



**KubeCon**



**CloudNativeCon**

**Europe 2023**





KubeCon



CloudNativeCon

Europe 2023

# Sustainability Through Accountability in a CNCF Ecosystem

*Yuval Lifshitz, @lifshitz\_yuval, IBM*  
*Huamin Chen, @root\_fs, Red Hat*





Please scan the QR Code above  
to leave feedback on this session

# Agenda

- Motivation
- Accountability
- Problems and Solutions
- Architecture
- Rook
- Kepler
- Analytics
- Demo
- Future work

# Motivation

## Energy Efficiency

### ENERGY SCALE

Global electricity demand

**20,000** TWh

Electricity use by ICT

**2,000** TWh

Data-centre electricity demand

**200** TWh

Bitcoin use by mid-2018

**20** TWh

©nature

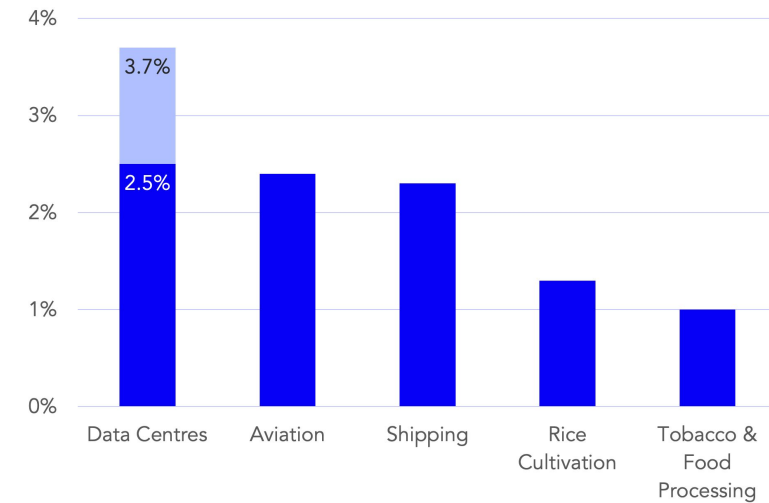
Figures are approximate.

## Carbon Footprint



### Global cloud computing emissions exceed those from commercial aviation

Share of global CO<sub>2</sub> emission generated by sector/category



Source: Climatiq Analysis, The Shift Project, OurWorldinData



# Accountability - Who is responsible?

Accounting



## We Need Better Carbon Accounting. Here's How to Get There.

by Robert S. Kaplan and Karthik Ramanna

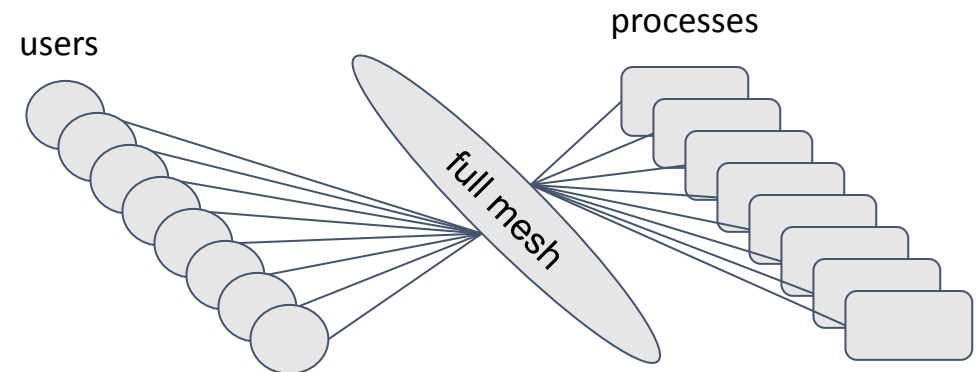
April 12, 2022



*California*  
LEGISLATIVE INFORMATION

**SB-260 Climate Corporate Accountability Act.** (2021-2022)

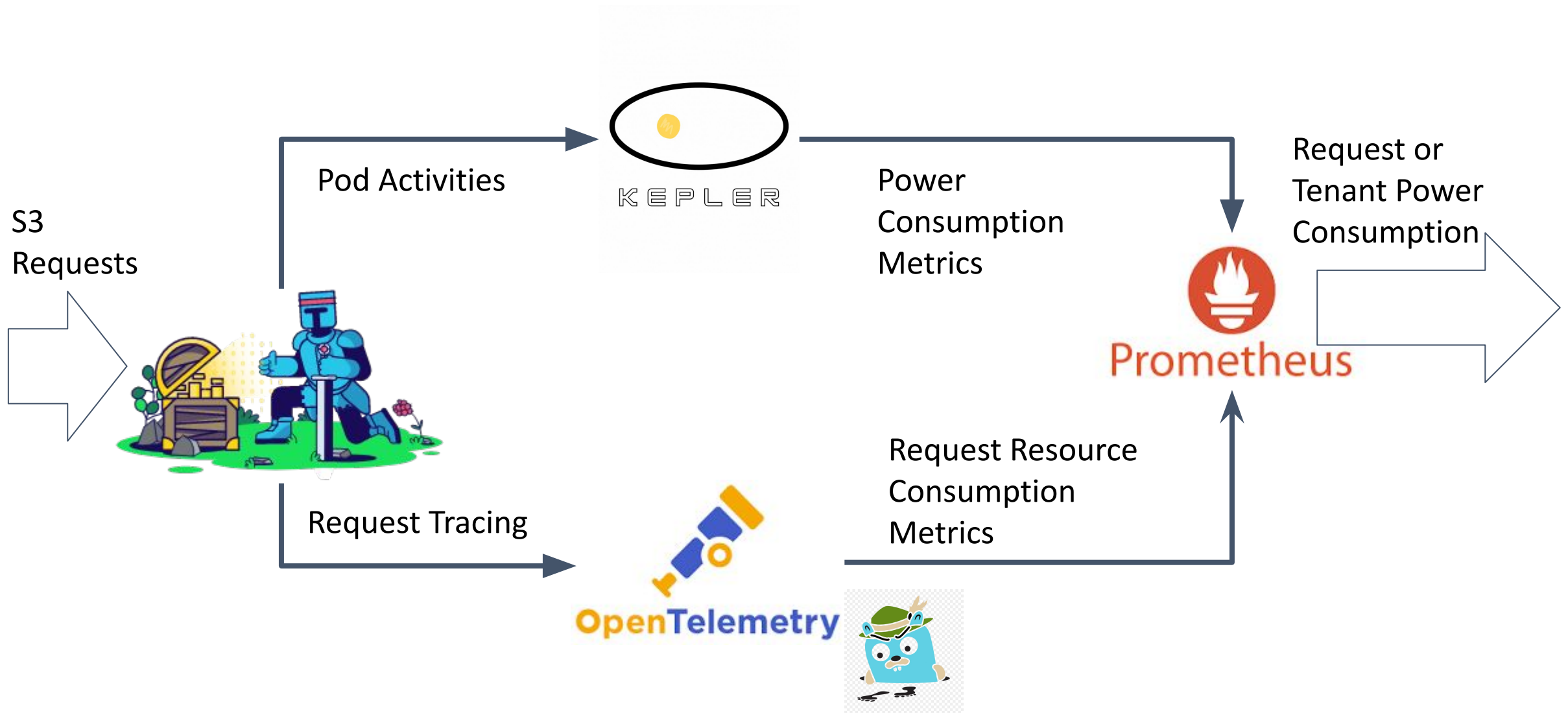
**Untangle** the ties  
between **users** and the  
**energy** consuming parts



# Problems and Solutions

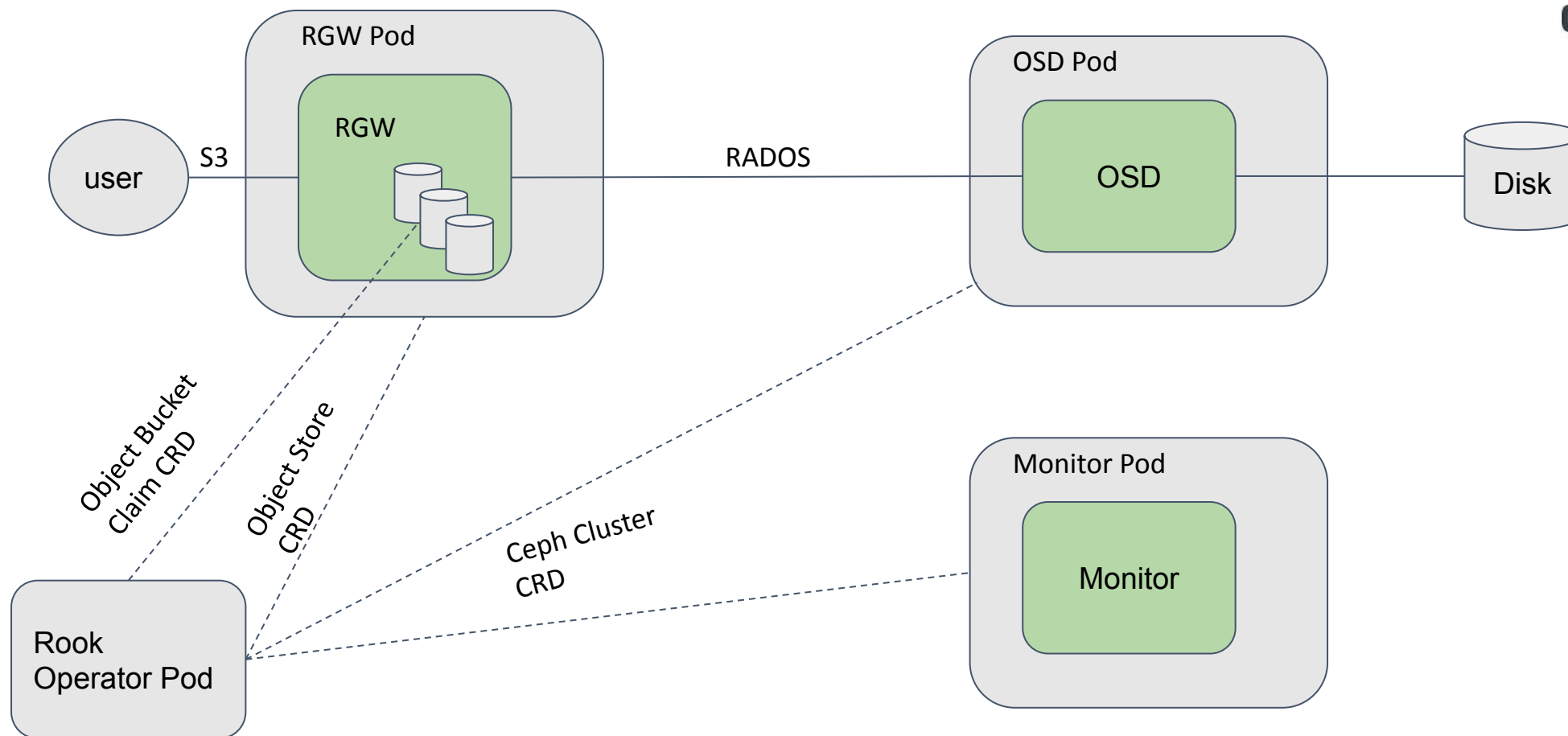
Problem	Solution
Workload Power Consumption Measurement	Use Kepler to measure Pod level power consumption
S3 Request Processing Tracing	Use OpenTelemetry+Jaeger to trace S3 request processing inside Ceph cluster containers
S3 Request Power Consumption Attribution	Attribute power consumption to S3 requests based on the resource usage and Pod power consumption

# Architecture

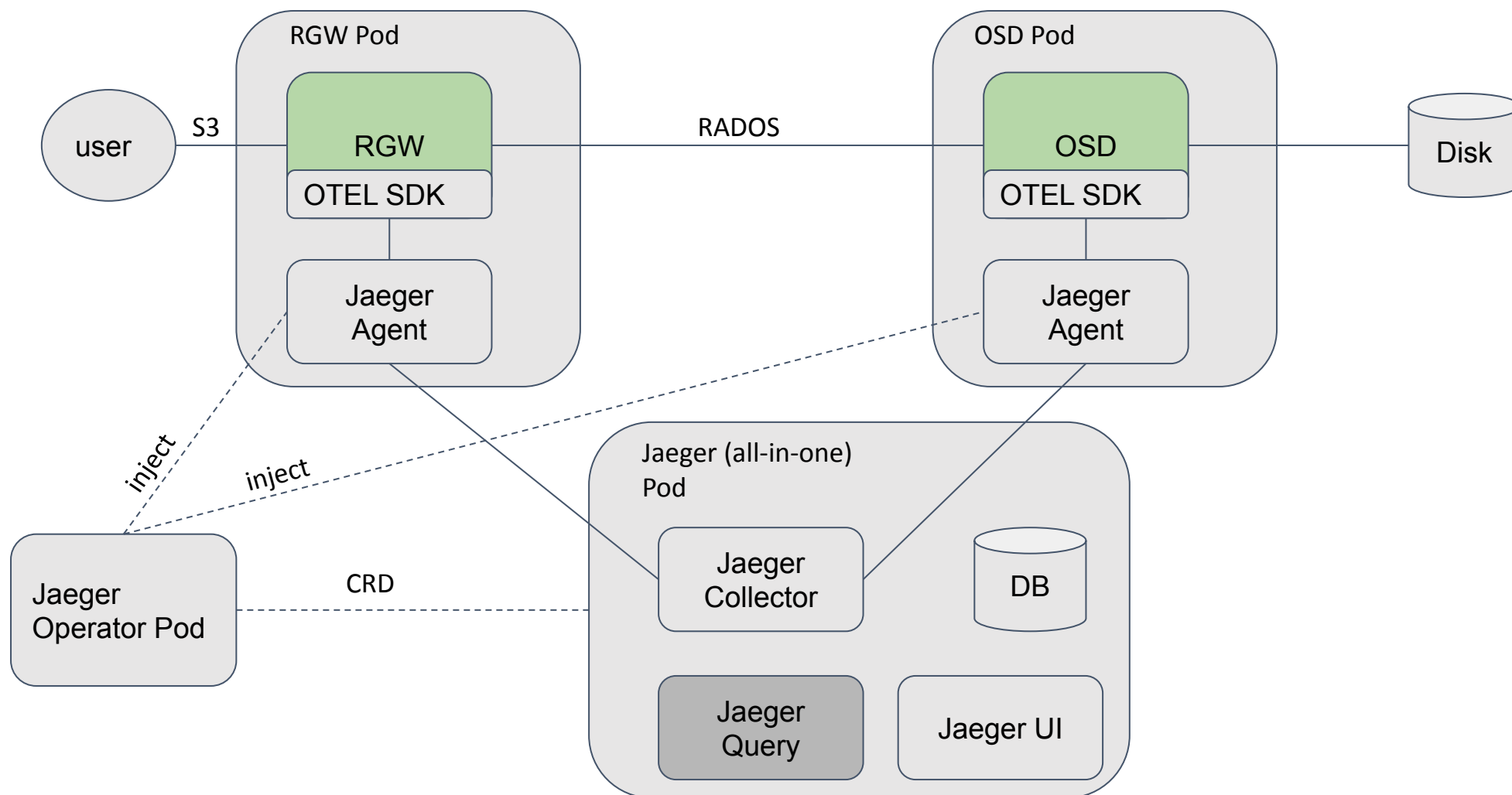




# Rook + Ceph



# Open Telemetry + Jaeger



# Kepler

Data Presentation

Energy stats as Metrics Counters



Data Modeling

$$\text{Energy}_{\text{Pod}} = \sum (\text{CPU stats, Memory stats, GPU stats, cgroups stats, hwmon stats})$$

Model Server

Data Aggregation

RAPL,  
hwmon,  
etc

cgroups  
stats

GPU  
stats(nvml)

hwmon stats

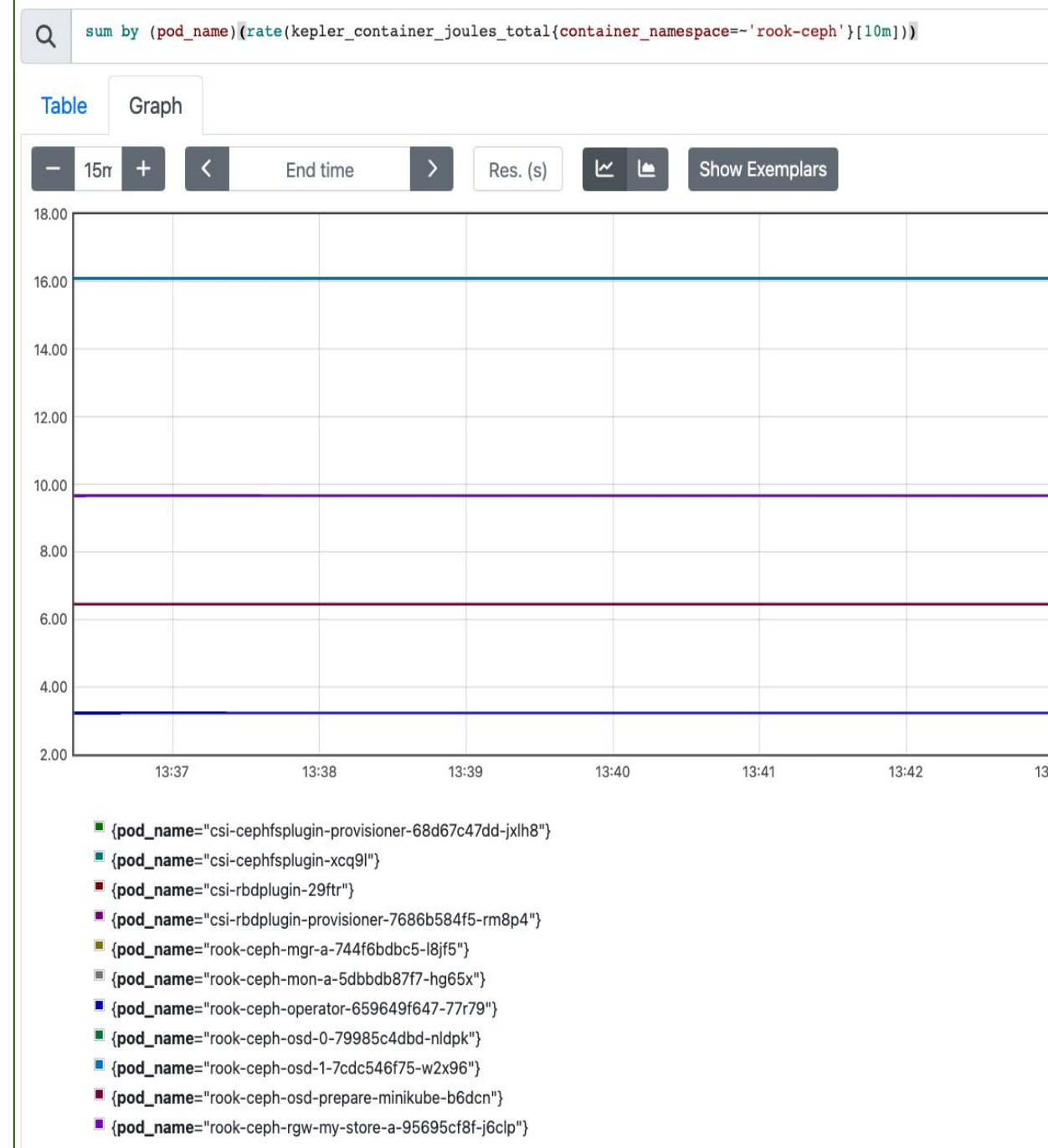
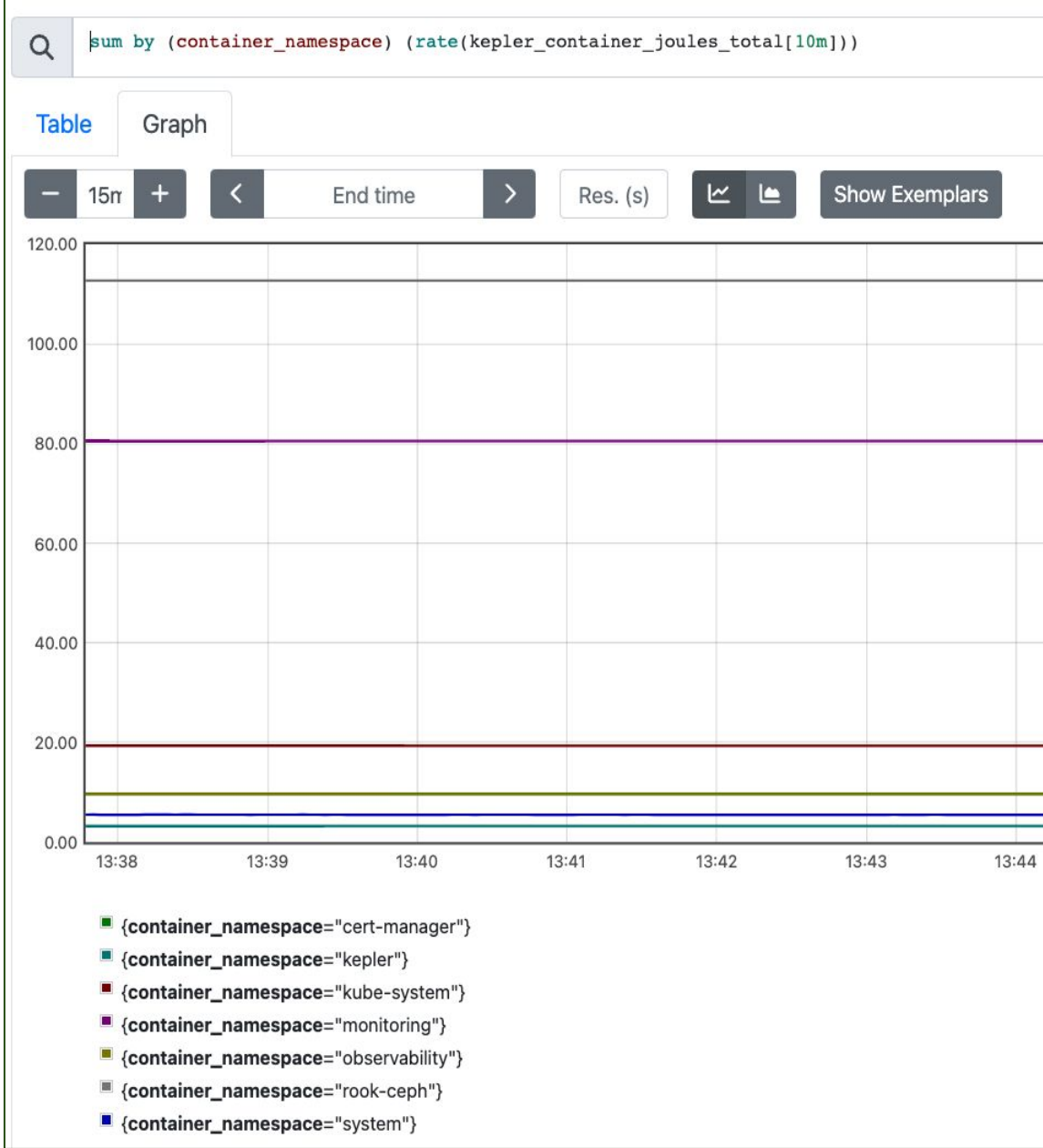
PID, Cgroup ID  
CPU cycles  
CPU time  
CPU instructions  
Cache misses  
....

Data Collection



Linux Kernel  
Tracepoint

Performance Counter



# Analytics

Trace:

Span:  
Pod: rgw1  
User: userX  
Duration: 200

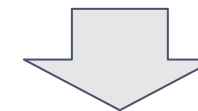
Span:  
Pod: rgw1  
Duration: 100

Span:  
Pod: osd2  
Duration: 50

Span:  
Pod: osd2  
Duration: 150



pods	users	duration	%
rgw1	userX	12000	30
rgw1	userY	5000	12.5
rgw1	userZ	23000	57.5
osd1	userX	1200	35.2
osd1	userY	1500	44.2
osd1	userZ	700	20.6
osd2	userX	800	17.4
osd2	userY	2100	45.6
osd2	userZ	1700	27



pods	energy
rgw1	56000
osd1	35000
osd2	30000

$$E(user) = \sum_{pods} E(pod) \cdot Duration(user, pod) / Duration(pod)$$



user	energy
userX	16800+12320+5220 = 34340
userY	7000+15470+13680 = 36150
userZ	32200+7210+8100 = 47510

# Demo

**demo cast**



<https://ascinema.org/a/5lIQQxcBLkZ0DPRReLHMUaPTE>

**demo repo**



<https://github.com/yuvalif/sustainability-via-accountability>

# Future Work

- Overhead. How much overhead does tracing incur and how to mitigate the overhead.
- Automation. Can energy attribution be automated, reported, and visualized?
- Optimization. Can resources be reconfigured to reduce energy consumption?
- Policy. Can energy consumption limiting be used as a bucket policy?



**KubeCon**



**CloudNativeCon**

Europe 2023

# Thank You

