BTP
– OTel shared semantics beyond BTP

Powered by OpenSearch

Product Owner - Logging Service @ SAP's Business Technology Platform (BTP)

Distributed Systems Engineer (M.Sc. Technical University of Darmstadt)

Over 11 years of experience in the Software Industry, with a steadfast focus on SAP's logging/observability needs for the past 8 years

In his previous role as an Architect, Hariharan spearheaded the replatforming of OpenSearch-based logging services to Kubernetes at scale

Hariharan Gandhi
Product Owner
Logging Service

Agenda

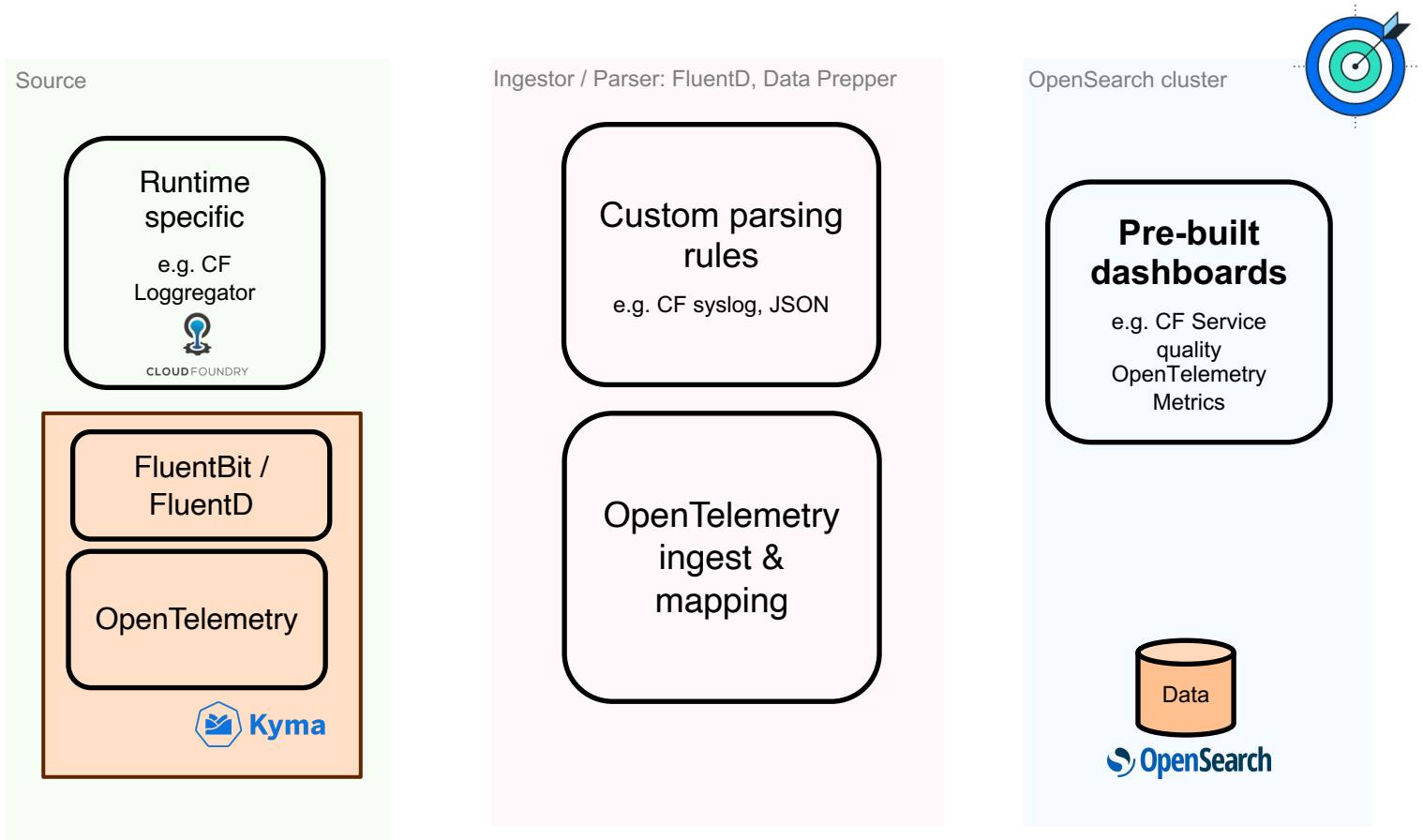
Our OpenSearch Setup

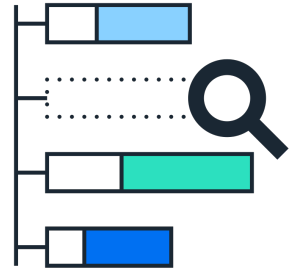
Use cases

Challenges

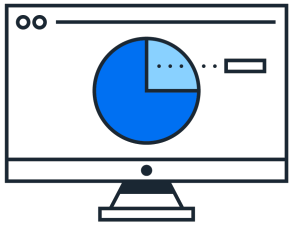
Team & OS Community

OpenSearch Setup





SAP specific correlation



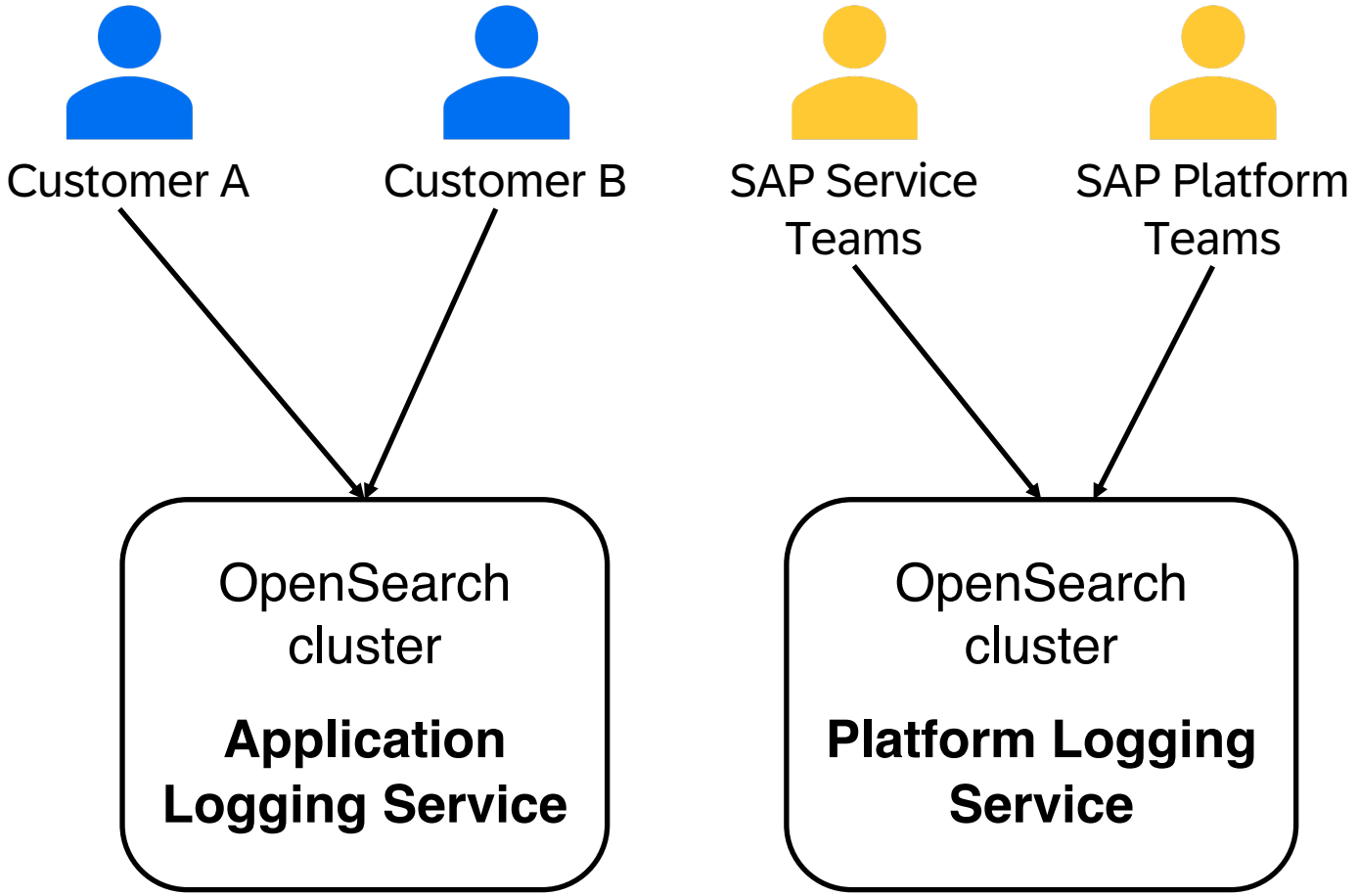
Pre-defined contents

OpenSearch Setup 1

Large central clusters on VMs

Shared stack

Multi-tenancy on top



OpenSearch Setup 1: Large central clusters on VMs

Shared stack

50+ Global

Both for runtimes and end use application logs

200k - 1Mi

Events per seconds

~100 Billion

Events: logs & metrics stored for 14 days

~1k

CPUs on Largest Cluster – Data Nodes

~1.5TB

Memory on Largest Cluster – Data Nodes

0.5PB

60-70% usage. Retention 14 days

OpenSearch Setup 1: Large central clusters on VMs

Shared stack

50+ Global
Both for runtimes and end use application logs

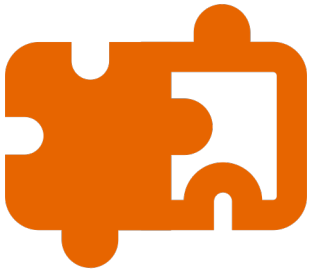
200k - 1Mi
Events per seconds

~100 Billion
Events: logs & metrics stored for 14 days

~1k
CPUs on Largest Cluster – Data Nodes

~1.5TB
Memory on Largest Cluster – Data Nodes

0.5PB
60-70% usage. Retention 14 days



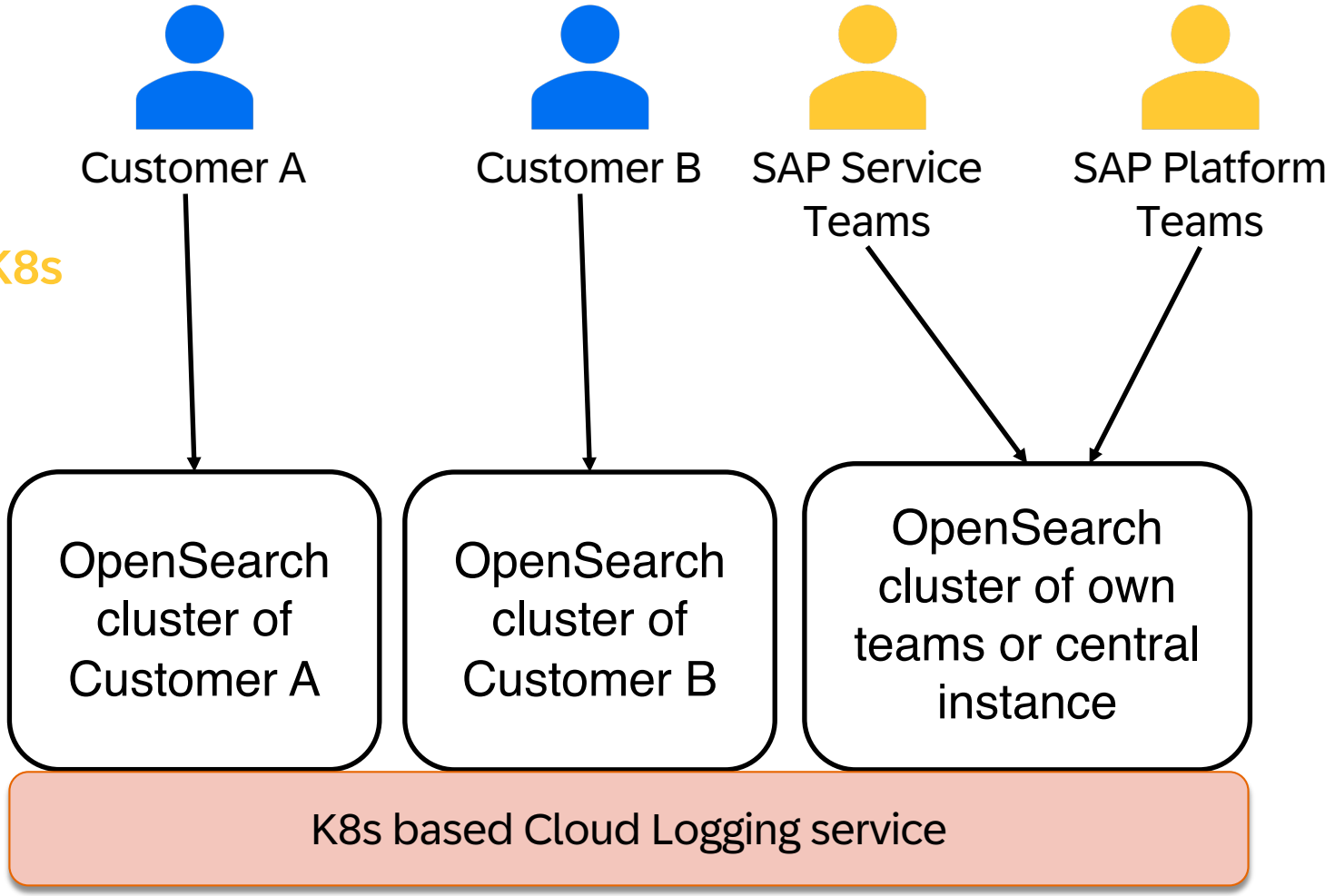
- Complex Multitenancy setup
- Limited exposure of OpenSearch features
- No flexibility in config: retention, dashboards etc

OpenSearch Setup 2

Dedicated OpenSearch clusters on K8s

Dedicated stack

BYO IDP



OpenSearch Setup 2: Dedicated clusters on K8s

>7400

OpenSearch Clusters

50000

Persistent Volumes

>3PB Storage

(Hot) With Auto-Scaling

>1700 EC2

Spread equally across 3 Zones

50K CPU / 220 TB of Memory

35 K8s

Clusters using Gardener

10 Global

Deployments

OpenSearch Setup 2: Dedicated clusters on K8s

>7400

OpenSearch Clusters

50000

Persistent Volumes

>3PB Storage

With Auto-Scaling

>1700 EC2

Spread equally across 3 Zones

50K CPU / 220 TB of Memory

35 K8s

Clusters using Gardener

10 Global

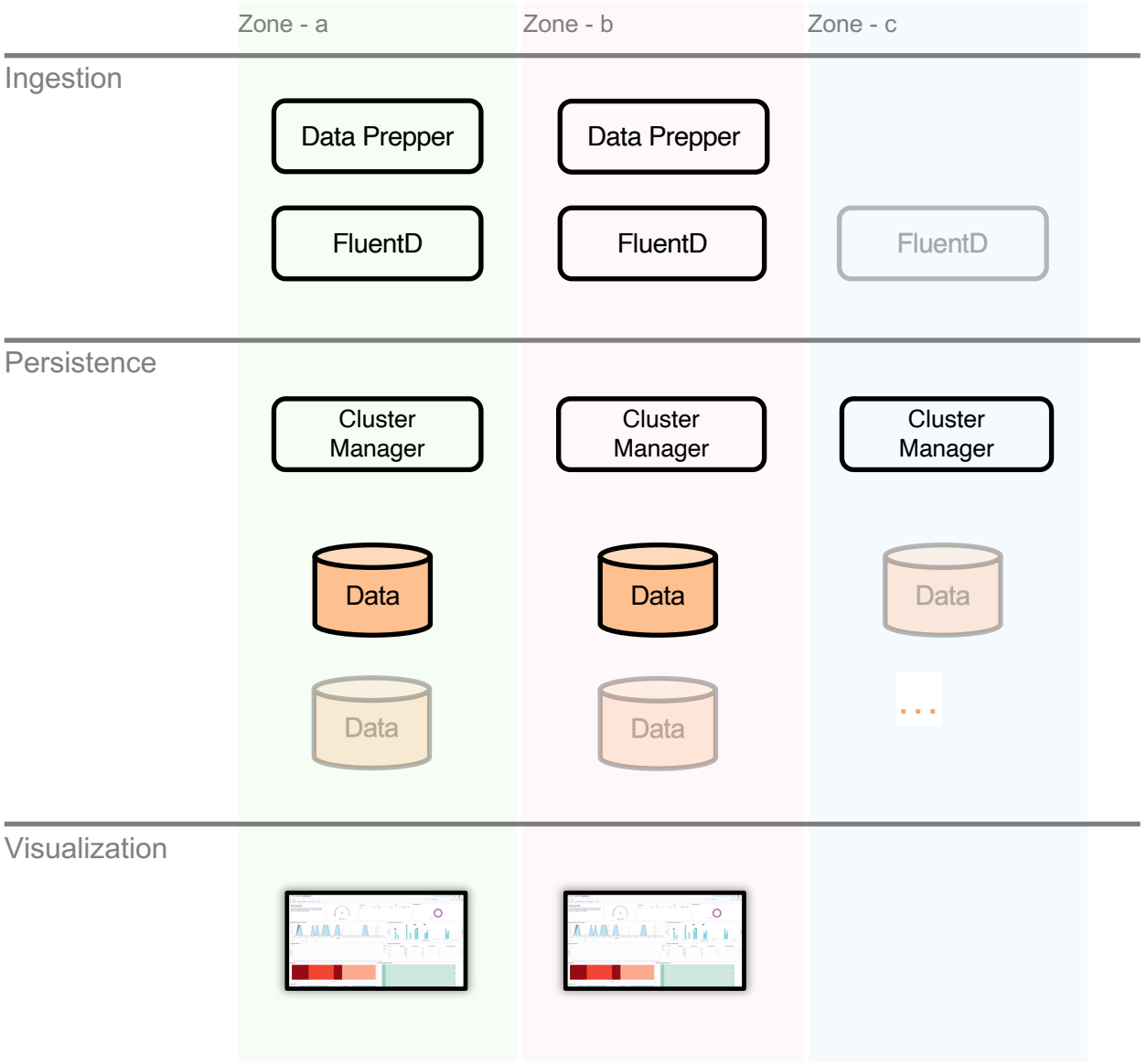
Deployments

- Complex operational setup

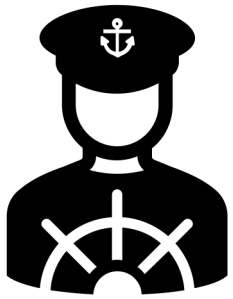
OpenSearch Setup 2

Control Components

Monitoring Components



OpenSearch Setup 2



Control Components

Monitoring Components

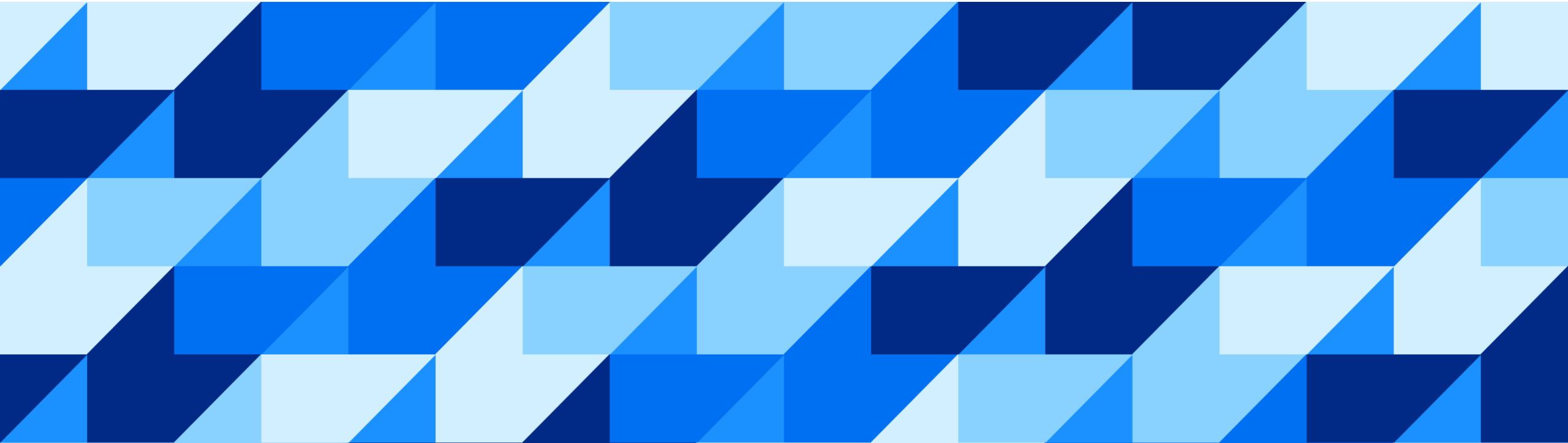
- Instance Lifecycle: Provisioning, Binding, Configuration
- Auto Scaling based on Disk usage w/ safe draining
- Troubleshooting / Self Healing
- Contents (dashboards, alerts) Manager
- Backup and Restore of user contents
- Certificate and Credential Rotation
- Metering / usage reporting
- Health:
  - Availability reporting
  - Self Monitoring
  - Alerting

OpenSearch Setup 2: single instance to complete landscape

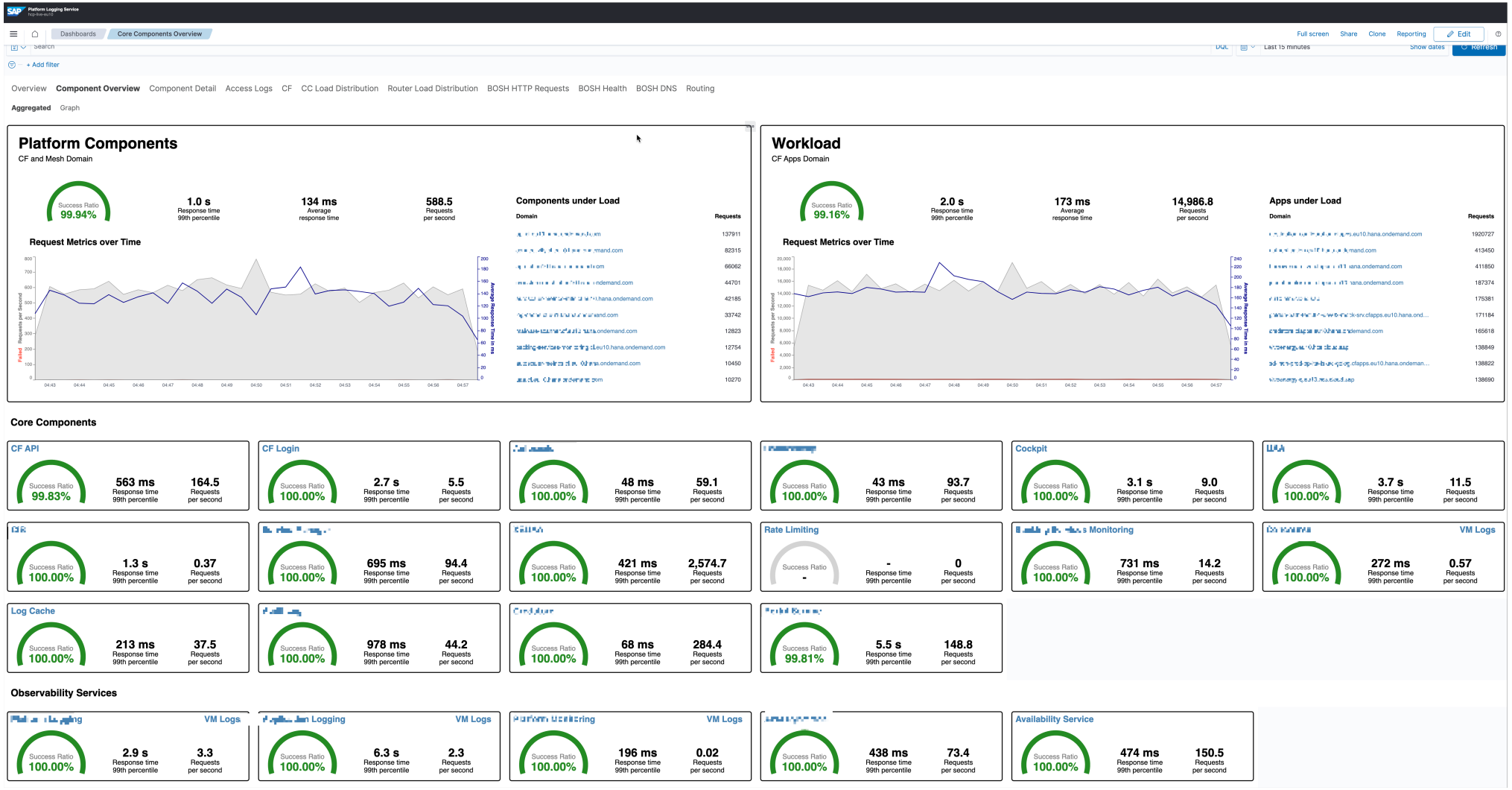
Control Components

Monitoring Components

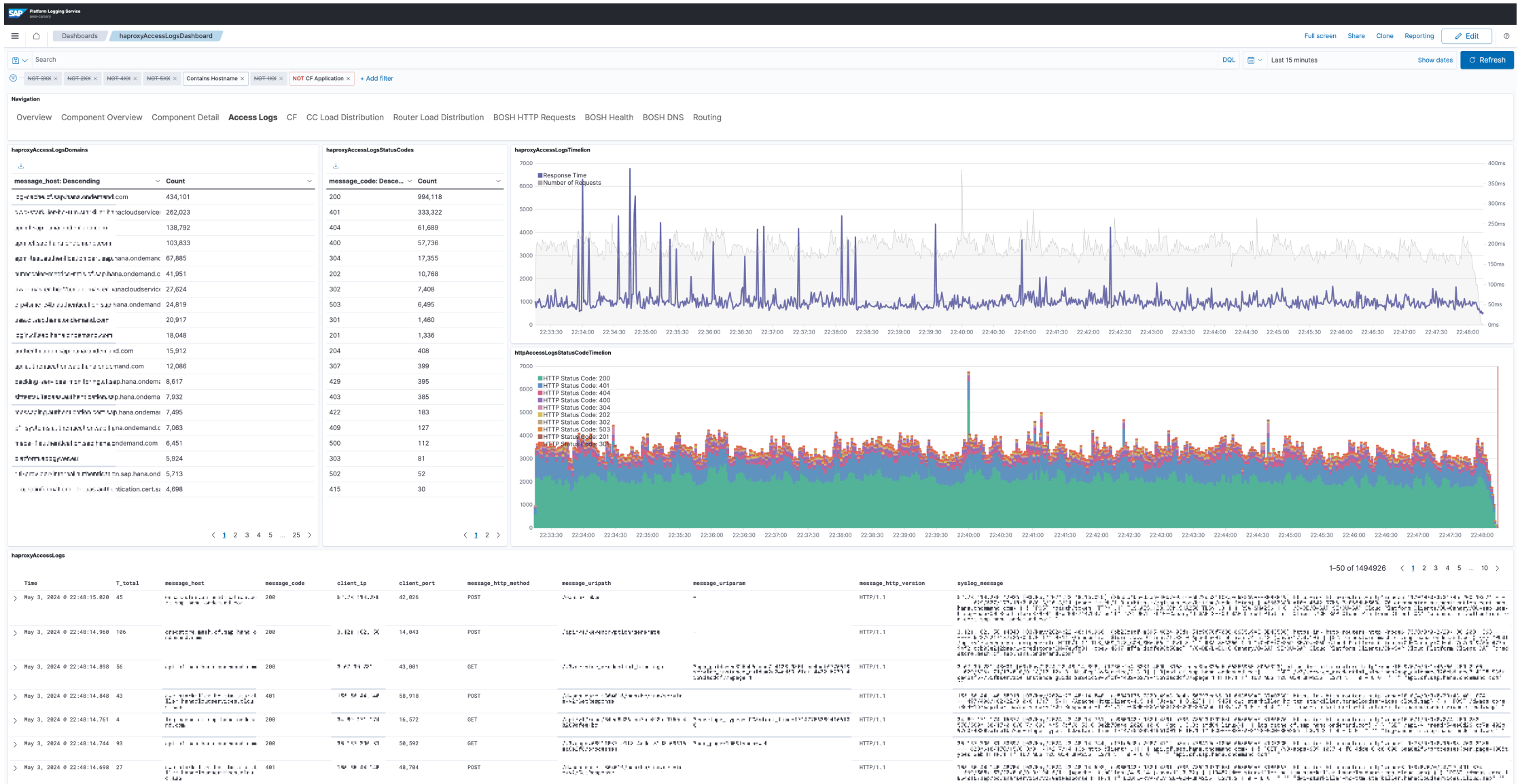
Use Cases



Log Analytics Use cases: RCA



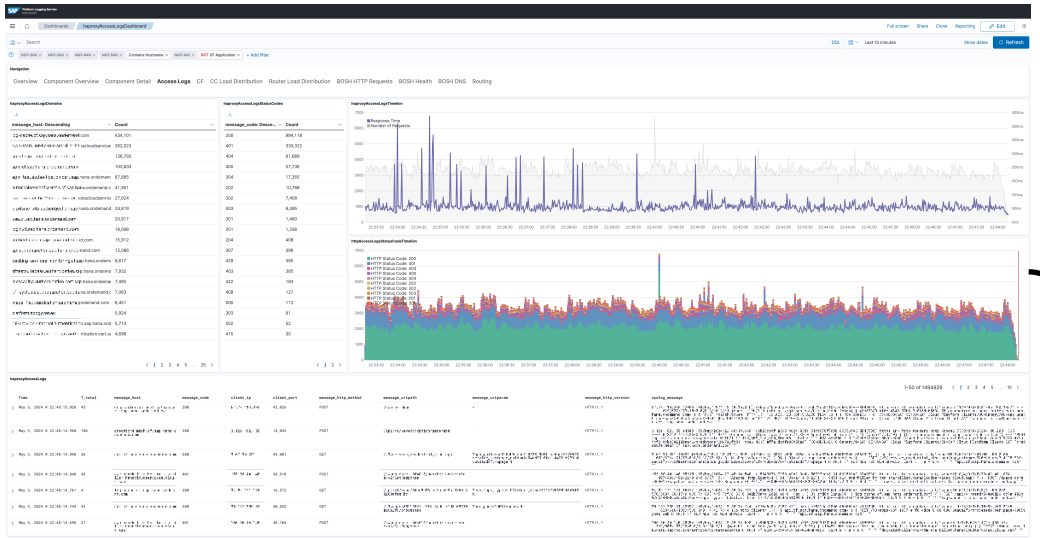
Log Analytics Use cases: RCA

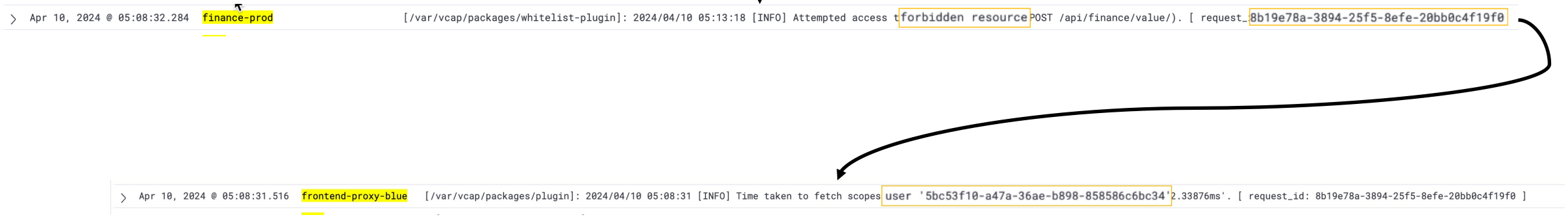


Log Analytics Use cases: RCA

SAP Service Teams

SAP Platform Teams

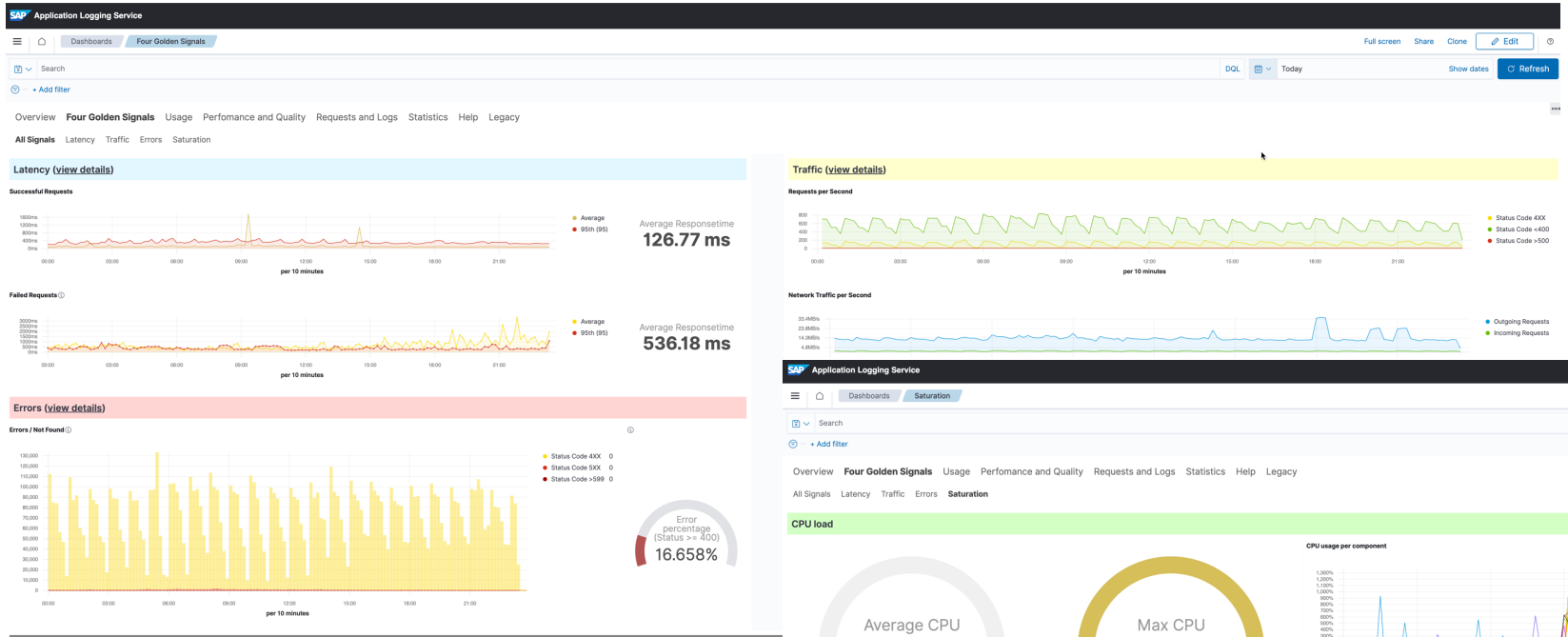


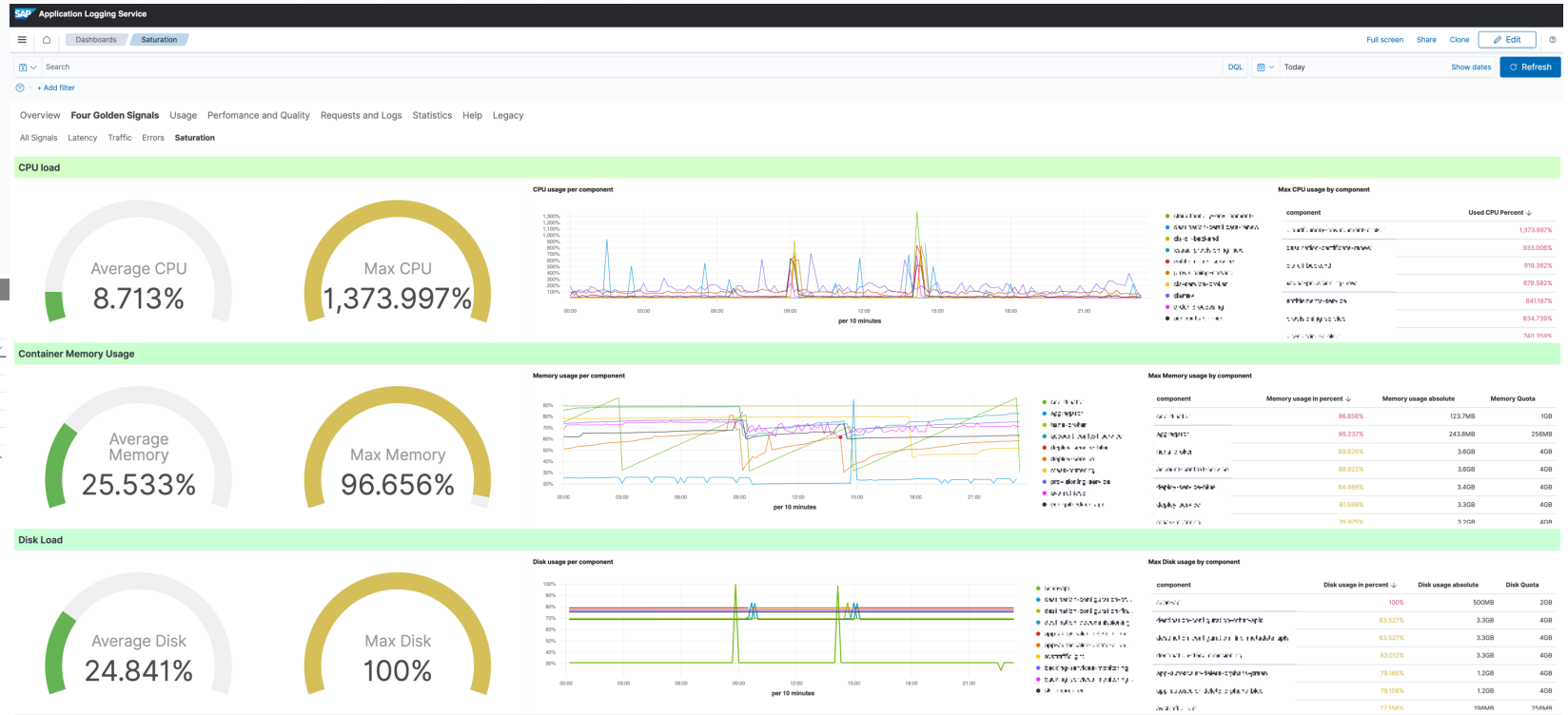


SAP Service Teams

External Customer

Use cases: Service Performance / Quality





SAP Service Teams | SAP Platform Teams | External Customer

Use cases w/ Open Telemetry: Logs, Metrics, Traces

SAP Cloud Logging cls

Dashboard / [OTel] K8s Container Metrics

Full screen | Share | Clone | Reporting | Edit

Namespace Name	#Pods
kyma-system	23
example	20
istio-system	3
test	1
	47

Export: Raw Formatted

Pod Name	#Containers
otel-accountingservice-6f86dcf4b6-872kr	2
otel-adservice-7bd4c79675-wfml9	2
otel-cartservice-55447ff54f-cd2dr	2
otel-checkoutservice-5b8df4cccd-m85wv	2
otel-currencyservice-c9dbd8b6f-thcj7	2
otel-emailservice-7795877686-r2lrt	2
otel-featureflagservice-775bc88578-jdtpf	2
otel-ffspostgres-6dff6c5f65-kq4v8	2
otel-frauddetectionservice-5955696b6f-hq96b	2
otel-frontend-7c5dc94ff8-6wrqc	2
	51

Export: Raw Formatted

1 2 3 4 »

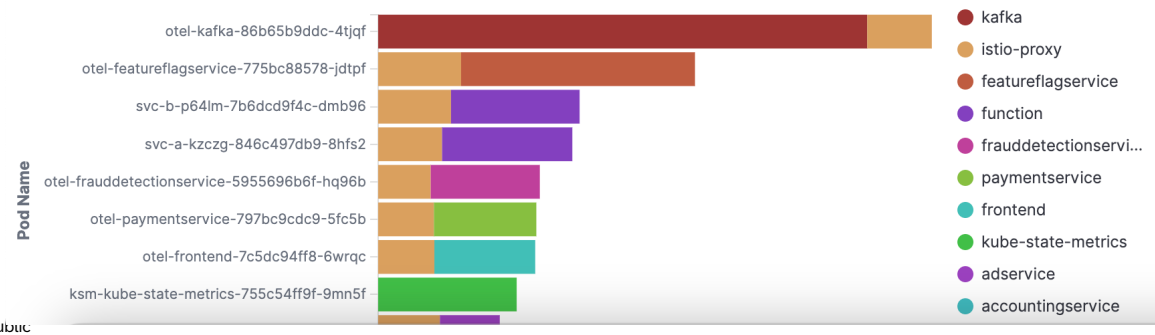
Container Name	#Metrics
istio-proxy	20
accountingservice	11
adservice	11
cartservice	11
checkoutservice	11
currencyservice	11
emailservice	11
featureflagservice	11
ffspostgres	11
frauddetectionservice	11
	240

Export: Raw Formatted

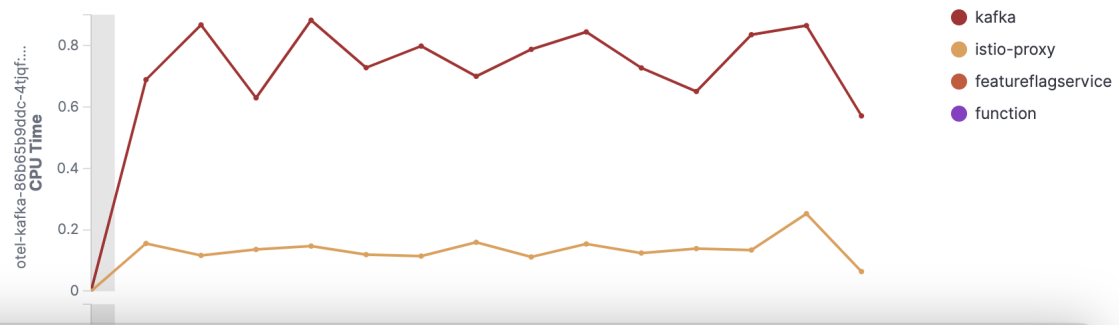
1 2 3 »

CPU

Top 10 Average CPU Usage

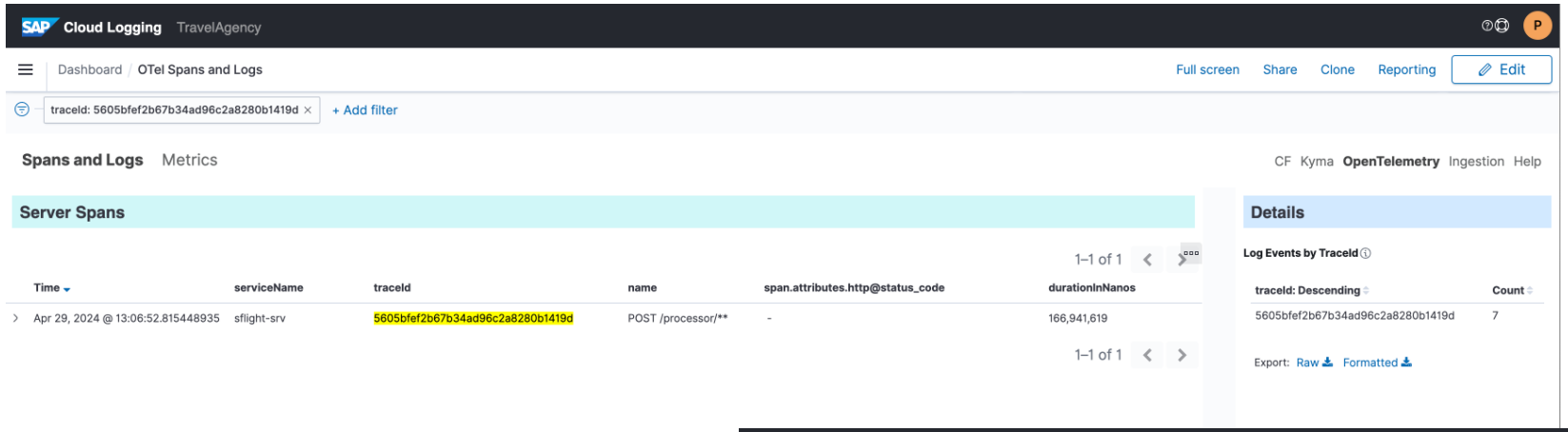


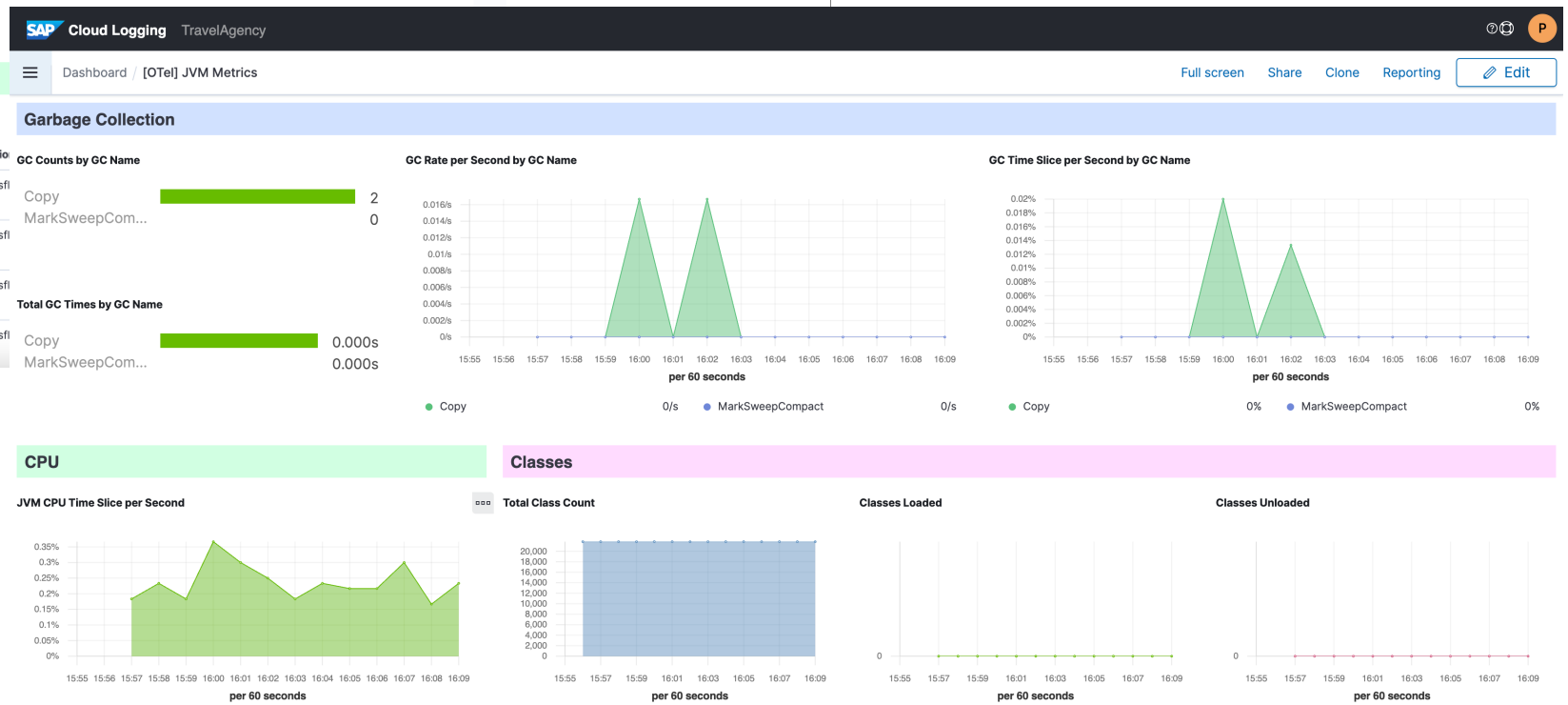
Top 3 Pods Container CPU Time



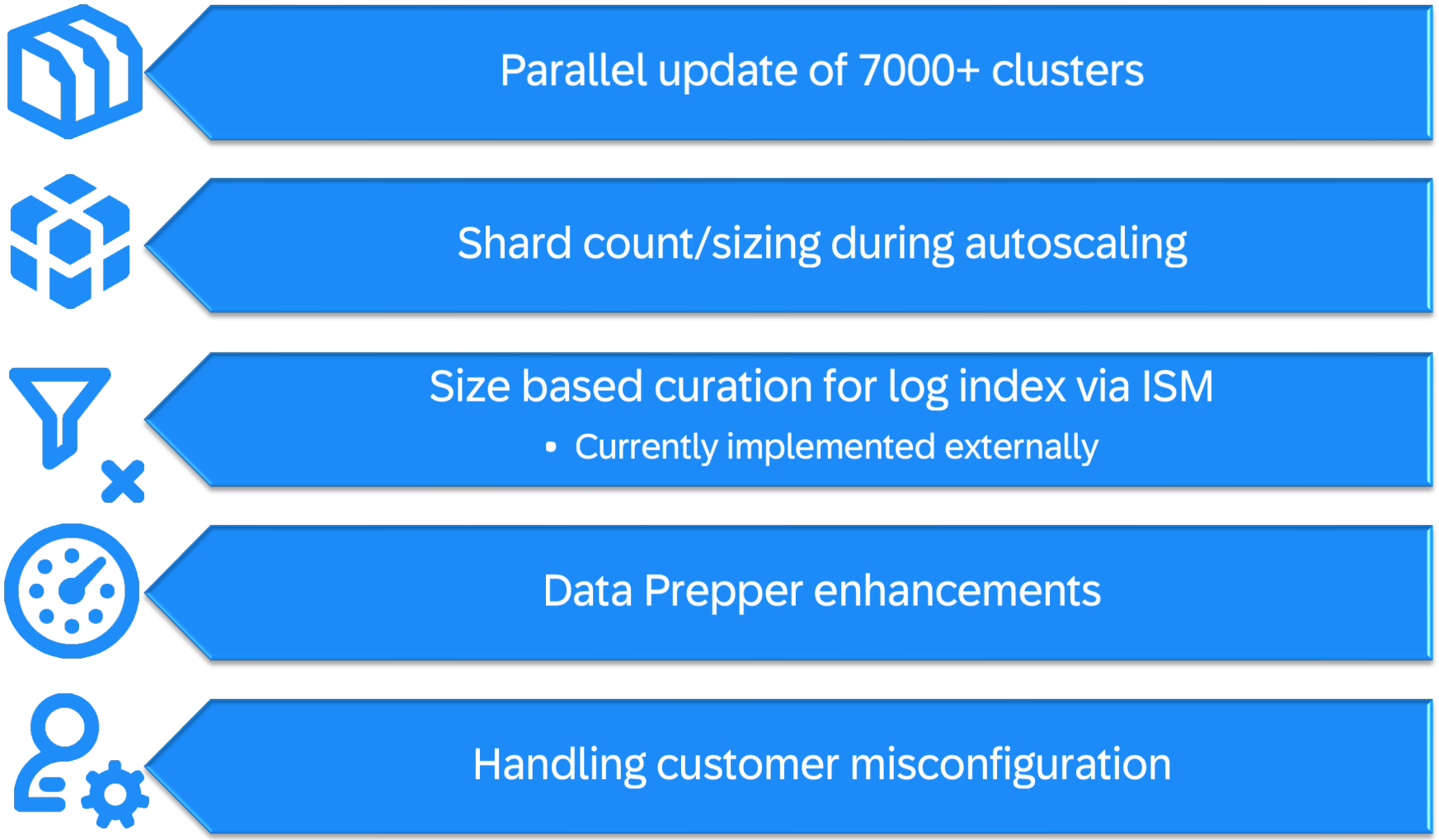
Use cases w/ Open Telemetry: Logs, Metrics, Traces

SAP Service Teams | SAP Platform Teams | External Customer





Known Challenges

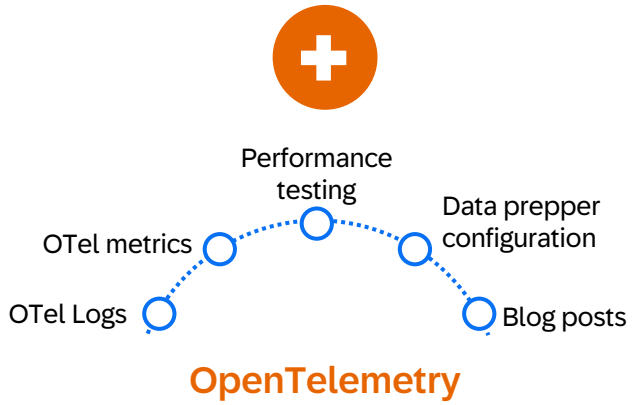


SAP Cloud Logging team + OpenSearch community

SAP customers expect a unified, business-centric and open SAP Business Technology Platform. Our observability strategy uses Elasticsearch as a major enabler. OpenSearch provides a true open source path and community-driven approach to move this forward.

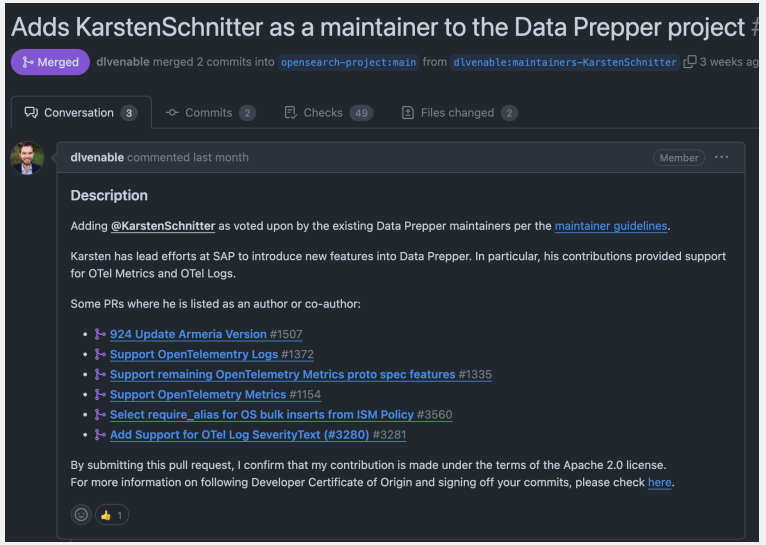
JAN SCHAFFNER
SVP AND HEAD OF BTP FOUNDATIONAL PLANE
SAP

GitHub Issues	Technical collaboration	Semantics contribution	Sharing results	FIPS Compliance



Tech Talk: How SAP analyzes OpenTelemetry signals using Data Prepper

Maintainer for Data Prepper



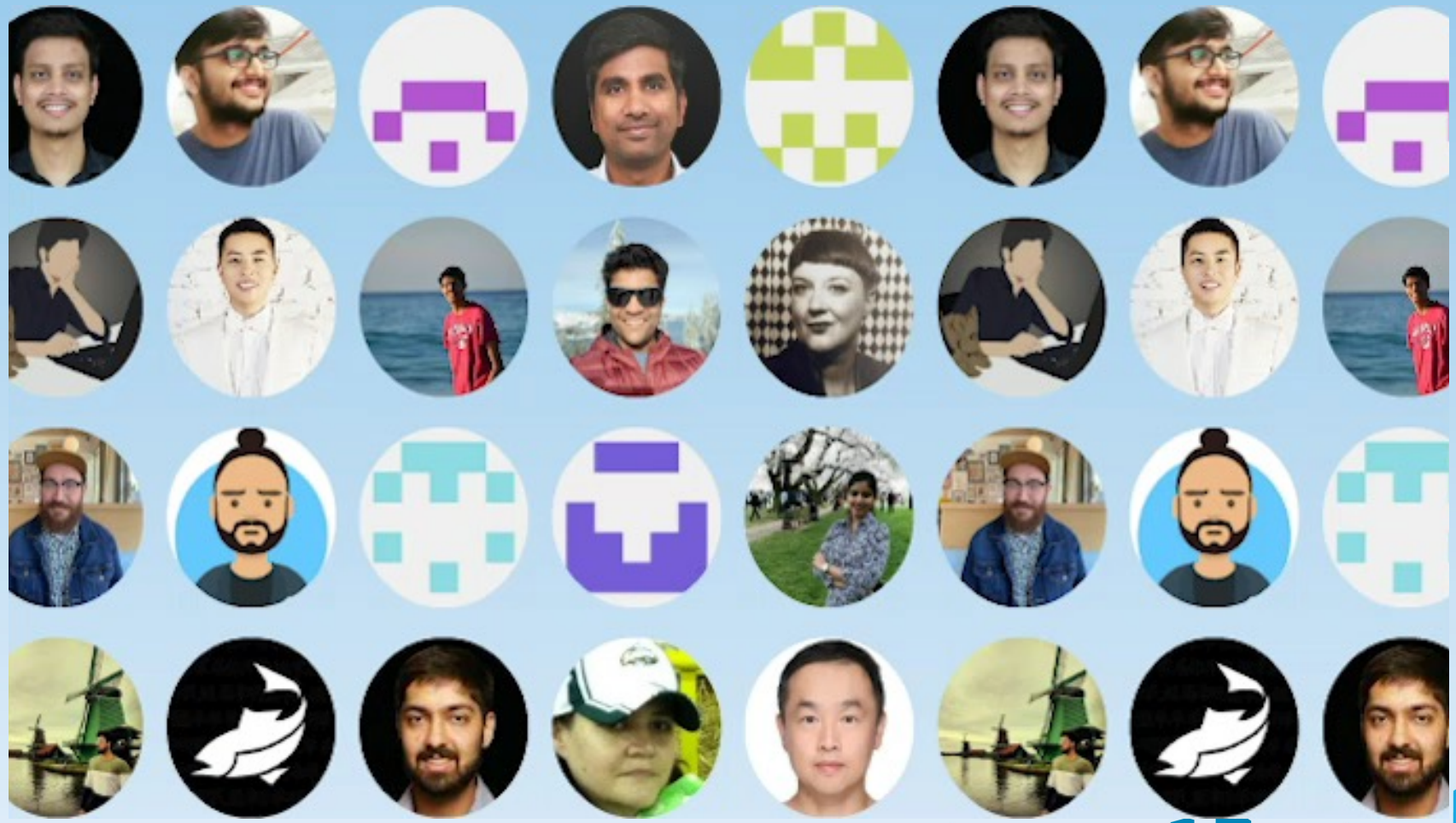
Thank you

Contact information:

Hariharan Gandhi
hariharan.gandhi@sap.com

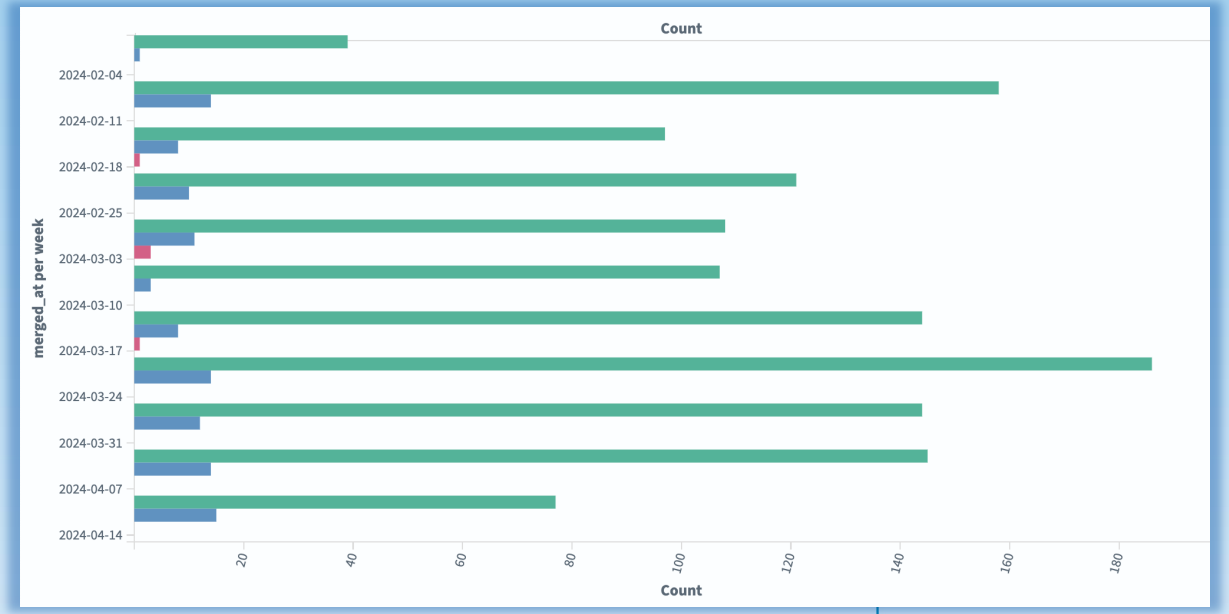
@hariharangandhi

Community-powered innovation

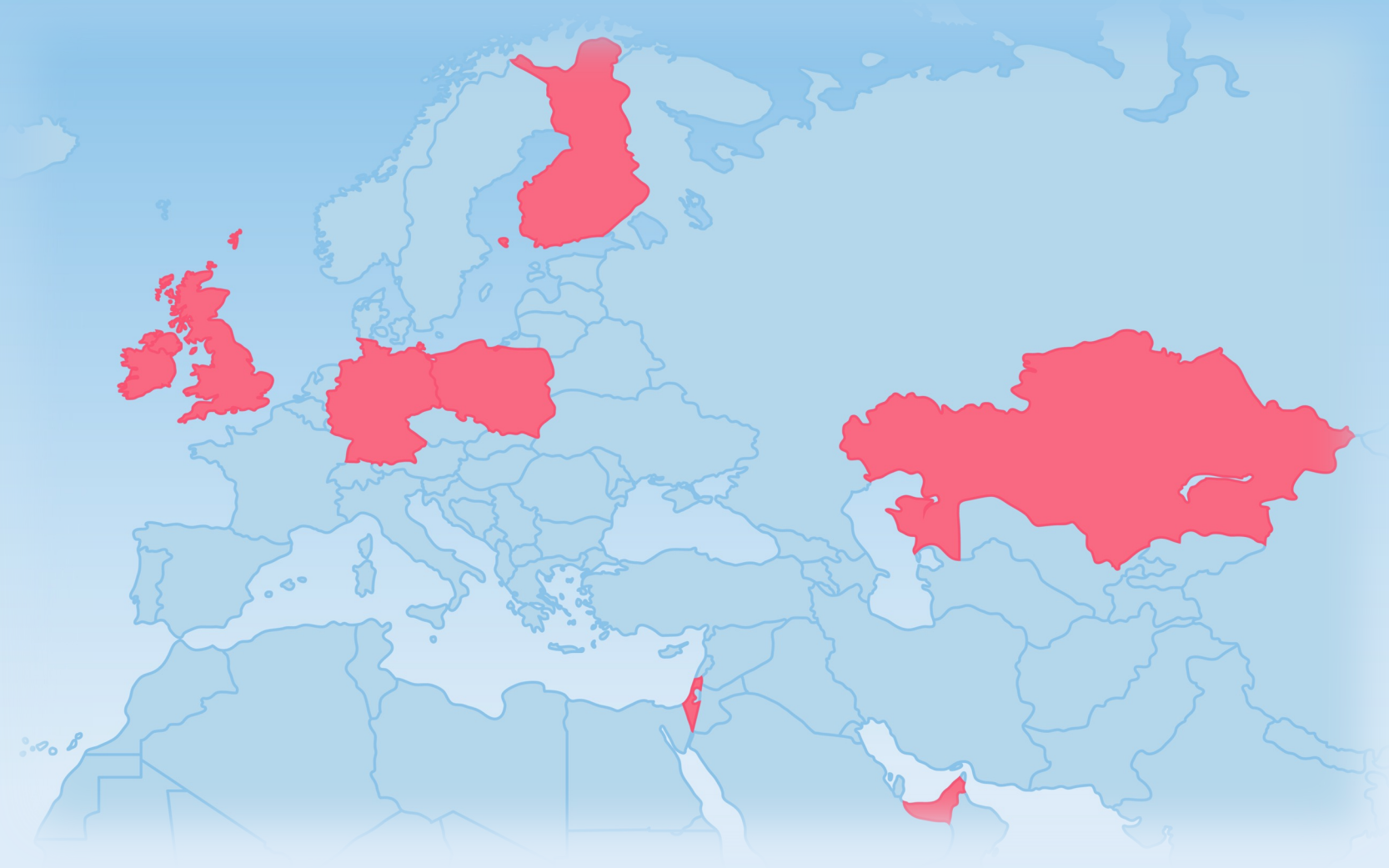


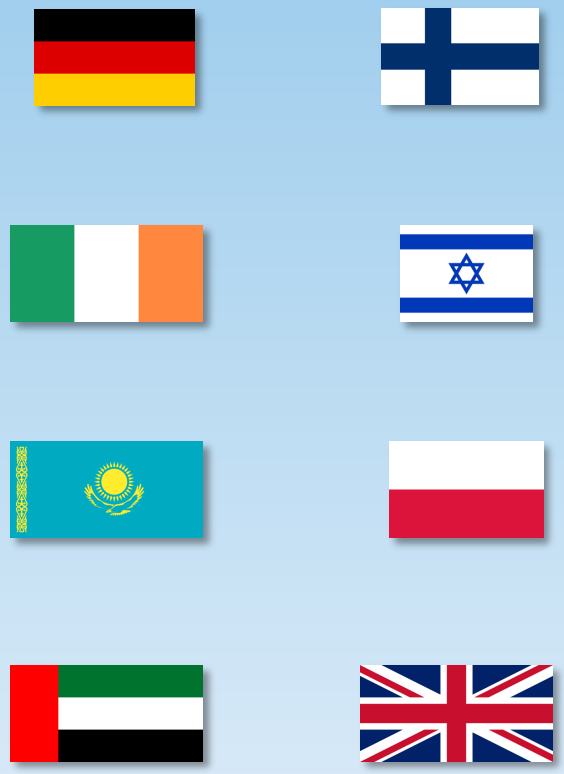
Community contributions

- 2000+ Contributors outside of AWS
  - *Oracle, Aiven, Aryn, ByteDance, SAP, Intel, IBM, and more*
- Weekly average of 100+ contributions

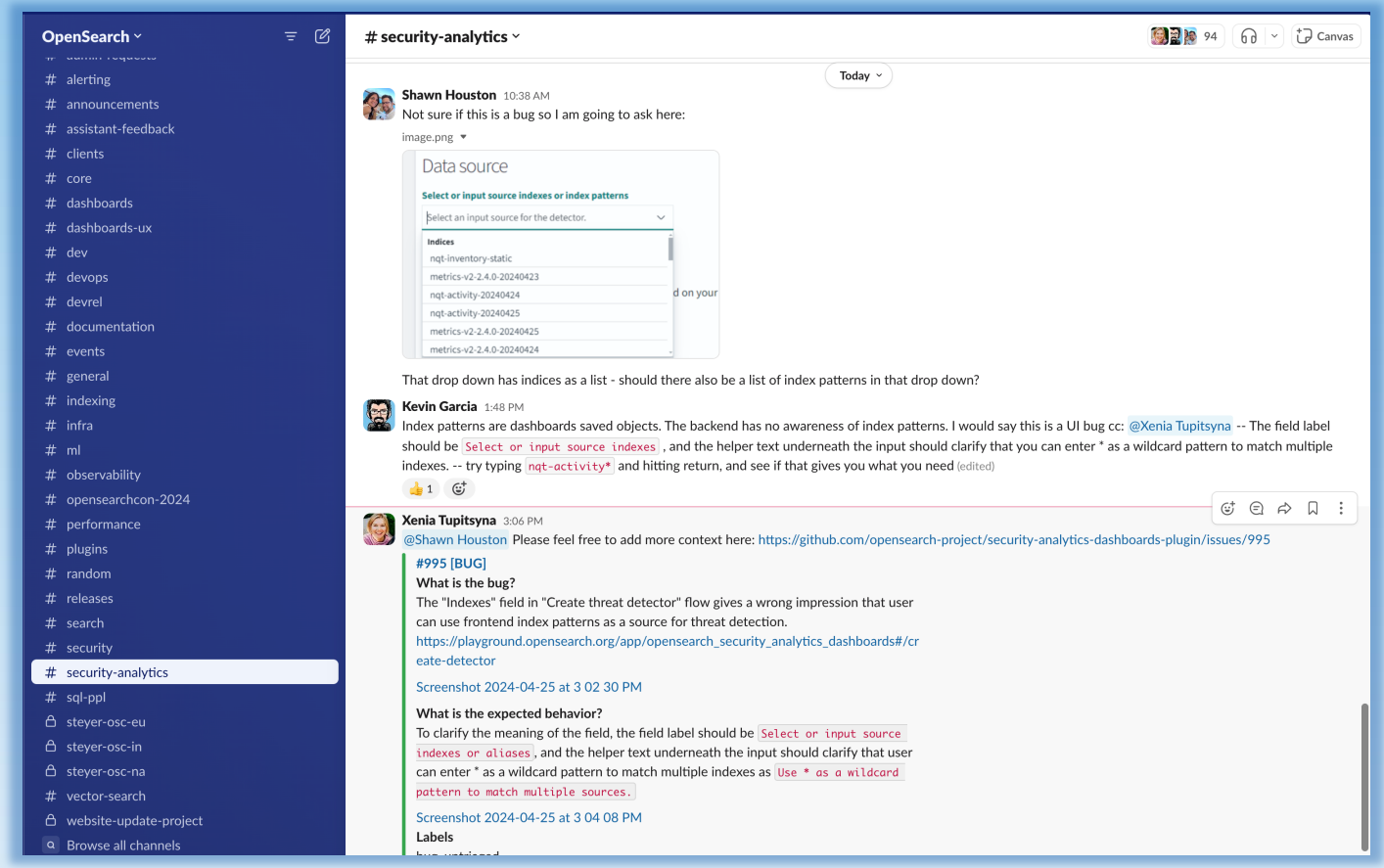


Community contributions





Community-led dialogue



Global gatherings

Downloads to date

70+ Partners and Counting

A more open OpenSearch

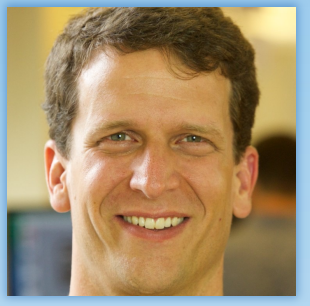
Amitai Stern


Anandhi Bumstead


Andriy Redko


Charlotte Henkle


Eli Fisher


Grant Ingersoll


Jonah Kowall


Kris Freedain


Mehul A. Shah


Nick Knize


Nithya Ruff


Samuel Herman

OpenSearchCon

EUROPE | Berlin 2024

THANK YOU | DANKE

EUROPE | BERLIN 2024