

Unlocking Multi-Cloud Observability

The Case Study of EO4EU Project's Observability Platform

Armagan Karatosun – ECMWF Francesco Maria Cultrera – CINECA Lucía Rodríguez Muñoz – CINECA

Who are we?



19 March 2024 | Paris, France







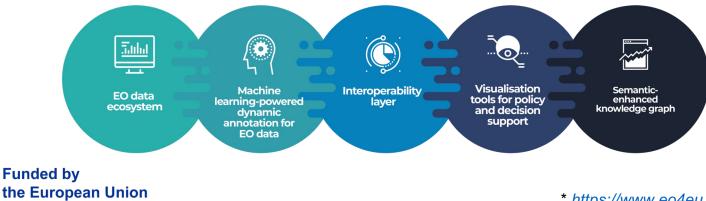
Armagan Karatosun Cloud Computing Engineer ECMWF

Introduction: EO4EU Project

Earth Observation for European Union (EO4EU)* is a **European Union-funded** initiative that aims at creating an advanced platform for searching, discovering, processing and analyzing Earth Observation (EO) data and it is currently being developed by a consortium of European partners.

The EO4EU platform is based on a series of innovative technologies which allow to:

- Access** EO data from different sources (e.g., Copernicus, Galileo, ECMWF)
- Support a sophisticated representation of data through a semantic-enhanced knowledge graph
- Use machine learning (ML) from marketplace to EO data processing
- Visualize EO data through easy-to-use graphical interfaces and extended reality (XR) web applications



Work funded by the EU Horizon Europe grant 101060784 (EO4EU)

* <u>https://www.eo4eu.eu/platform</u>

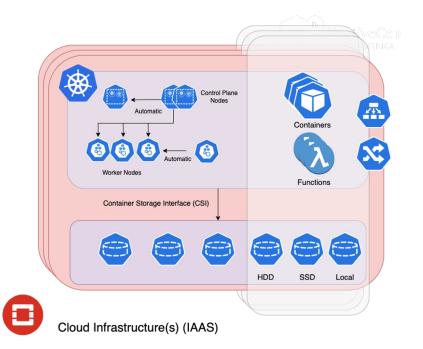
** Public user access soon-ish (Q1/2024)

Introduction: EO4EU Architecture

The EO4EU Platform uses a **multi-cloud** and **multicluster** architecture that leverages on 2 different OpenStack infrastructures:

- CINECA Supercomputing infrastructure provides High Performance Computing (HPC) and Cloud capabilities with Leonardo and ADA Cloud systems
- WEkEO*, a part of the Copernicus Data and Information Access Services (DIAS)

offering Infrastructure as a Service (IaaS) functionalities, and multiple Kubernetes clusters distributed across them.





Pillars of Observability



Observability pillars	Standards	
Metrics are numerical values that show how a system is performing over a period of time.	Prometheus with Alertmanager - de-facto standard for metrics in Kubernetes - often referred as kube-prometheus .	
Logs are immutable and timestamped records of events that happened over time.	Fluentd (with Fluent Bit) and document based database such as OpenSearch/ElasticSearch - often referred as EFK Stack .	
Traces represent a series of causally related distributed events that encode the end-to-end request flow through a distributed system.	OpenTelemetry gained a lot of popularity and adaptation, while still having options to integrate with other established solutions like Jaeger.	

Metrics: Single-cluster Case

Prometheus (kube-prometheus)

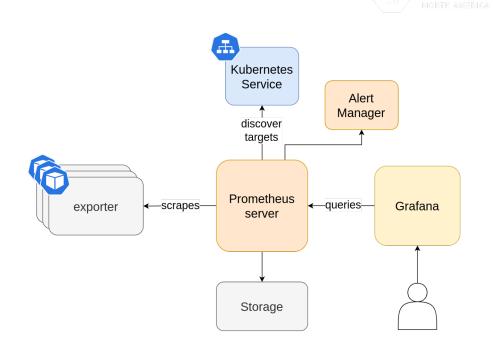
Monitoring and alerting system which collects metrics as time series data.

Grafana (grafana-operator)

Popular application to interactively visualize and analyze data.

Challenges in multi-cloud/cluster case:

- Scalability
- Availability
- Historical data
- Centralized alerts



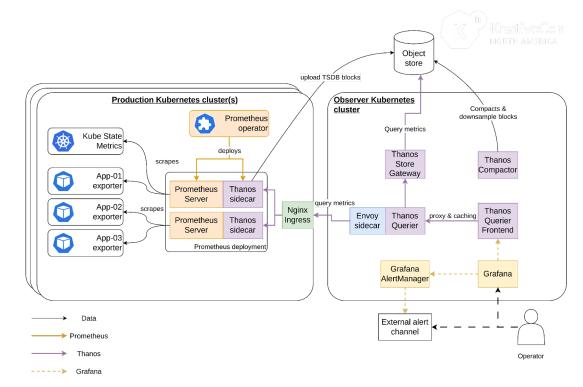
Metrics: Multi-cloud/cluster Case

Thanos (Bitnami Helm chart) provides:

- Prometheus compatibility (Grafana DataSource)
- Global Query View
- Unlimited retention
- Long-term storage compaction
 - Archival data
 - Faster metrics visualization

General issues:

- Sidecar approach management and resource consumption
- Direct query (data stored from 2h)
- Querier requires reverse proxy to reach clusters



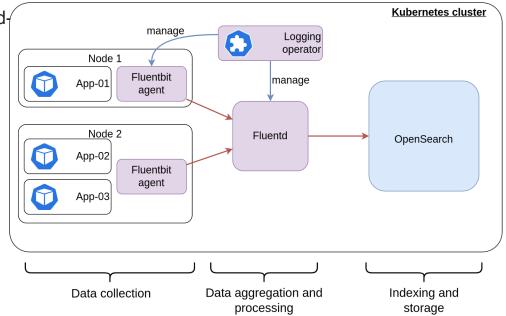
Logs: Single-cluster Case

Fluentd/Fluent Bit (kube-logging/loggingoperator)

- De-facto standard for logs (especially on cloudnative)
- Many source and output plugins
- Built-in filters and parsers available for common use-cases (e.g, nginx)
- TLS encryption support

Opensearch (opensearch-k8s-operator)

- API driven document-based database
- Highly Available and scalable
- Distributed architecture
- Advanced features like storage tiers, index templates and state management policies



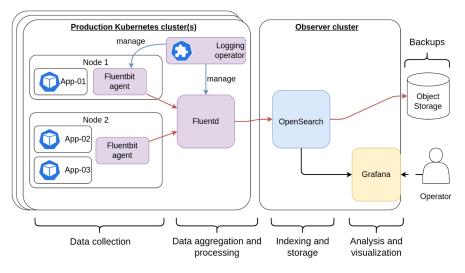




Logs: Multi-cloud/cluster Case

We implemented an approach for all our production Kubernetes clusters to send their logs to a centralized Opensearch cluster, and define standards to increase the efficiency and performance of Opensearch

- Index naming convention for each **cluster**
- Field mappers, State management and Index Templates based on application type
- **logging-operator** applies config changes and log flows to clusters on commit
- ClusterFlow and ClusterOutput manage log routing to generic or special (e.g., ingress) indexes and templates



Observability Day

Example indexes:

- logs-ingress-cineca-eo4eu-ope-2024-02-27
- logs-generic-cineca-eo4eu-ope-2024-02-27

Traces: Single-cluster Case

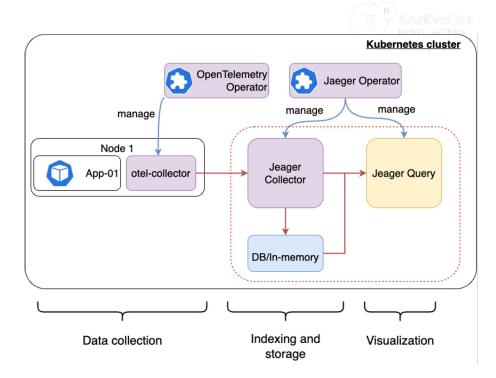


OpenTelemetry (opentelemetry-operator/otelcollector)

- Vendor-agnostic proxy that can collect observability data
- De-facto standard
- Processors to enrich data (e.g. k8sattributes)
- Supports broad variety of backends
 - Jaeger
 - Prometheus/Thanos (prometheusremotewrite)

Jaeger (jaeger-operator/jaeger-all-in-one)

- Focused on distributed tracing in microservices
- Service dependency graphs
- Multiple storage backends
 - Cassandra
 - Opensearch/Elasticsearch



Traces: Multi-cloud/cluster Case

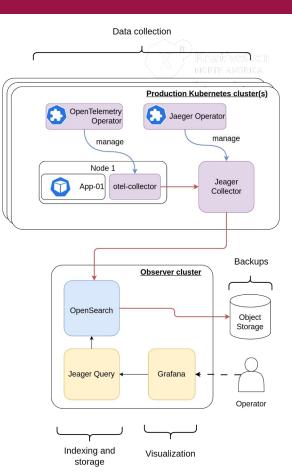
OpenTelemetry (opentelemetry-operator/otel-collector)

- Configured to send traces to local Jaeger collector via OTLP exporter
- Using processors to enrich data
- ingress-nginx in each cluster is configured sent traces to otelconnector
 - "otel-service-name" = unique \bigcirc
 - "use-forwarded-headers" Ο
 - Additional opentelemetry attribute(s) Ο

Jaeger (jaeger-operator/jaeger-production)

- Using Opensearch as a centralized storage backend for all collectors
- Traces are visualised by Grafana, in combination with the logs and metrics
- Each collector uses the same indices that Jaeger creates
 - Multi-tenancy is an ongoing discussion* Using es.index-prefix=traces* Ο
 - 0

* https://github.com/jaegertracing/jaeger/issues/3881



Observability Day EUROPE

Our Multi-cloud Journey

Challenges

- How to integrate several tools?
- Deploy and manage tools at scale (requires automation)
- Multi tenancy for heterogeneous teams access (e.g., RBAC)
- Missing useful documentation (we contributed to this*)
- Networking solution (e.g., ingress or cross-cluster connectivity and service discovery)

Design principles

- Open-source and open-license ecosystem
- Automation through GitOps and Infrastructure as Code (IaC)
- Kubernetes-first approach:
 - Official operators as first choice
 - Well-supported Helm charts
 - Custom solutions when needed (e.g., Grafana Dashboards and Alerts)

We started with GitLab CI and Terraform, then replaced with ad-hoc solutions and GitOps tools.

*Contributions:

Documentation: PVC and Volumes example · Issue #1391 · grafana/grafana-operator

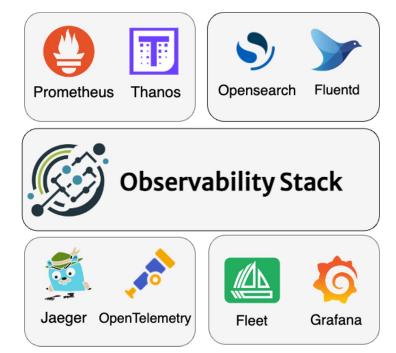


Update rbac-for-monitoring.md by armagankaratosun · Pull Request #1104 · rancher/rancher-docs

Introducing Observability Stack

By leveraging the available open-source ecosystem, we come up with the idea of Observability Stack

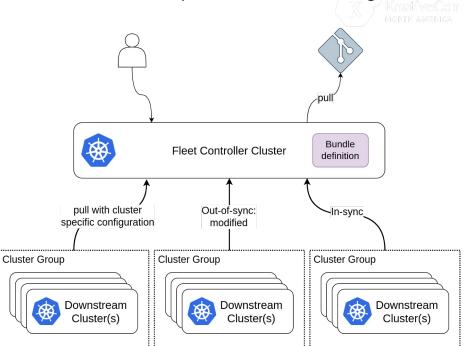
- Not a product an umbrella project, to provide a platform agnostic and flexible observability toolkit for administrators and developers to build their solutions tailored to their unique requirements.
- Automated Multi-cloud deployments by embracing GitOps as the 4th Pillar of Observability, we automatically configure clusters in our multi-cloud infrastructure to transmit metrics, logs, and tracing data to a central "observer" cluster, ensuring simplified deployment, day 2 operations and consistency.



GitOps as the 4th Pillar

Observability stack uses **Fleet (by Rancher)** for the distribution of its components and the management of their lifecycle.

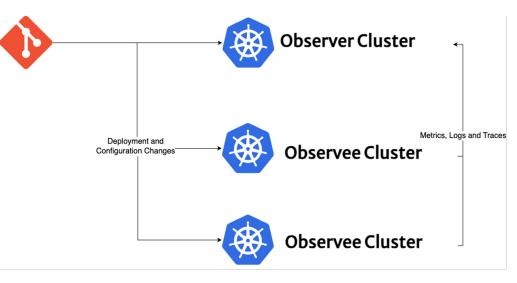
- Cluster Labels: Each cluster registered as a downstream cluster to Fleet Manager can be labelled to form groups
- Cluster Group: Clusters connected to the same Fleet Manager can be organized into groups with a *matchLabels* selector
- **Observability Stack** uses this mechanism to form Cluster Groups with labels:
 - o observability-role: observer
 - o observability-role: observee



GitOps as the 4th Pillar

By adopting a logical separation of "**observee**" and "**observer**" cluster labels and cluster groups, each cluster is configured to transmit its observability data, ensuring that the **observer** cluster has a centralized dataset to analyze.

- **Simplifies the deployment** of the Observability Stack to the downstream clusters
- Auto-enroll newly provisioned cluster as
 Observee
- Automates the configuration changes (e.g., adding new ClusterFlow) and version updates
- Ensures consistency within the multicloud/cluster setup
- **Plug-and-play architecture -** components of the Observability Stack can be changed easily



EO4EU Observability Platform

The EO4EU Observability Platform uses Observability Stack to observe the entire multi-cloud/cluster from infrastructure to application level and it leverages on a observee and observer cluster role is subdivision.

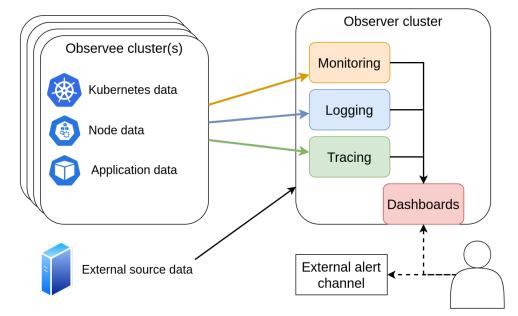
The **Observer** is the central cluster which allows to query, collect and visualize all data coming from **Observee** clusters which host all the EO4EU platform services.

Features:

- Scalable
- Highly available
- Full view on all sources
- Single access point for operators and developers

Tools:

- Monitoring (Prometheus, Thanos)
- Logging (Fluentbit, Fluentd)
- Tracing (OpenTelemetry, Jeager)
- Indexing (OpenSearch)
- Visualization (Grafana)









Let's hope it will work :)

If not we have backup pictures







Join us in overcoming the challenges on multi-cloud observability at





Please leave feedback

Thank you!



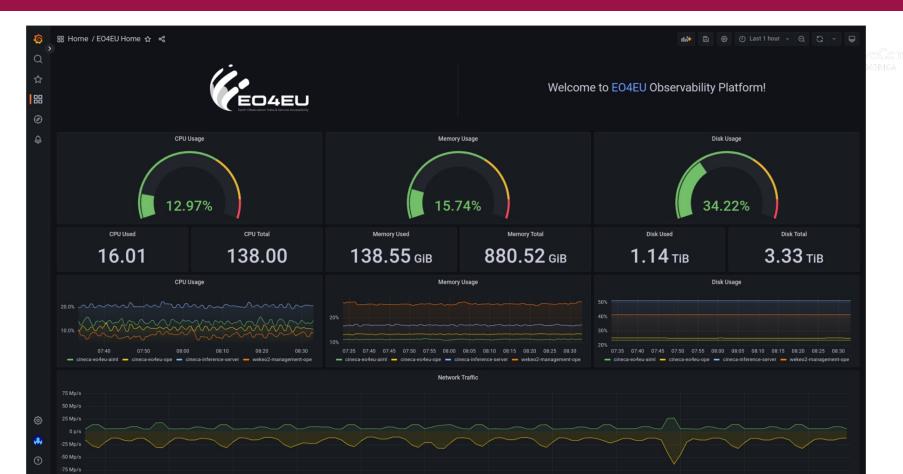


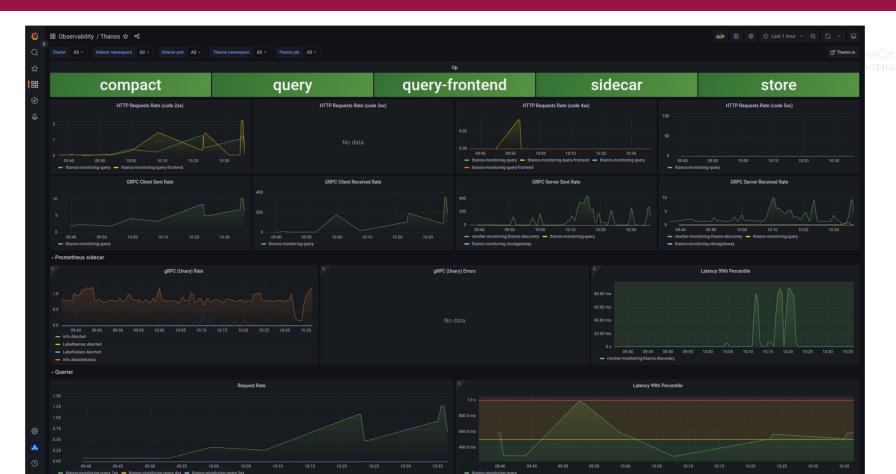


Work funded by the EU Horizon Europe grant 101060784 (EO4EU)

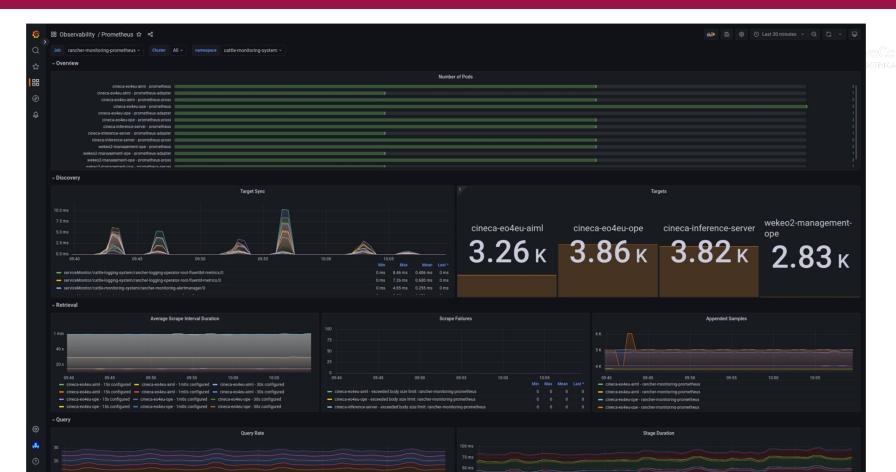
Backup Demo



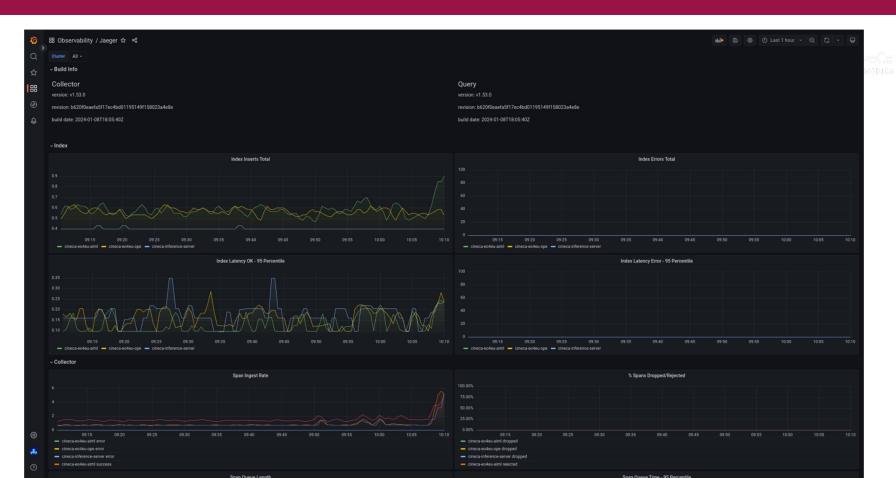






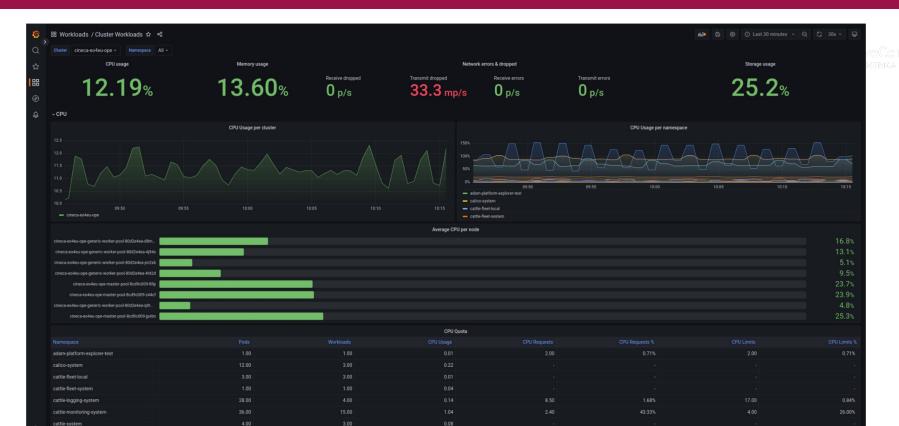










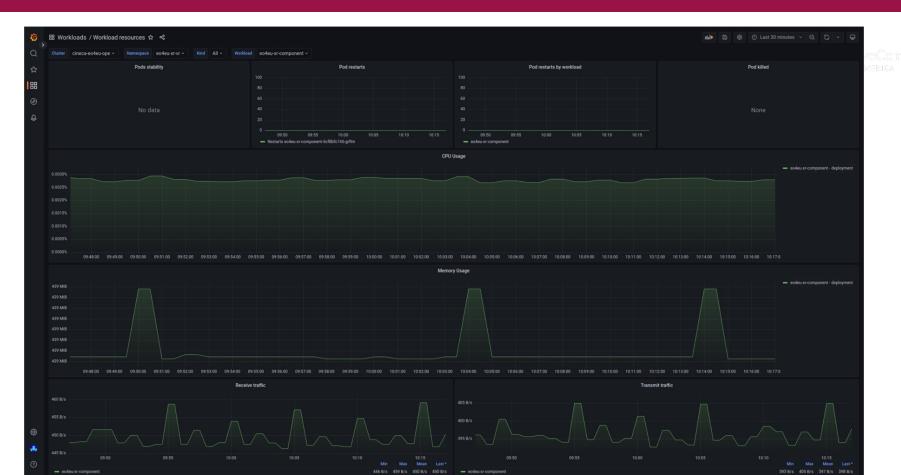


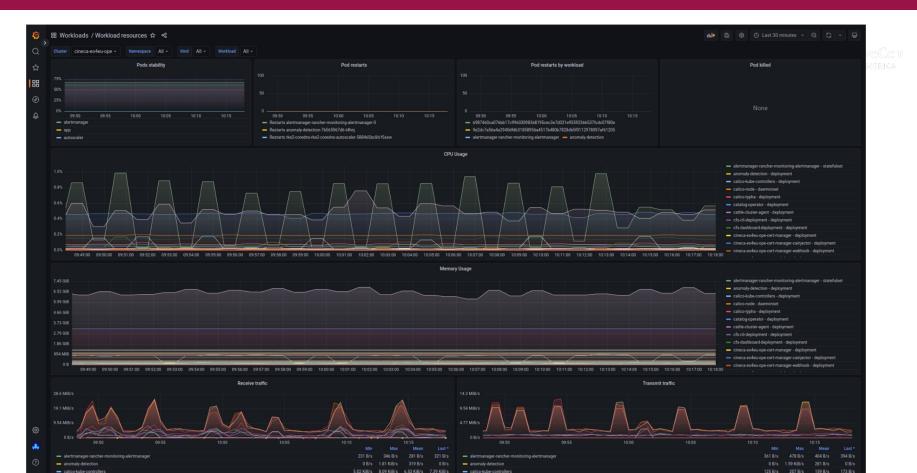
.

~ Memory

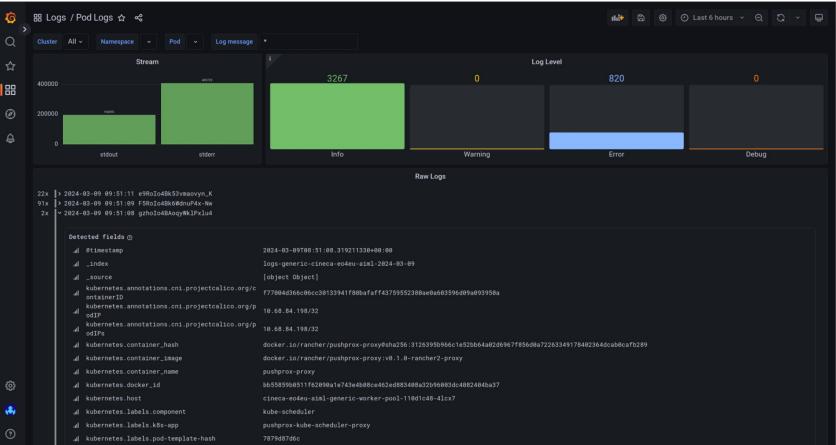
customer-facing-services

Memory Hears (w/e ear





Dashboards: Logs



EUROPE

Dashboards: Logs

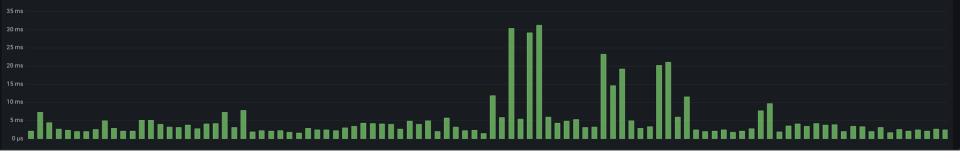
G → B Logs / Pod Logs ☆ ≪							
Q	Cluster cineca-eo4eu-ope ~ Namespace	tigera-operator ~ Pod All	 Log message 				
~	Stream			Log l	Level		
☆			6	0	0	0	
	6						
	4						
Ø	2						
¢							
~	0			Wenter		Debug	
	stdout	stderr	Info	Warning	Error	Debug	
				Raw Logs			
	6x 【✔ 2024-03-09 04:39:46 periodic-5m0s-re	econcile-event					
	Detected fields 💿						
	. _{II} @timestamp	2024-03-09T03:39:46.370551	065+00:00				
	" Request.Namespace						
	"II _id P1pMIY4BV4e_RqDWigYb						
	ull_index logs-generic-cineca-eo4eu-ope-2024-03-09 ull_source [object Object] ull kubernetes.container_hash docker.io/rancher/mirrored-calico-operator@sha256:1a68378fa3c2d9ac1c4db7c8c6cb5d72933a033b9a19a71062c2d67a79c13bc4						
	"I kubernetes.container_image	docker.io/rancher/mirrored	-calico-operator:v1.30.7				
	"" kubernetes.container_name tigera-operator						
	"II] kubernetes.docker_id 6decd76bb47191fe1a8053fc60131d619b6abe82b71802547e19ead472257873						
	الله kubernetes.host cineca-eo4eu-ope-generic-worker-pool-80d2e4ea-pn2xb						
	"" kubernetes.labels.k8s-app tigera-operator						
n	""į kubernetes.labels.name tigera-operator						
<u>6</u> 2	kubernetes.labels.pod- "W template-hash						
6 ,	"" kubernetes.namespace_name	tigera-operator					
_	Image: wider netes.pod_id 609cb0c7-b72a-49a7-9582-0c47ca681b34						
?	"" kubernetes.pod_name	tigera-operator-6f6d6ddf76	-9r98n				
	d loval	info					

Dashboards: Logs & Traces

Select log request ID to select only traces related to logs

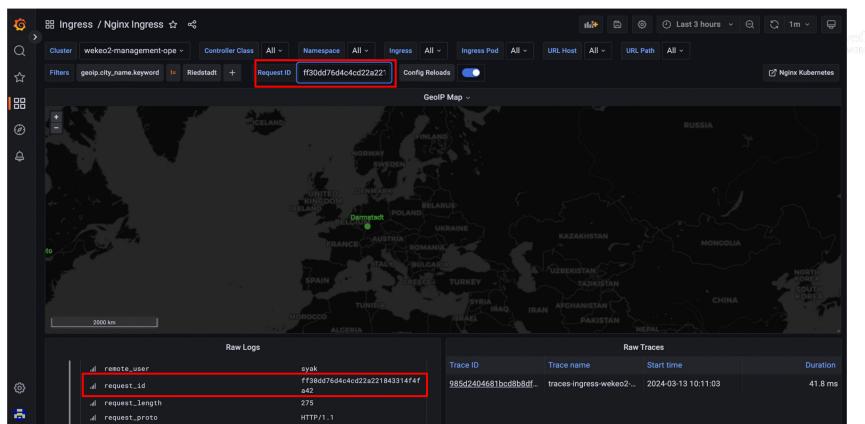
Raw Logs		Raw Traces				
				Start time		
kubernetes.labels.app.kubernetes.io/part-of kubernetes.labels.app.kubernetes.io/version	cineca-eo4eu-ope	50a69081f4c94ecc0110475ceaa	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:43	2.50 ms	
"" kubernetes.labels.controller-revision-hash	695b6c9ddd	1ec0c317a882d2f375dd586d9bf	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:42	2.64 ms	
"" kubernetes.labels.helm.sh/chart	rke2-ingress-nginx-4.8.200	d349d7750c3238cbfe8b0828edf	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:37	2.16 ms	
" kubernetes.labels.pod-template-generation		344f095935777d4aed67a1ba6c5	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:32	2.41 ms	
"" kubernetes.namespace_name	kube-system					
"I kubernetes.pod_id	2631096b-6c5a-451d-8cd5-6ae09d72375f	<u>f71727b4061e1c1cdc353b17bb8</u>	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:27	2.14 ms	
"I kubernetes.pod_name	rke2-ingress-nginx-cineca-eo4eu-ope-controller-hgkpv	4506539cc2607eb4291eff4113f1	traces-ingress-cineca-eo4eu-ope:	2024-03-09 08:34:24	2.57 ms	
"" logtag		abbe0e1c1de868df56487cc7f7a4	traces-indress-cineca-eo4eu-one:	2024-03-09 08:34:22	1 75 ms	





Dashboards: logs & traces

Observability Day



,|| request_query service=git-upload-pack

 \bigcirc

|| request_time 0.041

Dashboards: Traces

ල Explore ද 🗯 Traces ∨		Split 🔡 Add to dashboard	 Last 1 hour ~ G 	ि 🕄 Run query 🗸
				மு⊚ ⊞ ∷
Query type Search TraceID JSON file				
Trace ID e16e6abaf2a160da0564e751eadd43c5				
+ Add query 🕤 Query history 💮 Inspector				
Trace View				
traces-ingress-cineca-eo4eu-ope: HTTP POST prometheus-operated	//thanos.info.lnfo/lnfo e16e6abaf2a160da0564e751eadd43c5			
Trace Start: 2024-03-09 10:00:12.537 Duration: 2.49ms Services: 1 Depth: 1 Tota				
Орь 621.5µз	1.24ms	1.86ms		2.49ms
Service & Operation	621.5µs	1.24ms	1.86ms	2.49ms

Dashboards: Tracing view

Observability Day EUROPE

ம ⊚ ம் ∷

7.48ms

\$	@ Explore 😪 🗯	Traces ~ CI	ose 🔡 Add to dashboard	୦ ୪ ୧ ୧ ୦ ୪	🗯 Traces 🗸	X Close ⊞ Add to da	shboard < 🕘 ~	<i>∂</i> > Q
Q				© ⊚ ∰ ‼				¢ ©
☆	Query type Se	arch TraceID JSON file			Query type Search Tracell	D JSON file		
88	Service	traces-ingress-cineca-eo4eu-ope			Trace ID 53687f4500d	17a91242dd37ac52774bd1		
	Operation	All			+ Add query 🗿 Query histo	ry 🛈 Inspector		
Ø	Tags							
¢	Advanced options				Trace View			
	+ Add query	Ouery history	ator					
	T Add query				traces-ingress-cineca-eo4eu-o	ope: HTTP POST prometh	eus-operated /than	os.Store/Series
	Table					7 Duration 7 49ma Comission	Denthi 1 Tatal Coons	1
					Trace Start: 2024-03-09 09:57:04.02	Duration: 7.40ms Services: 1 3.74ms	Depth: Total Spans: 5.61r	
	Trace ID	Trace name	Start time	Duration				
	24b5e1845513fee5d		4 2024-03-09 09:57:08	3.81 ms				
	8366bf35f2626f5ea3		4 2024-03-09 09:57:08	3.45 ms	Servi 🗸 🔪 💥 🖓 🖓 Servi	1.87ms	3.74ms	5.61ms
	e68ba83cd80e4dabe		4 2024-03-09 09:57:07	2.43 ms 3.46 ms	traces-ingress-cineca-			
	1aff0fb49fb16a7f3c		4 2024-03-09 09:57:07	3.46 ms				
	880646289136d0401		4 2024-03-09 09:57:07 4 2024-03-09 09:57:07	3.48 ms				
	<u>66cf93fdd2465f487f</u>		4 2024-03-09 09:57:07	5.11 ms				
	53687f4500d7a9124		4 2024-03-09 09:57:04	7.48 ms				
	86cab38f896ace163	-		2.03 ms				
Ô	4a7b24a0fdec862d0		4 2024-03-09 09:57:02	8.00 ms				
	b1867fe2f2524998a		4 2024-03-09 09:57:00	3.74 ms				
	2e9ab618697fd9ad4		4 2024-03-09 09:57:00	3.87 ms				
?		h7f2 traces-ingress-cineca-eo		2 91 ms				