DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Unlocking Sustainable Exascale Computing with Data Centre Digital Twin

Are current infrastructures suitable for extreme data processing? Technologies for data management

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION

The Data Centre Digital Twin

Digital twins are virtual replicas of physical systems that can be used for simulation, monitoring, and optimization. In the case of data centers, a digital twin can help us better understand and manage the complex interplay between hardware, software, and environmental factors.

By creating a digital twin of a data center, we can test different scenarios, identify potential issues before they occur, and optimize resource utilization. This is especially important as we move towards exascale computing, where the amount of data being processed and stored is increasing exponentially. With a digital twin, we can ensure that our data centers are running efficiently and sustainably.

Holistic Monitoring Frameworks in Datacenters

Holistic monitoring frameworks are becoming increasingly popular as data centers grow in size and complexity. One example is **Examon**, a tool designed and maintained by the University of Bologna. Examon provides three types of data sources: real-time system-level metrics, application-level metrics, and energy consumption data.

Collected data

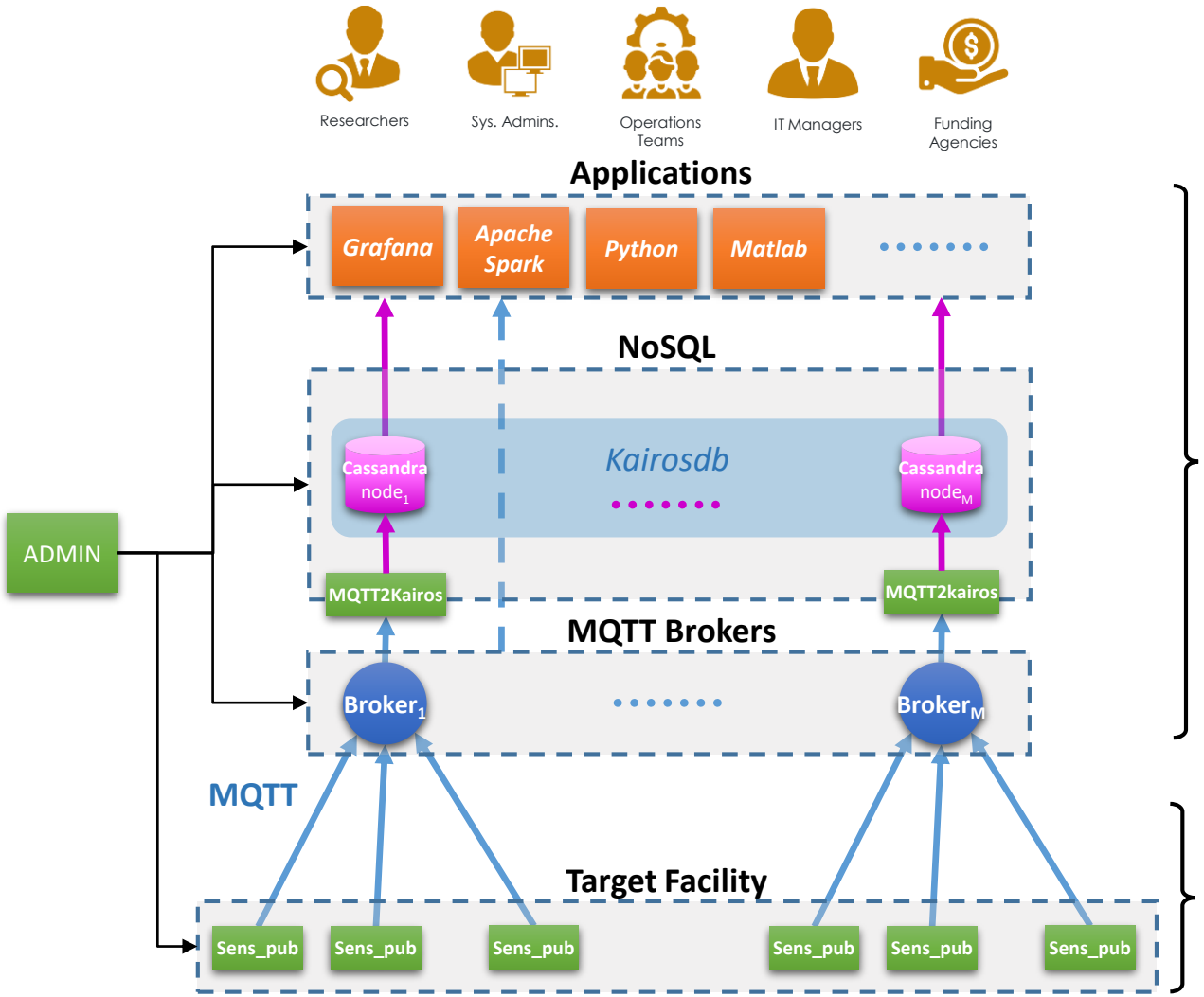
The data actually covers :

- the computing nodes internal information (980+ computing nodes) such as core loads, temperatures, frequencies, memory write/read operations, CPU power consumption, fan speed, GPU usage details, etc.,
- the system-wide information, including the liquid cooling infrastructure, the air conditioning system, the power supply units, workload manager statistics,
- job-related information,
- system status alerts, and
- weather data.

10 months of M100 system, 50 TB, Open dataset available at: https://zenodo.org/record/7541722

Plugin	#Metrics	#Plugin-specific columns	Description
Vertiv	25	1	Mainly collects data from the air-conditioning units (CDZ) located in room F (Marconi 100) of Cineca. The plugin uses the RESTful API interface available on the individual devices to extract the most interesting metrics.
Schneider	164	1	Dedicated data collector designed to acquire data from an industrial PLC by accessing its HMI module (from Schneider Electric). The PLC controls the valves and pumps of the liquid cooling circuit (RDHx) of Marconi 100. It consists of two (redundant) twin systems controllable by two identical HMI panels, Q101 and Q102.The ExaMon plugin extracts and stores all the metrics available on both panels.
IPMI	104	1	Collects all the sensor data provided by the OOB management interface (BMC) of cluster nodes.
Ganglia	177	1	Connects to the Ganglia server (gmond), collects and translates the data payload (XML) to the ExaMon data model.
Logics	37	2	Data collection system already installed at Cineca. It is specialized for collecting power consumption data from equipment in the different rooms, typically using multimeters that communicate via Modbus protocol. The ExaMon plugin dedicated to collecting this data interfaces to the Logics database (RDBMS) via its REST API. NOTE: Since the translation process is fully automated, the same inconsistencies present in the original db may result in the ExaMon database: e.g., metric names in the Italian language, units of measure as metric name, etc.
Weather	10	0	Collects all the weather data related to the Cineca facility location (Casalecchio di Reno) using an online open weather service (https://openweathermap.org).
Nagios	1	5	Interfaces with a Nagios extension developed by CINECA called "Hnagios", collects and translates the data payload to the ExaMon data model.
SLURM	54	4	Collects aggregated data from the SLURM server; these information is gathered through ad hoc scripts created by CINECA system administrators.
Job table	1	89	Collects information regarding the jobs executed on the cluster (and store in the SLURM database); the information collected are those provided by users at submission time.

ExaMon Architecture



Front-end

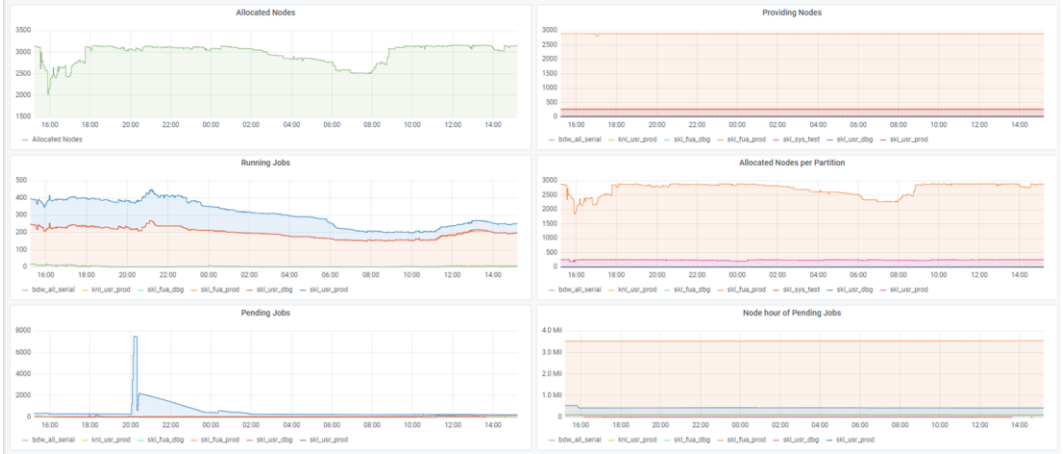
- **MQTT** Brokers
- Data **Visualization**
- **NoSQL** Storage
- **Big Data** Analytics

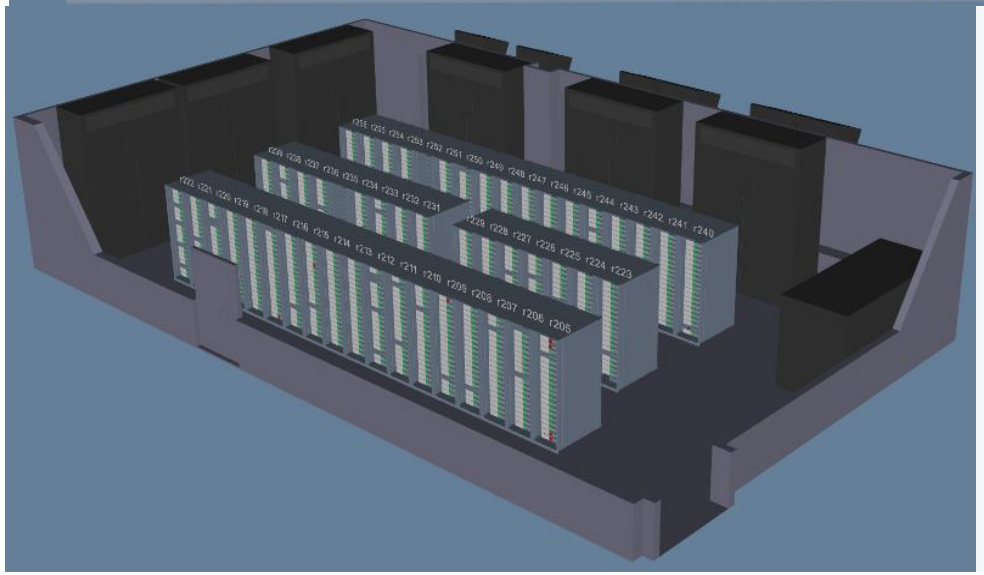
Back-end

- **MQTT**–enabled sensor collectors

Spark
DASK
presto
trino
Kubeflow
Grafana
jupyter
redash
Apache Superset
babylon.js
cassandra
KairosDB
MQTT

Digital twin of Marconi 100





CPU Util.		Memory Util.		Free CPUs	
bdw_all_serial	2.78%	bdw_all_serial	5.21%	bdw_all_serial	70
knl_usr_prod	22.38%	knl_usr_prod	20.05%	knl_usr_prod	60384
skl_fua_dbg	0%	skl_fua_dbg	0%	skl_fua_dbg	576
skl_fua_prod	99.34%	skl_fua_prod	93.45%	skl_fua_prod	912
skl_sys_test	0%	skl_sys_test	0%	skl_sys_test	192
skl_usr_dbg	11.11%	skl_usr_dbg	7.64%	skl_usr_dbg	768
skl_usr_prod	83.59%	skl_usr_prod	68.84%	skl_usr_prod	2016

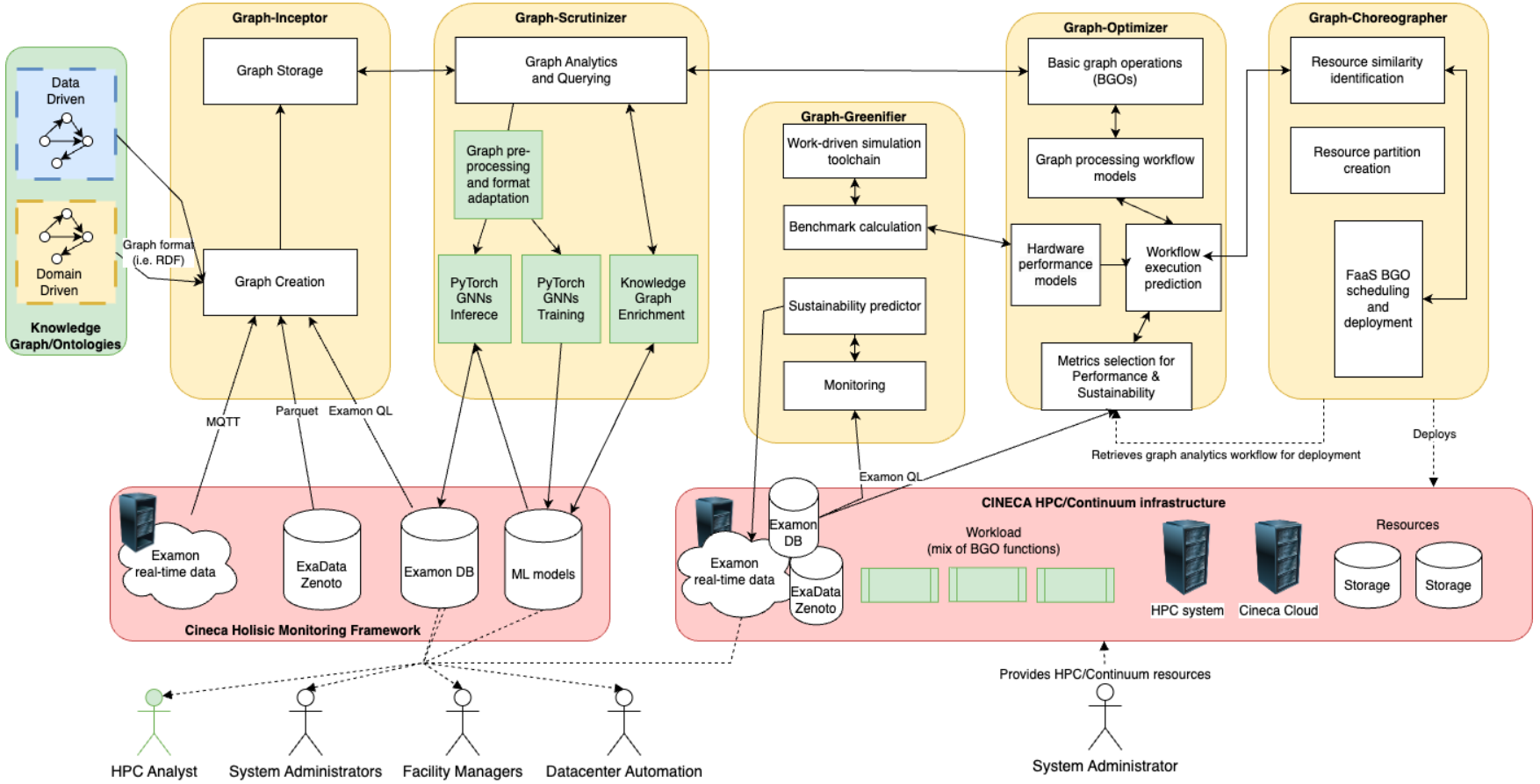
Graph representation of monitoring data

Graph models can capture the spatiotemporal dependencies between computation, nodes, and cooling equipment.

GNNs are well-suited for processing graph-structured data and leverage the spatial proximity of compute nodes in an HPC room.

Previous scientific research from UniBO showed that GNN outperforms other methods in anomaly detection (AUC = 0.86 vs 0.73, at 8 hours)

GM solution for the Data Centre Digital Twin



Identification of Operational Data Analytics (ODA) priorities

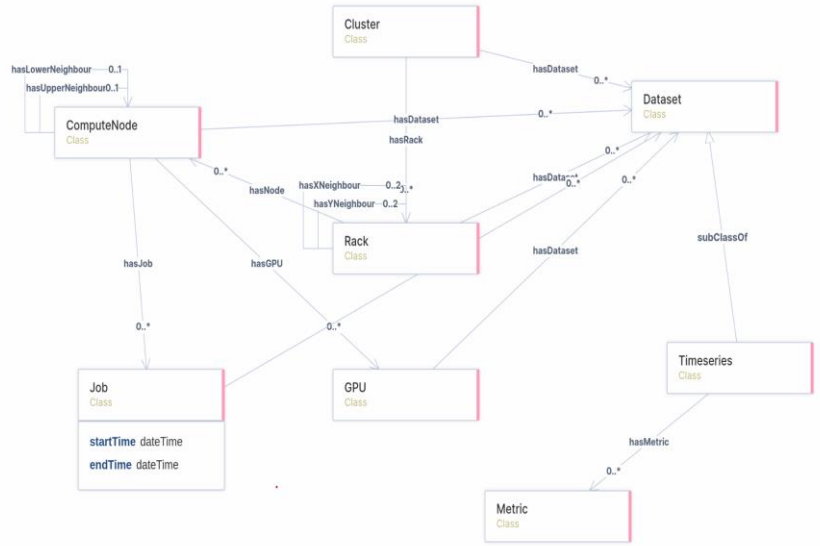
Internal stakeholders:

- HPC – production (user support)
  - Availability,
  - Waiting time,
  - Job performance
- System administrators
  - Predictive maintenance
- Facility manager
  - Energy efficiency (PUE),
  - Holistic predictive maintenance,
  - Capacity planning
- HPC – architecture (evolution)
  - Resource design,
  - Sizing of the network topology

- BIM, Building Information Modeling (3D model integrated with thermal and cooling data), not tailored for a data center room
- Need a real integration of facility data and computational resource data
- A predictive model of the PUE, Power Usage Effectiveness, would enable cooling optimization and avoid waste (anticipate cooling demand)

Current status: Knowledge graph

- What is the best graph representation? What is a node/vertex, what are its properties, what do edges mean?
- This has an impact on the graph dimension and on the efficiency of the BGO that we need to execute.
- The KG is also query driven (not only which data we want to represent but also to answer which questions)



Expected results and KPI

- Two-fold improvement in data center energy efficiency
- 25% lower GHG emissions
- Insights on the usage of the infrastructure and on dependencies
- GM framework and DC KG ready and available to address new ODA targets (predictive maintenance, strategic planning, …)
- DC KG openly released to be used by researchers on large graph analytics and on operational data analytics

Impact

- HPC market crucial for European industrial development
- European Commission invests over €8 billion in HPC
- CINECA is one of the leading European Supercomputing Centres and Big Data Hub
- CINECA is committed to improve societal impact and scientific throughput
- CINECA enables drug discovery (e.g. COVID-19) and fusion energy research
- Graph Massivizer tools used for modeling Leonardo supercomputer as a digital twin
- Graph-Greenifier tool for accurate sustainability-performance analysis

DATAWEEK

JOIN.LEARN.SHARE.GET VALUE

Thank you!

EUHUBS4DATA

BDV BIG DATA VALUE ASSOCIATION