E4 COMPUTER ENGINEERING

ENABLING AI DRIVEN HPC INFRASTRUCTURES WITH THE EXASCALE MONITORING PLATFORM

WHO WE ARE

Timeline:

- 2002 — E4 COMPUTER ENGINEERING
- 2005 — FIRST INFINIBAND CLUSTER
- 2009 — FIRST HYBRID CPU-GPU NVIDIA TESLA CLUSTER
- 2012 — SGI ITALIA ACQUISITION (sgi); PEDRAFORCA CLUSTER — Barcelona Supercomputing Center Centro Nacional de Supercomputación
- 2013 — FIRST ARM+GPU SERVER, FIRST ARM CLUSTER
- 2016-2017 — D.A.V.I.D.E. FIRST OPENPOWER SERVER & PETAFLOPS CLUSTER; TOP 500 CINECA, The GREEN 500, PRACE; 299° IN TOP500 (JUNE 2017), 14° IN GREEN500 (JUNE 2017)
- 2018 — E4 ANALYTICS, E4 ANALYTICS FOUNDATION
- 2018 — EPI SGA1 PARTNER OF (European Processor Initiative, epi)
- 2019 — EUROHPC JU CALLS; SW INTEGRATION & CI/CD FACTORY
- 2021 — E4 COMPUTER ENGINEERING DEUTSCHLAND, E4 DEUTSCHLAND FOUNDATION
- 2022 — EPI SGA2 (LIGATE, exaFOAM, ADMIRE, textarossa, ATTRACT, MAELSTROM); BOLOGNA TECNOPOLO (ECMWF | CINECA) + FIRST RISC-V BASED CLUSTER (Monte Cimone); (EUPEX, RISC-V)
- 2023 — DYNAMOS, epi, EUPEX; DATACENTER CHIRON LAB NVIDIA GRACE & NVIDIA GRACE HOPPER (CHIRON LAB); EUMaster4HPC, iSOLDE, TRISTAN, MAX DRIVING THE EXASCALE TRANSITION, DECICE, SPACE EDOCC
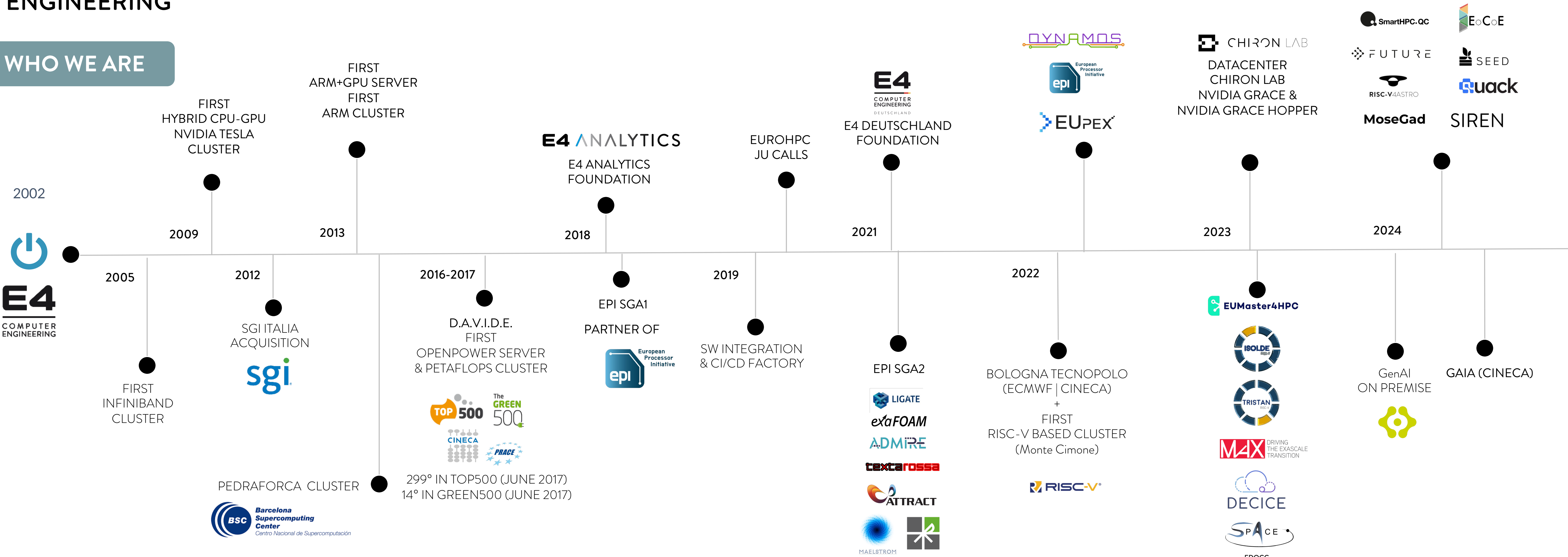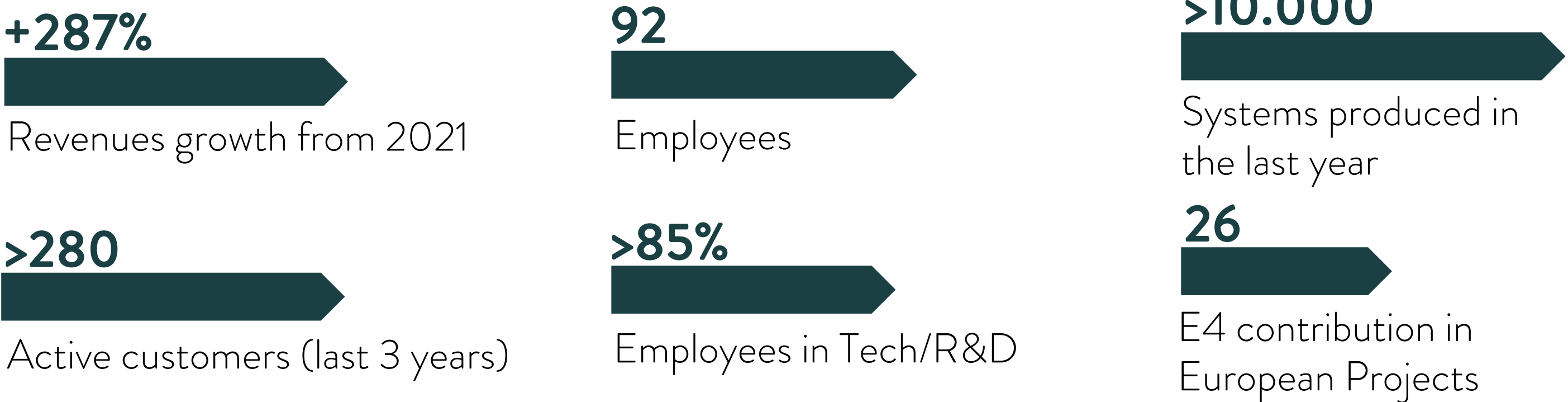- 2024 — SmartHPC-QC, EoCoE, FUTURE, SEED, RISC-V ASTRO, Quack, MoseGad, SIREN; GenAI ON PREMISE; GAIA (CINECA)

E4 is one of the main HPC service integrator in Europe; we work with all the main hardware vendors for all the components category and our job is codesign with our customers and deploy the infrastructures that best serve their needs. Most of the clusters are used in hard sciences or engineering simulations, in academia or enterprises from pharma to aero space.

E4 IN NUMBERS

- **+287%** Revenues growth from 2021
- **92** Employees
- **>10.000** Systems produced in the last year
- **>280** Active customers (last 3 years)
- **>85%** Employees in Tech/R&D
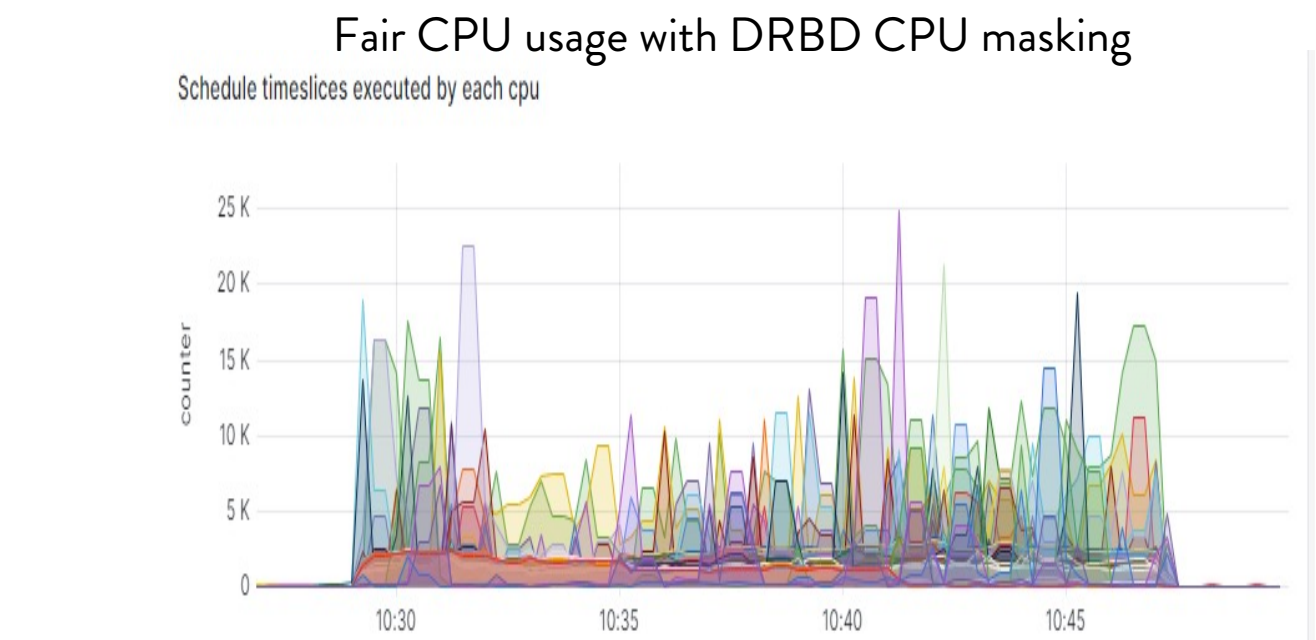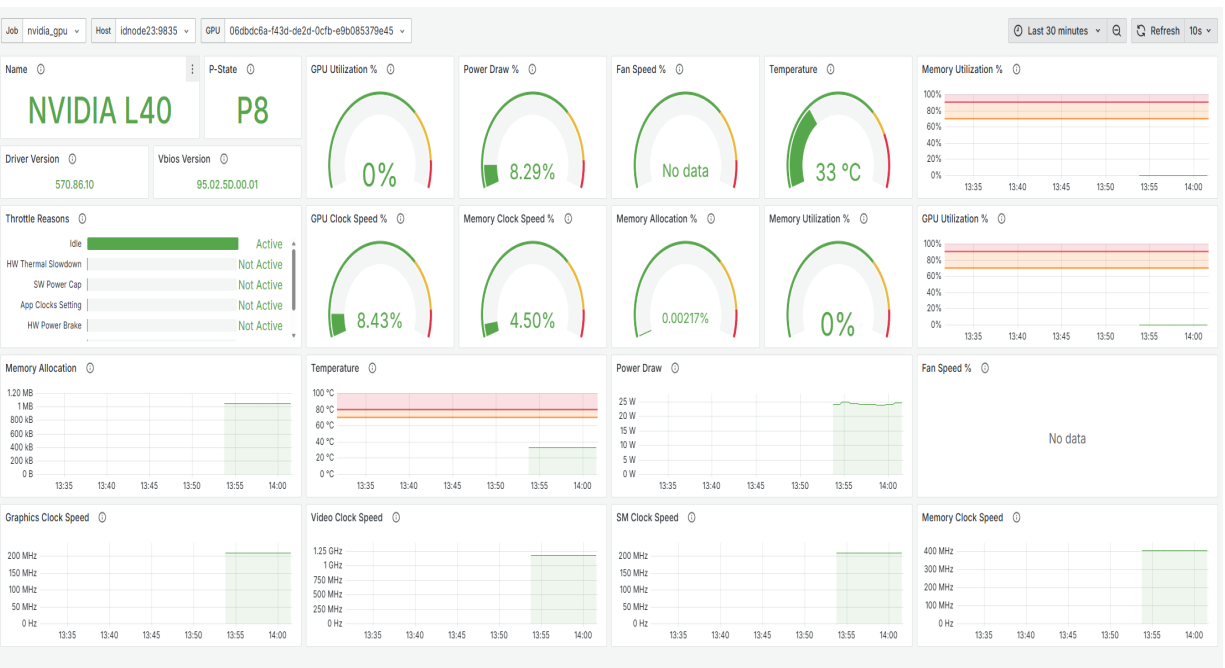- **26** E4 contribution in European Projects

TECH FACTORY

Integration Facility where our technicians build servers or storage systems.
Burn In Room to improve E4 systems reliability with at least 72 hours of test that involves all components. **R&D Lab** with 6 standard racks with heterogeneous systems, 100kW, remote access available on demand to perform benchmarking, co-design, prototyping.

CLUSTER MANAGEMENT
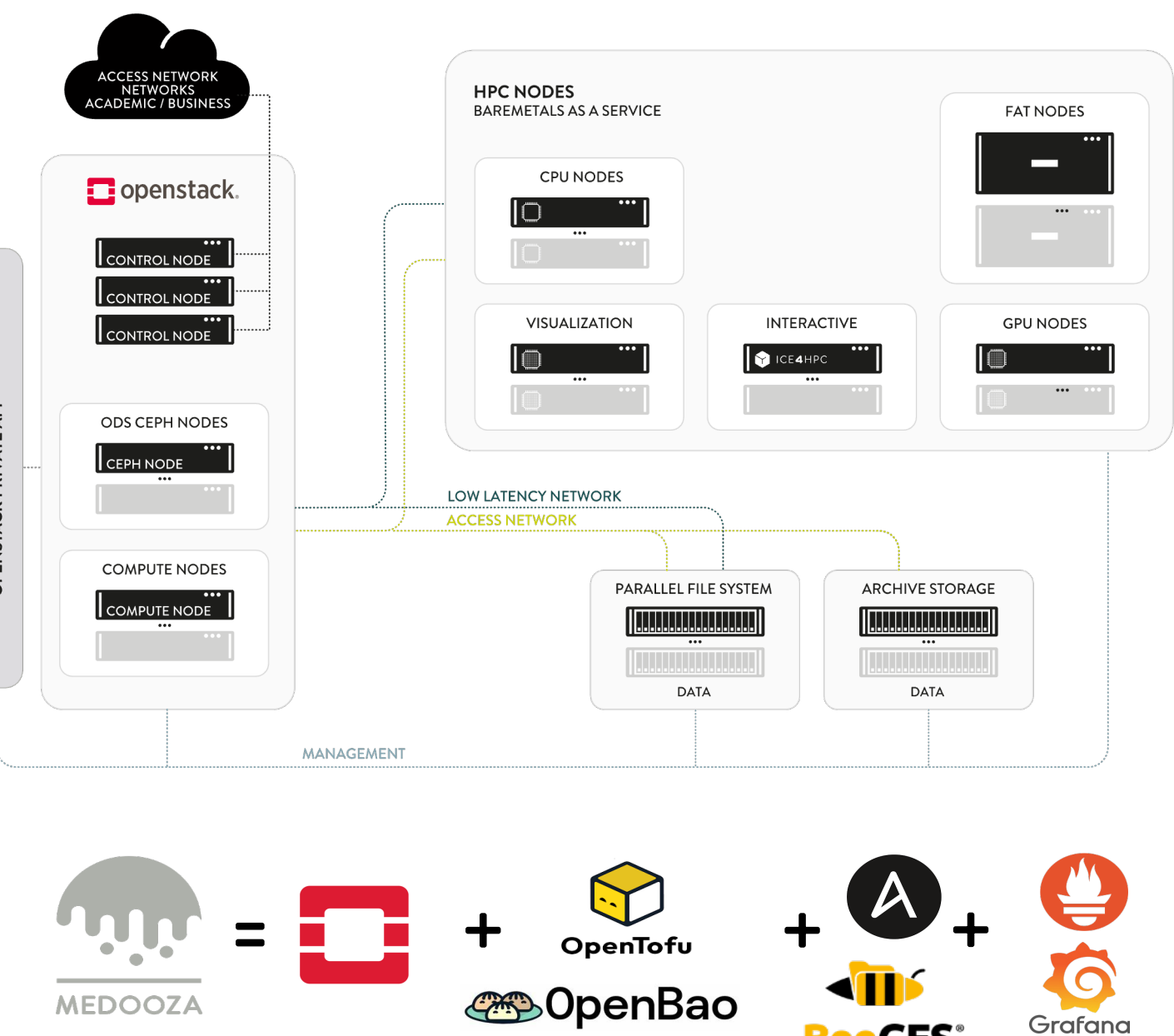
Medooza: General architecture and set of tools to manage and deploy HPC and private cloud. Use IaC tools to create resources and manage secrets, etc.

MEDOOZA = (OpenTofu, OpenBao) + (BeeGFS, Grafana) + EXAMON + ICE4HPC + EESSI

Storage: first to use NVME RAID solutions for BeeGFS. Tested the DRBD and LINSTOR solutions from LINBIT. **Argus:** monitoring and alerting tool for E4 storage solutions; uses OCP metrics and provides data about usage and carbon footprint.

Fair CPU usage with DRBD CPU masking
Schedule timeslices executed by each cpu
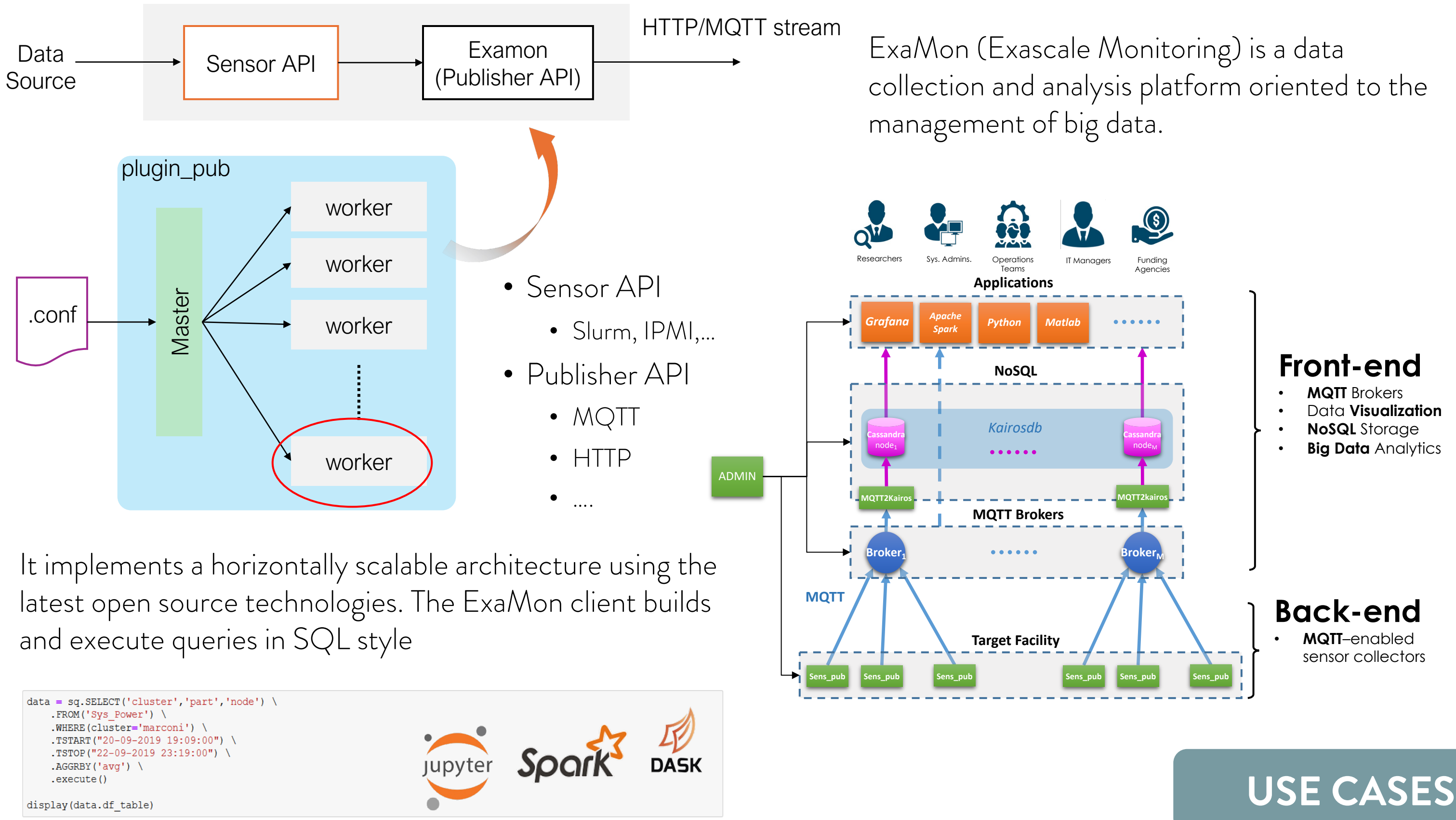
FLAGSHIP PROJECTS

ASTRI@MountTeide, Tenerife: ICT infrastructure for the pilot project of the Cherenkov telescope array.

Franklin@IIT, Raise@IIT, Genoa: HPC and Private cloud clusters based on Medooza with complete SW stack.

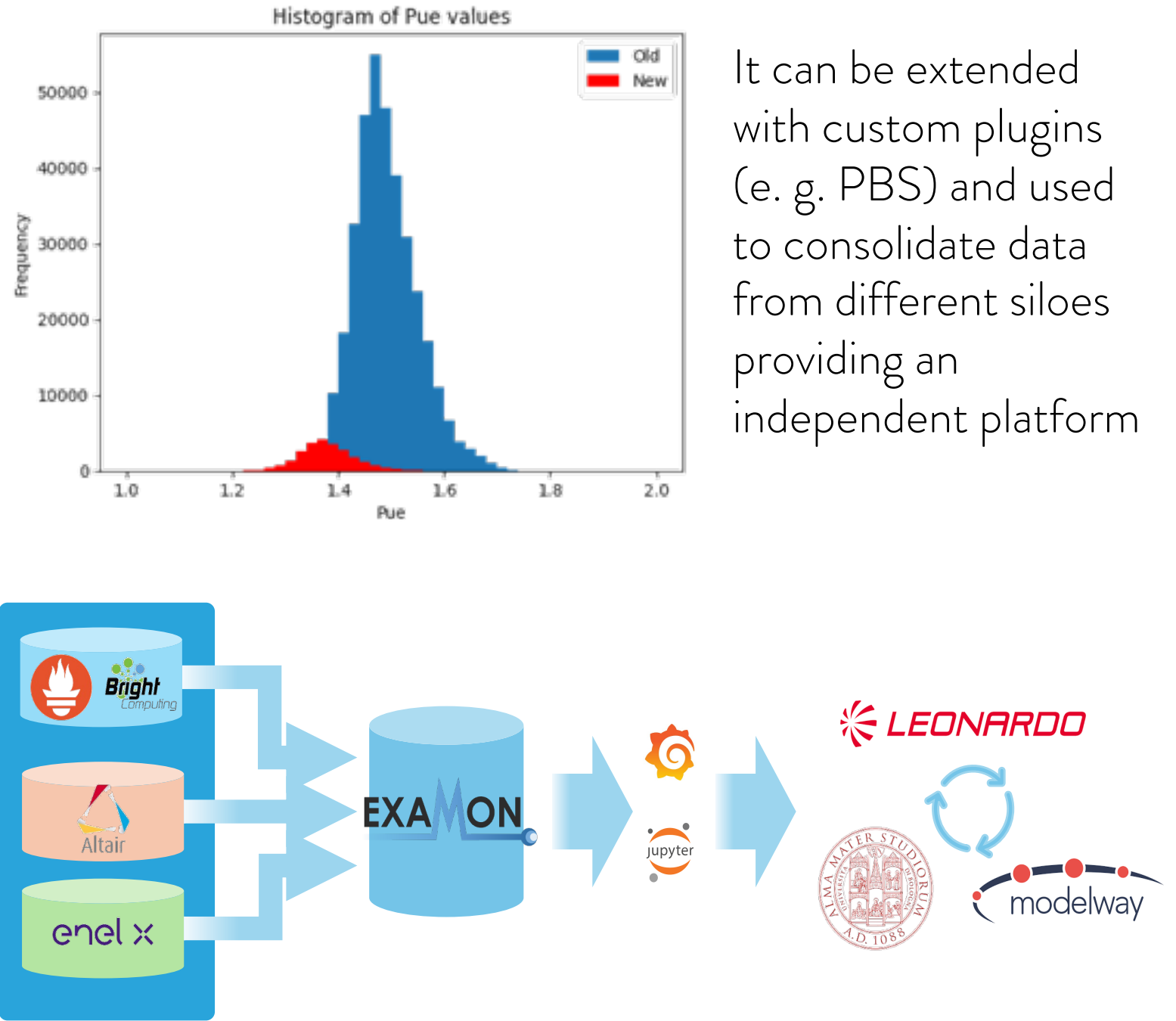**Gaia@CINECA, Bologna:** Expansion of Galileo100 called Gaia: cloud for public research in CINECA.

EXAMON PROJECT

ExaMon (Exascale Monitoring) is a data collection and analysis platform oriented to the management of big data.

- Sensor API
  - Slurm, IPMI,...
- Publisher API
  - MQTT
  - HTTP
  - ....

It implements a horizontally scalable architecture using the latest open source technologies. The ExaMon client builds and execute queries in SQL style

```
data = sq.SELECT('cluster','part','node') \
.FROM('hpc_Power') \
.WHERE(cluster='marconi') \
.TSTART("22-08-2019 19:09:00") \
.TSTOP("22-08-2019 23:19:00") \
.AGGRBY("avg") \
.execute()
display(data.df_table)
```

(jupyter, Spark, DASK)

Front-end
- MQTT Brokers
- Data Visualization
- NoSQL Storage
- Big Data Analytics
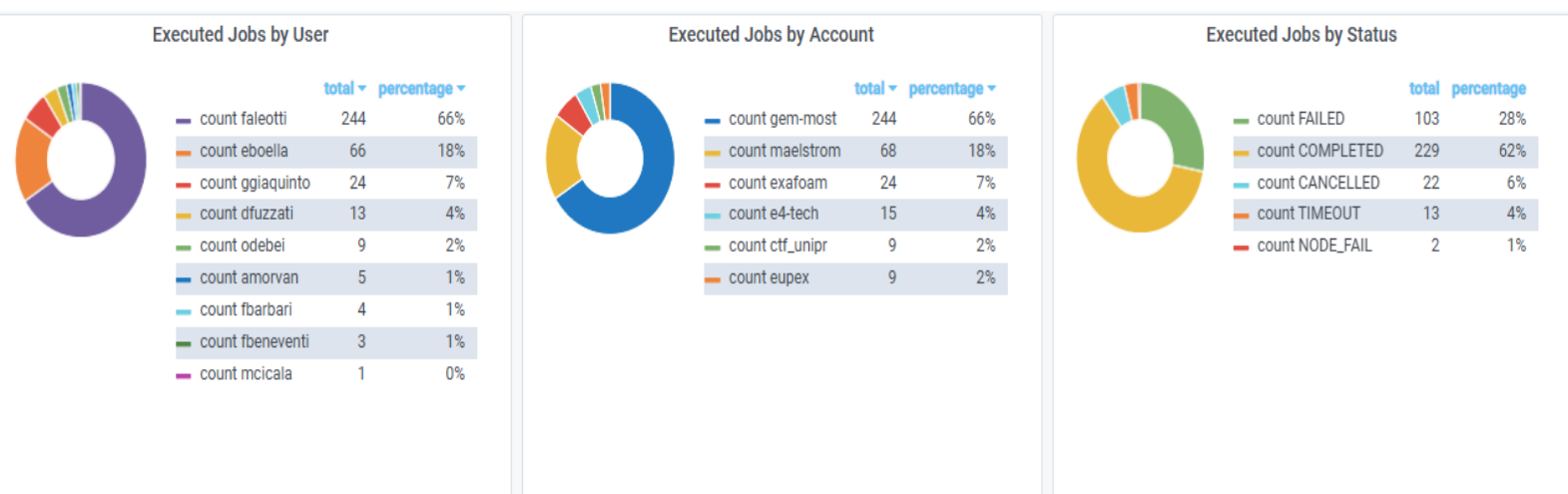
Back-end
- MQTT-enabled sensor collectors

Cluster digital twin: the efficiency curves obtained from historical data can be analyzed to define the optimal operating point of the devices as a function of load, temperature, and humidity.

USE CASES
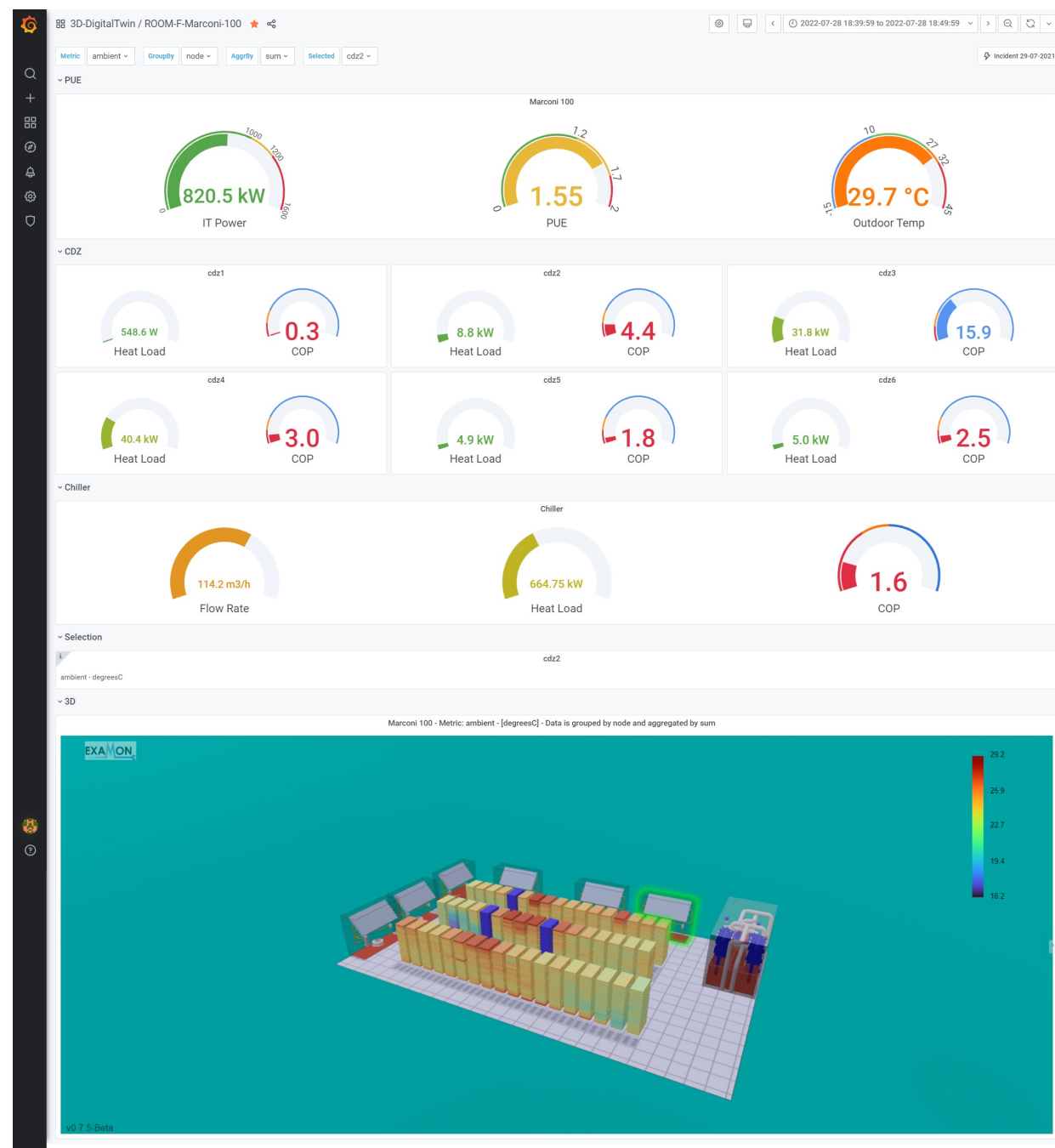
Thanks to the immediate feedback provided by the dashboards, the operators were able to set the individual set points of the devices optimally. PUE reduction of 8%

Histogram of Pue values

It can be extended with custom plugins (e. g. PBS) and used to consolidate data from different siloes providing an independent platform

(Bright, Altair, enel x) → EXAMON → jupyter → LEONARDO, modelway

SUMMARY AND PERSPECTIVES

ExaMon can be used to collect data per job (type) and per user and the data can be modeled with AI methods to provide predictive energy usage (scheduler hooks and policies can be optimized) and predictive anomaly detection (preventing costly restarts and allowing less checkpointing e. g. for training jobs).

Normal LLM models cannot be used for advanced AD but alternatives exists, see Molan et al. (https://doi.org/10.1016/j.future.2024.06.032). This would allow to create operational chatbots for HPC operators see J. A. Khan, M. Molan, e A. Bartolini, «EXASAGE: The first Data Center Operational Data Co-pilot»

Job Energy By User

user_name	data_quality	total_energy_Wh (sum)
eboella	100	25.10 kWh
gglaquinto	100	5.52 kWh
faleotti	100	3.95 kWh
odebei	100	1.63 kWh
dfuzzati	100	702.58 Wh

E4 EUROPEAN PROJECTS · EXAMON website · EXAMON Github archive · BEEGFS-ANSIBLE COLLECTION · POSTER DOWNLOAD

www.e4company.com