



Datacenter Automation: High-Performance Computing Monitoring and Management

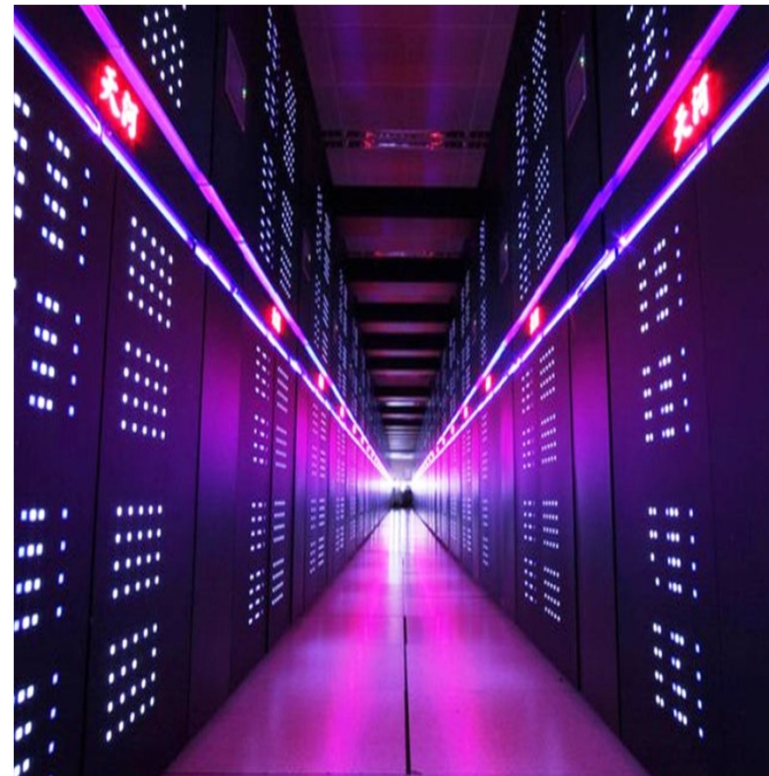
Andrea Bartolini, Luca Benini

The Energy Efficiency Embedded System (EEES) Lab.
Università di Bologna, DEI



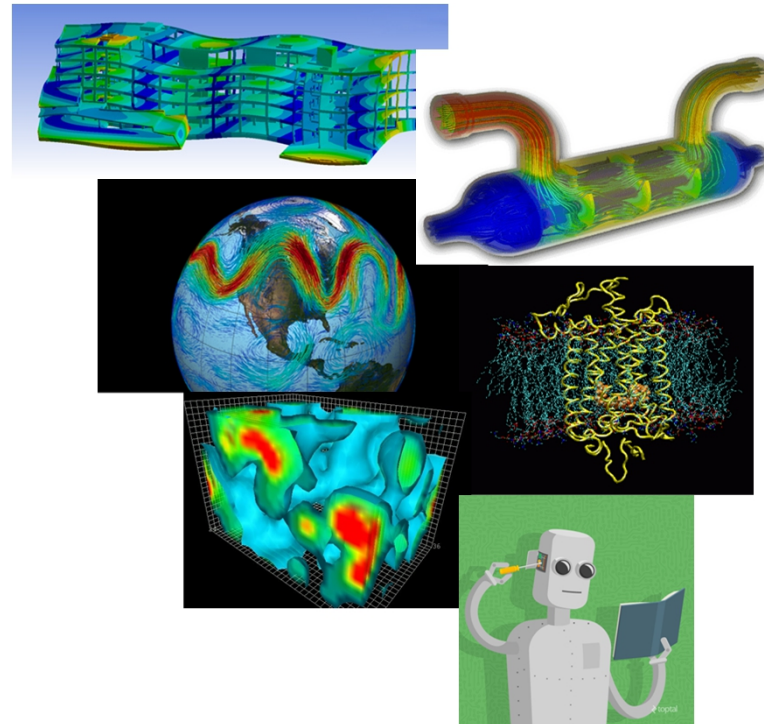
HPC – High Performance Computing

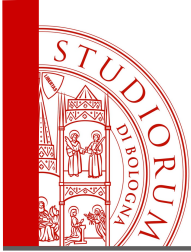
- High-performance computing (HPC) is the use of parallel processing for running advanced application programs efficiently, reliably and quickly
- Typical users: scientific researchers, engineers, data analysts



Typical applications

- Structural calculations
- Computational fluid dynamics
- Climate and weather research
- Geophysics
- Molecular dynamics
- Lattice QCD
- Astrophysics
- Data analytics
- Machine learning
-and many more

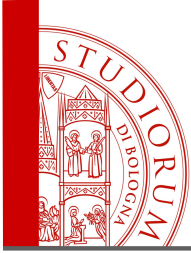




Clusters / HPC

- Clusters/HPC:
 - Composed of a collection of Desktop Computers and/or Servers which are called nodes
 - The computing nodes are interconnected by a local area networks
 - All the nodes works together as they was a single computer.
 - Each node has its own operating system and communicate with network protocols (eg: ethernet, infiniband)





Cost and performance...

- Total cost of ownership (TCO)
- CAPEX – capital costs
 - IT provisioning
 - Cooling infrastructure provisioning
 - Facility and Power provisioning
- OPEX – operational costs
 - IT and Cooling Energy Costs
 - Worker Salary



In supercomputing CAPEX is 2-4x OPEX

=> Performance & availability Matters!!!

Defenently not an embedded system, so why
@IWES?



Bologna the city of the three...



...Datacentres.



Bologna the city of the three...

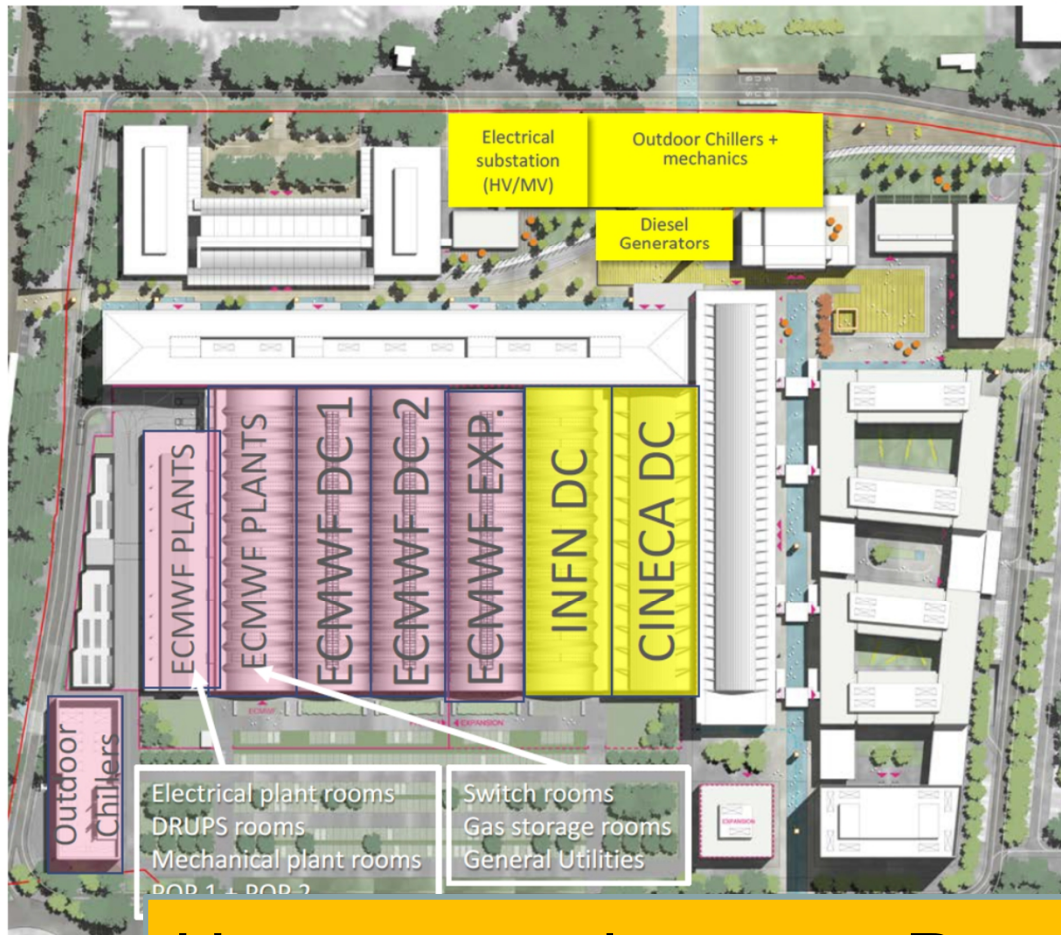


...Datacentres.



Why do we care - tomorrow

The data centers at the Science Park



ECMWF DC main characteristics

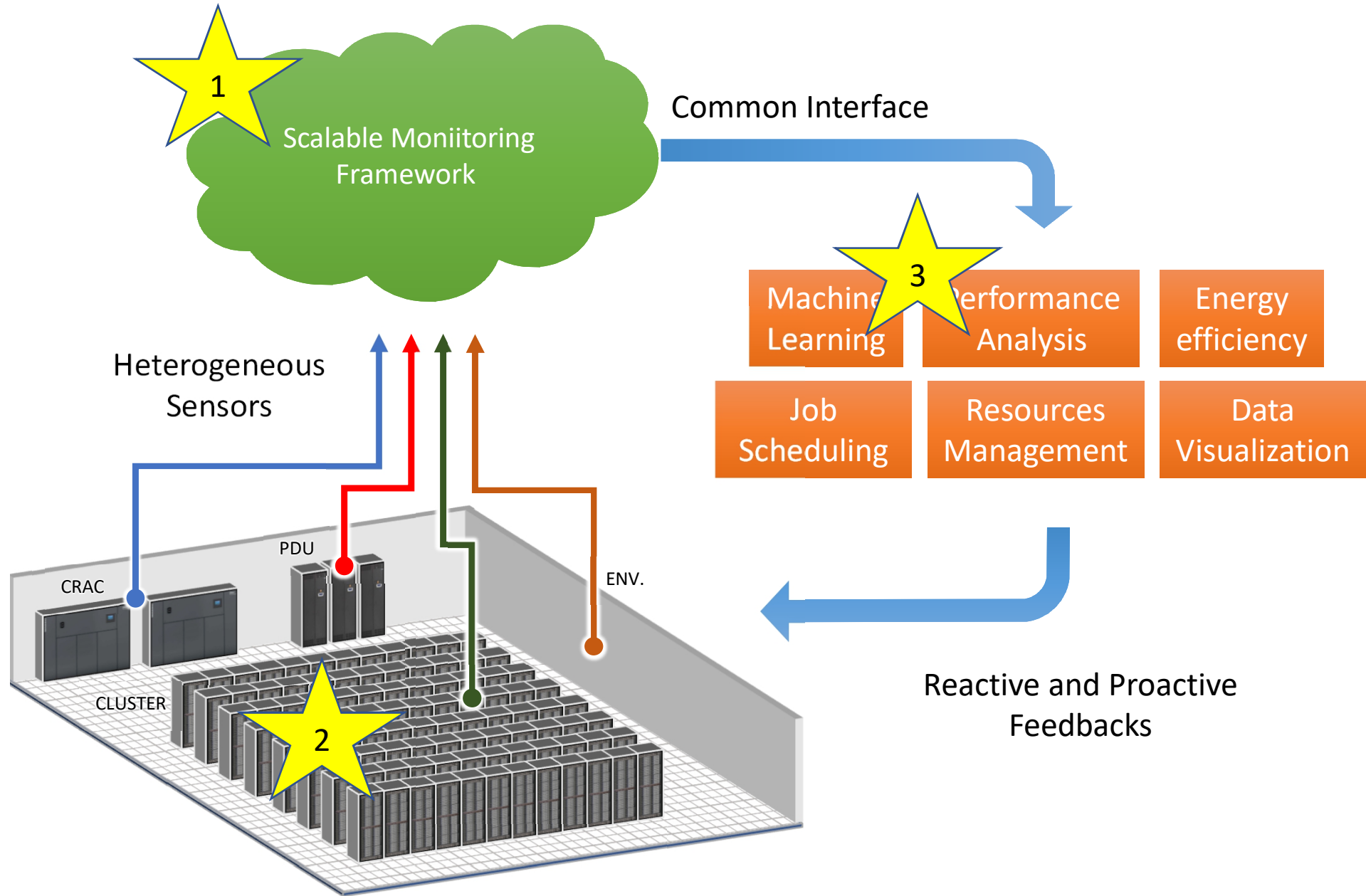
- 2 power line up to 10 MW (one bck up of the other)
- Expansion to 20 MW
- Photovoltaic cells on the roofs (500 MWh/year)
- Redundancy N+1 (mechanics and electrical)
- 5 x 2 MW DRUPS
- Cooling
 - 4 dry coolers (1850 kW each)
 - 4 groundwater welles
 - 5 refrigerator units (1400 kW each)
- Peak PUE 1.35 / Maximum annualized PUE 1.18

INFN – CINECA DC main characteristics

- up to 20 MW (one bck up of the other)
- Possible use of Combined Heat and Power Fuel Cells Technology
- Redundancy strategy under study
- Cooling, still under study
 - dry coolers
 - groundwater welles
 - refrigerator units

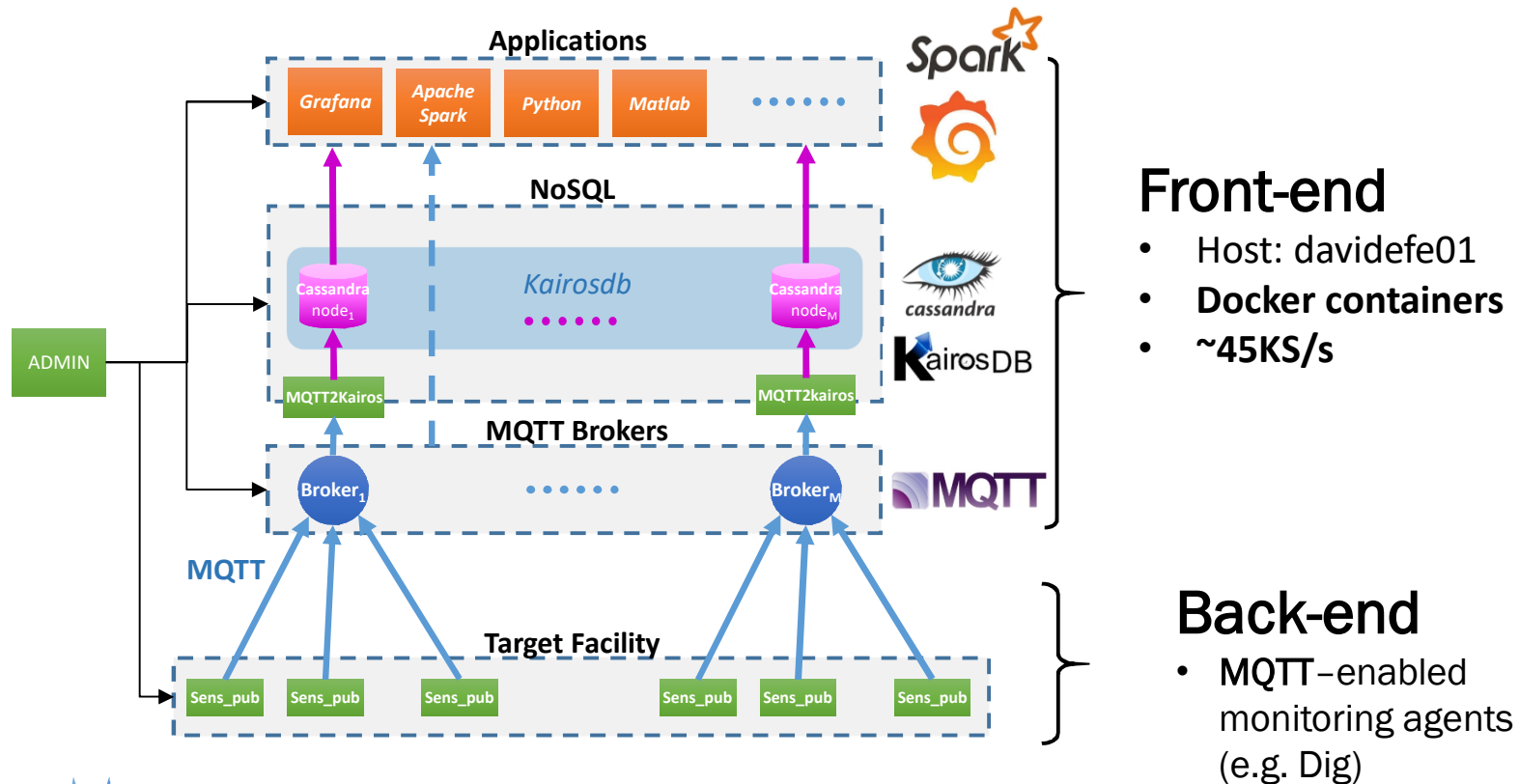
How can we increase Power and Cost efficiency?

A New Trend: Datacentre Automation

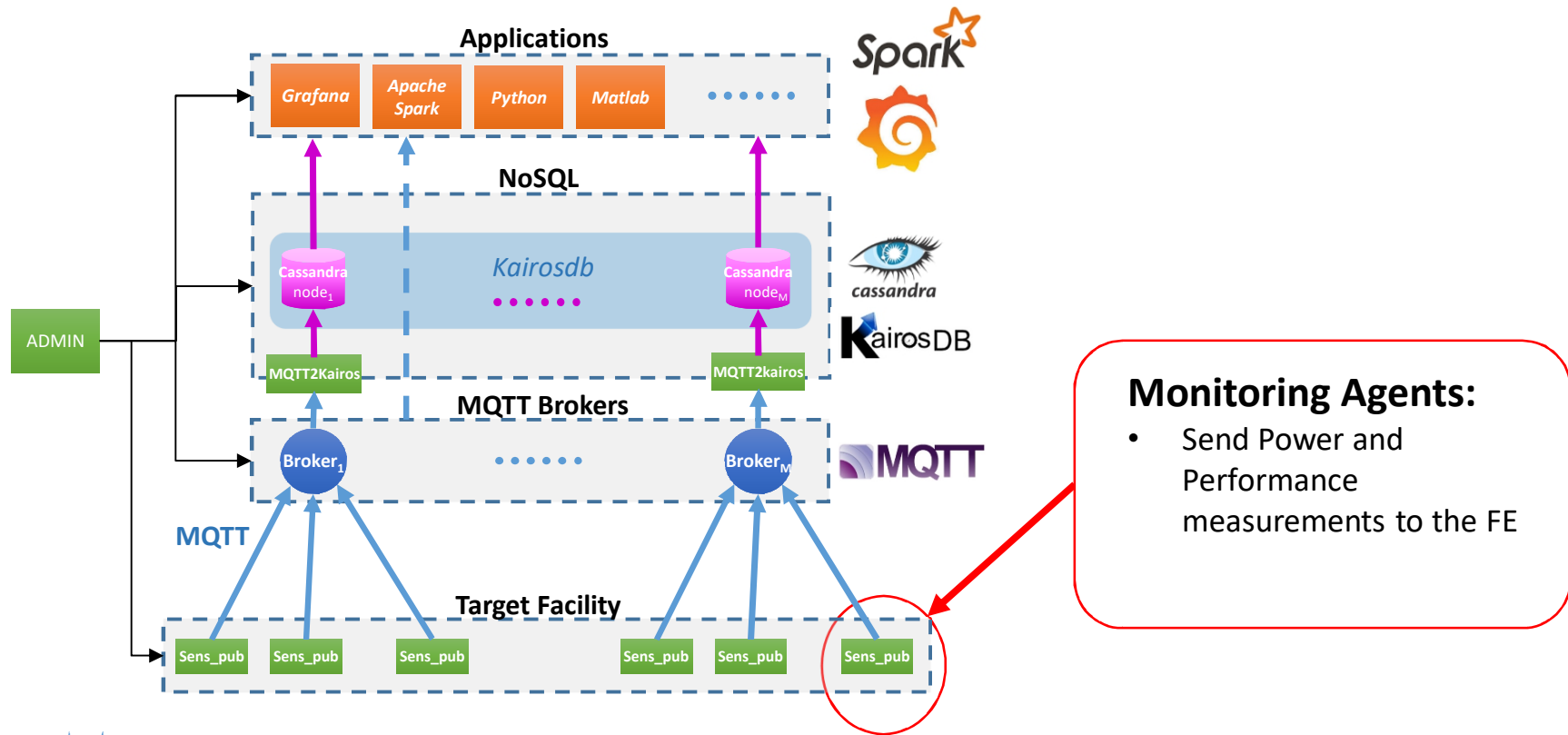




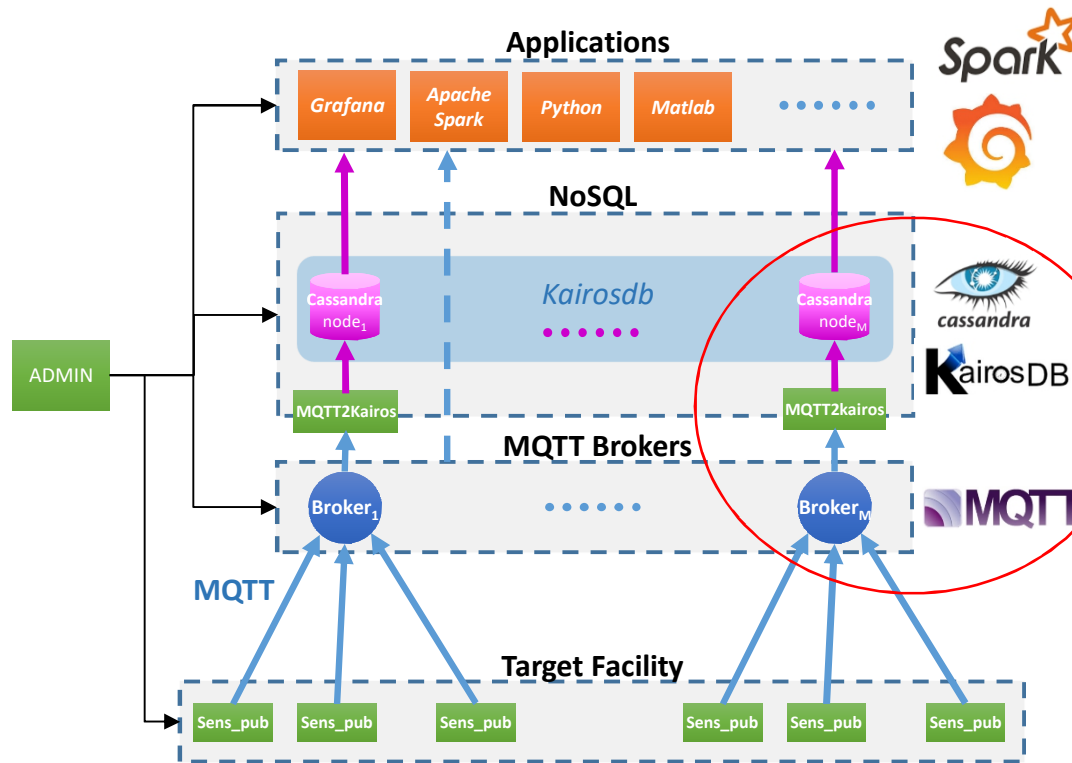
ExaMon: an Industry 4.0 approach to datacenter automation



ExaMon: an Industry 4.0 approach to datacenter automation



ExaMon: an Industry 4.0 approach to datacenter automation



Broker:

- Forward data to the listeners (e.g. kairosDB)

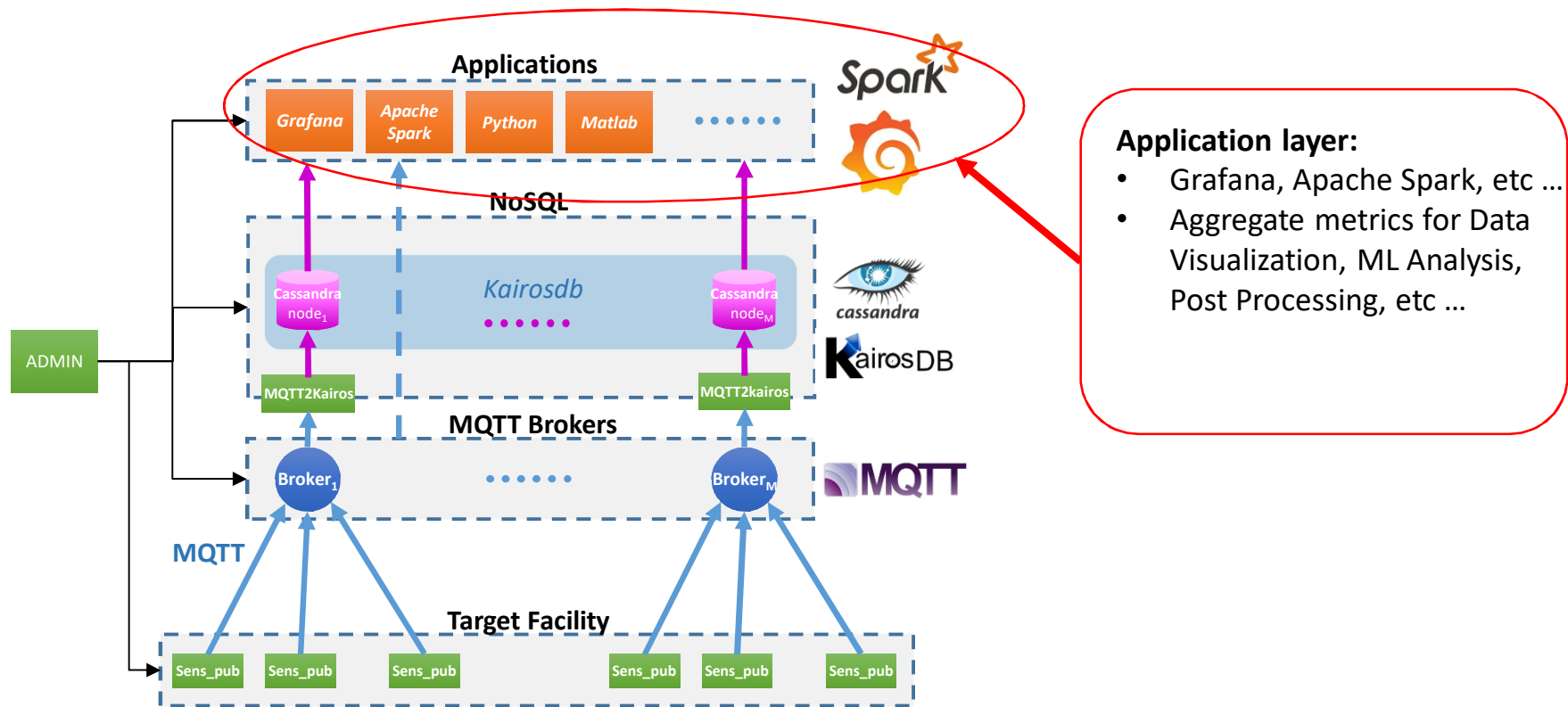
Mqtt2kairosdb:

- Interface between MQTT and KairosDB
- KairosDB is a front-end to handle time series in Cassandra

Cassandra:

- NoSQL database
- Highly scalable
- Optimized to balance the load on multiple nodes

ExaMon: an Industry 4.0 approach to datacenter automation



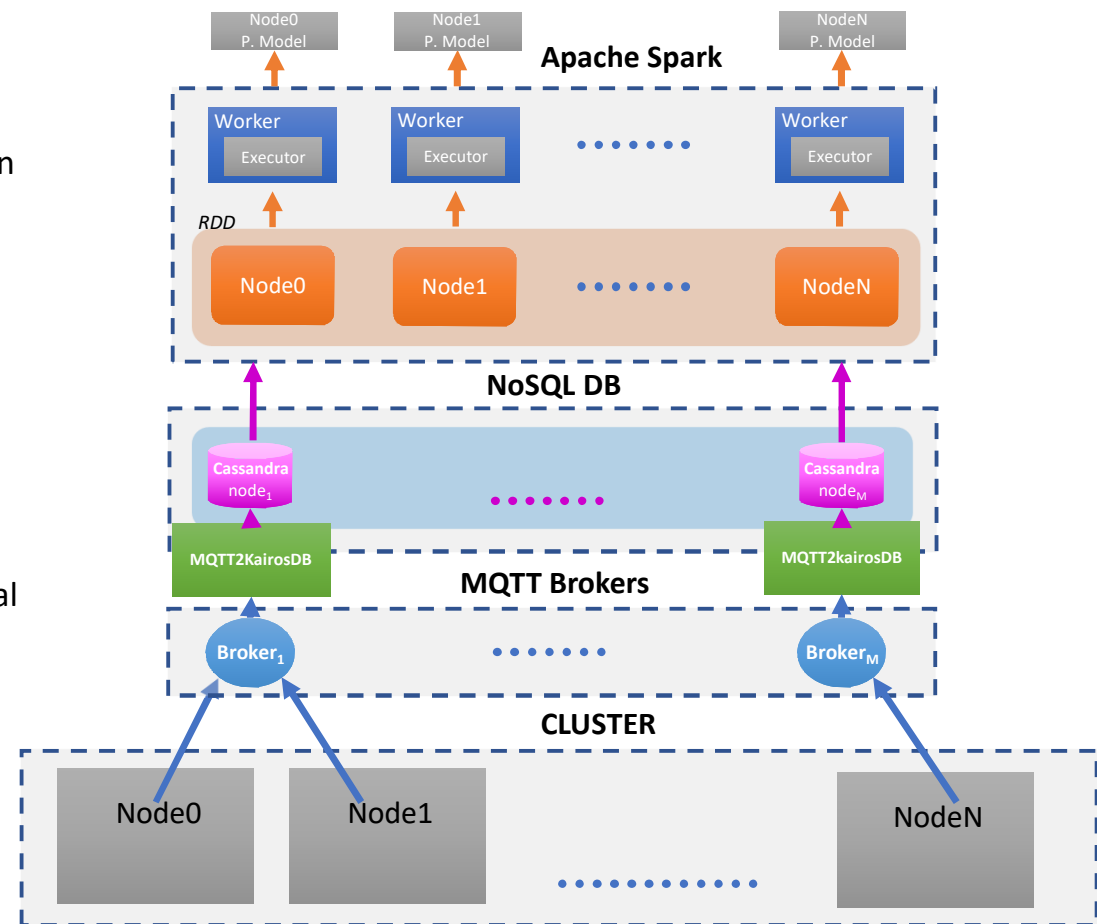
The Big Data & DL backbone

Computing clusters

- Not only the computing engine of Big Data solution
- Also a complex industrial plant and a growing industry in ER
- A compute nodes can produces $\sim 100/1000$ metrics/s * “peta/exa scale” = Big Data!

Datacenter automation – improve energy/cost efficiency and effectiveness – Industry 4.0 thanks to:

- Live collection and processing of large telemetry data (>100GB/day x cluster)
- On-line generation of “plant models” – a.k.a. digital tweens”, security break detection and HW fault prediction!





D.A.V.I.D.E. (#18 Green500 Nov'17)

E4

COMPUTER
ENGINEERING

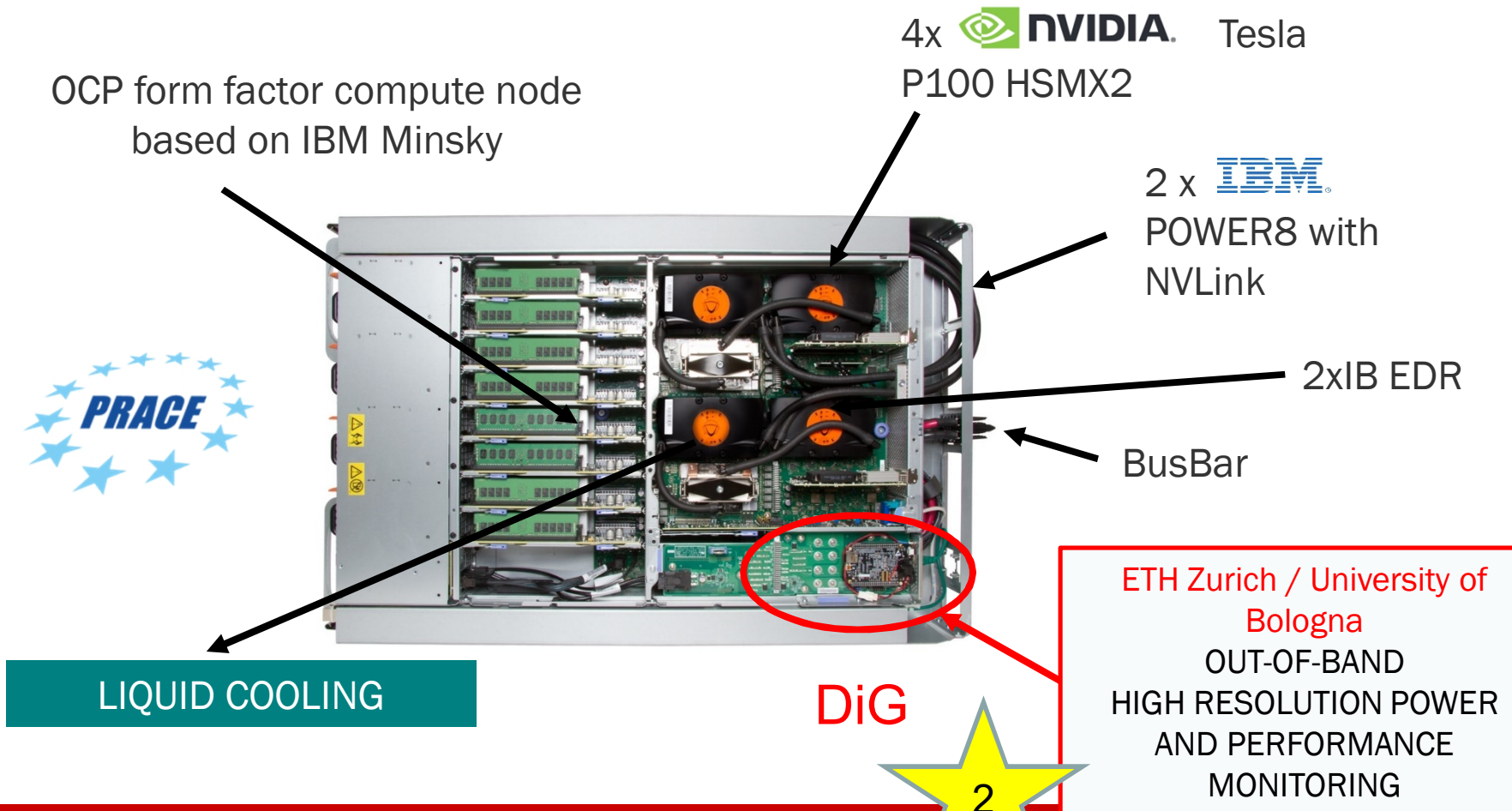
D.A.V.I.D.E.
SUPERCOMPUTER
(Development of an
Added
Value
Infrastructure
Designed in
Europe)





D.A.V.I.D.E. SUPERCOMPUTER

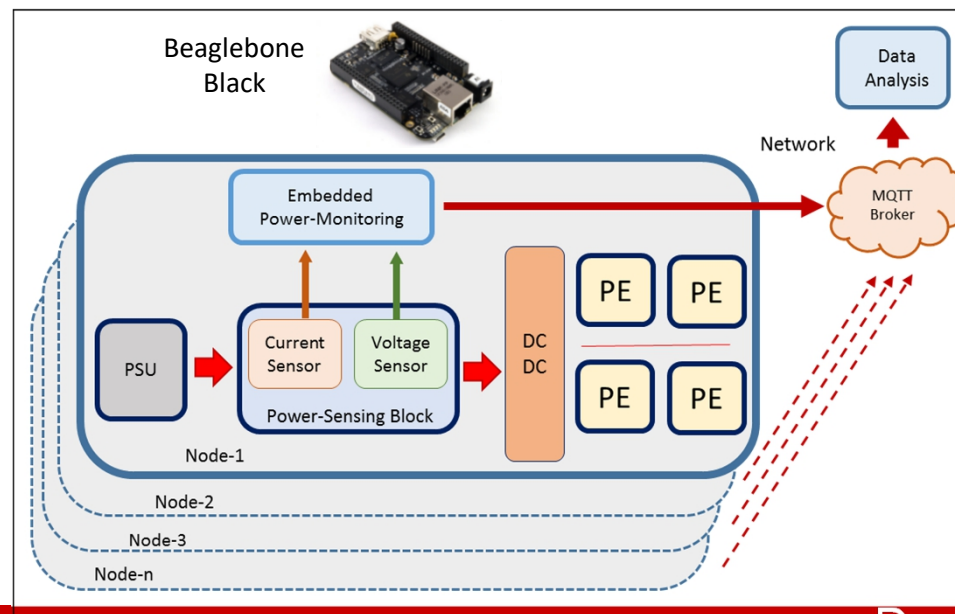
(Development of an Added Value Infrastructure Designed in Europe)





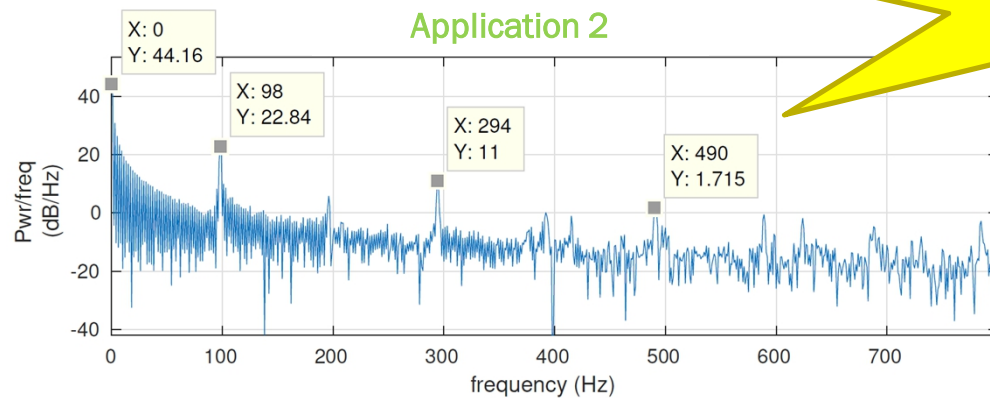
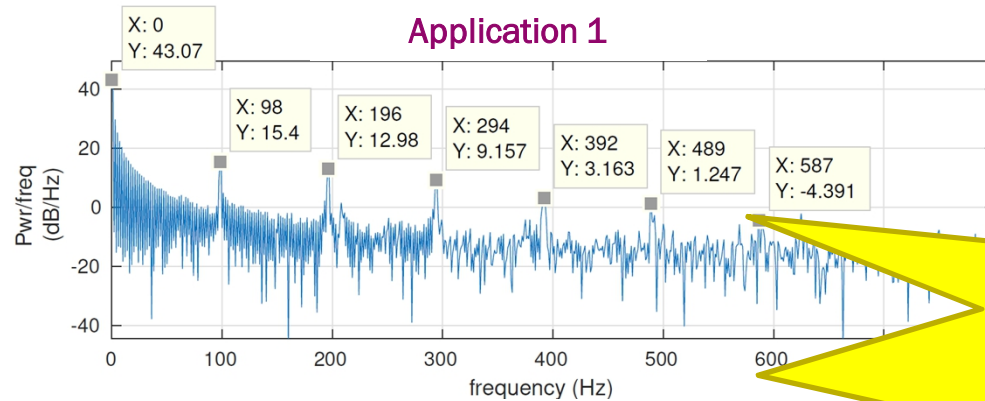
DiG: High Resolution Out-of-band Power Monitoring

- Out-of-band → Zero overhead
- Collect more than 1.5 kS/s, 7/7d, 24/24h, for all users
- Architecture independent (i.e. tested on Intel, ARM and IBM)
- Fine grain → down to ms scale (sampling @800 kS/s + avg)
- IoT communication technology (MQTT) → scalable
- Edge Computing!



DiG: live FFT on the power traces

Real-time Frequency analysis on power supply and more...a live oscilloscope

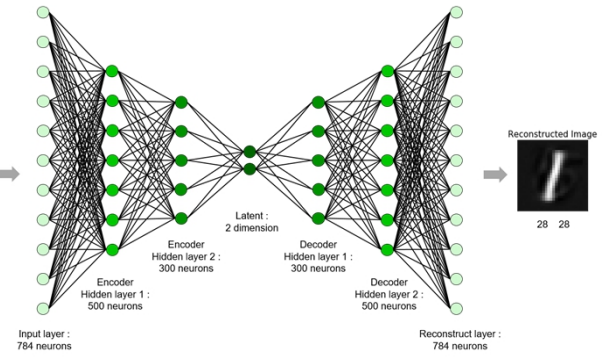
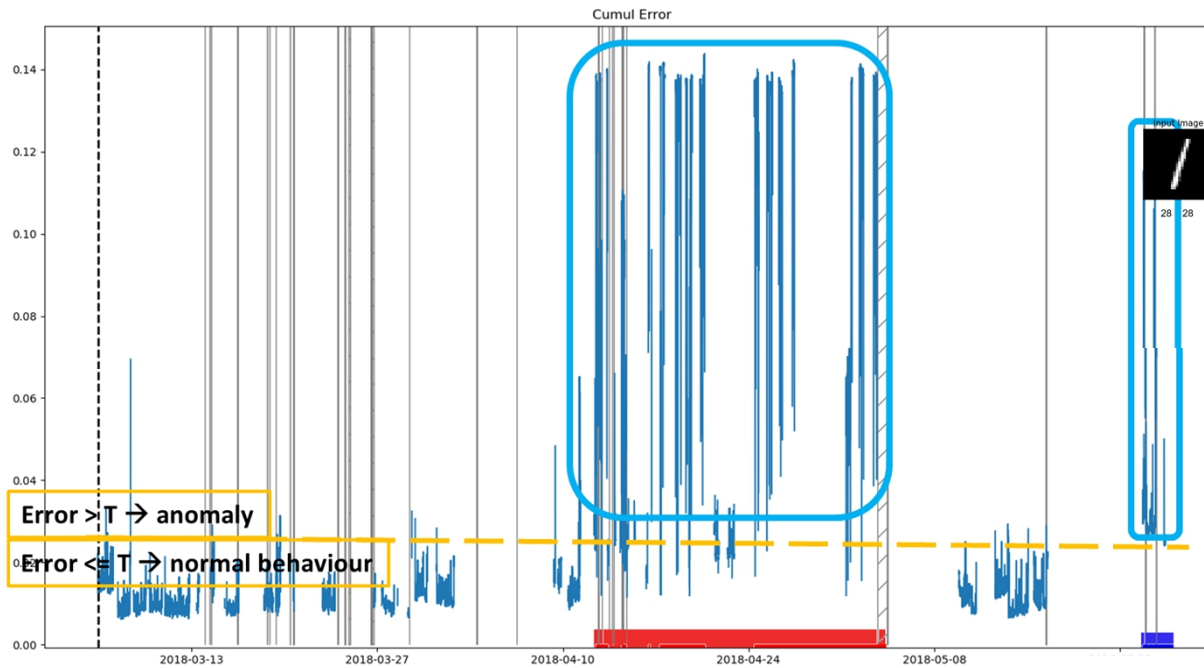


Upcoming - live
oscilloscope x 45
nodes @ 400KHz

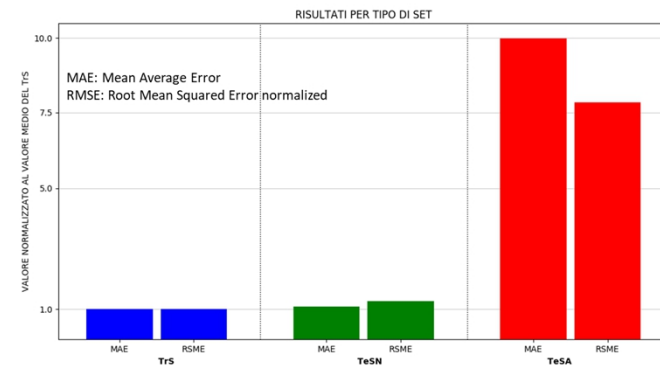


- **User** -> to discriminate **application phases**
- **Sys Admin** -> to detect **malicious users**
- **Designers** -> to debug and optimize **power delivery network**

AI+ Big Data on D.A.V.I.D.E.: Anomaly detection



Fault!! 



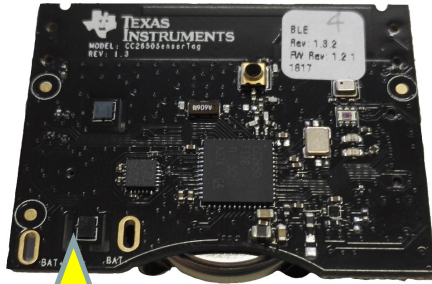
Train Set (no anomalies)

Test Set (no anomalies)

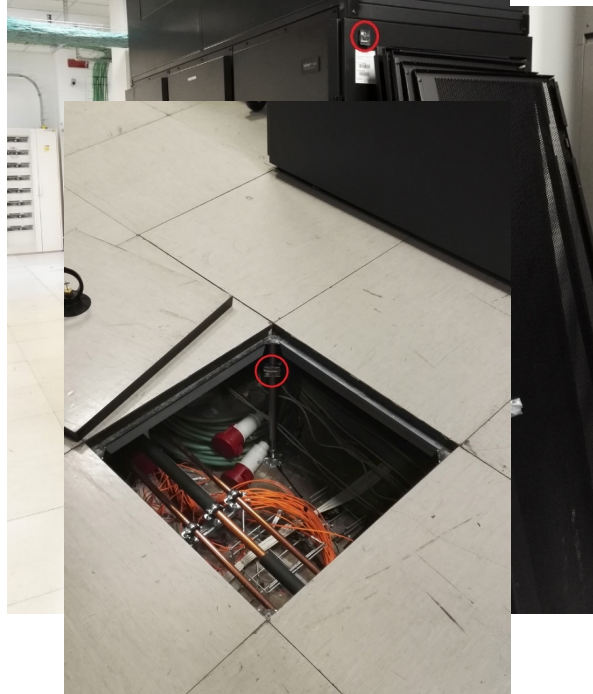
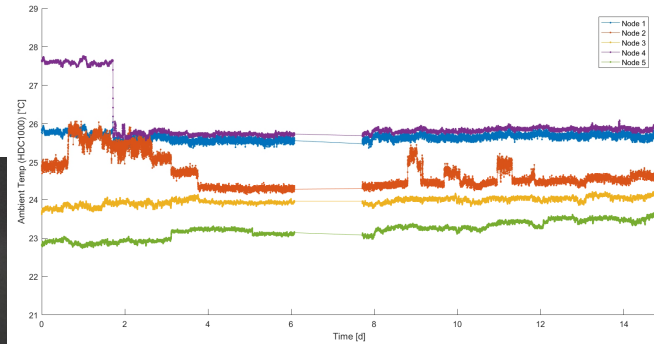
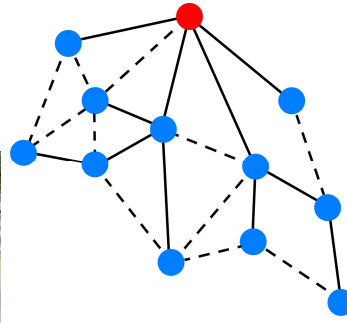
Test Set (with anomalies)



Datacenter monitoring network

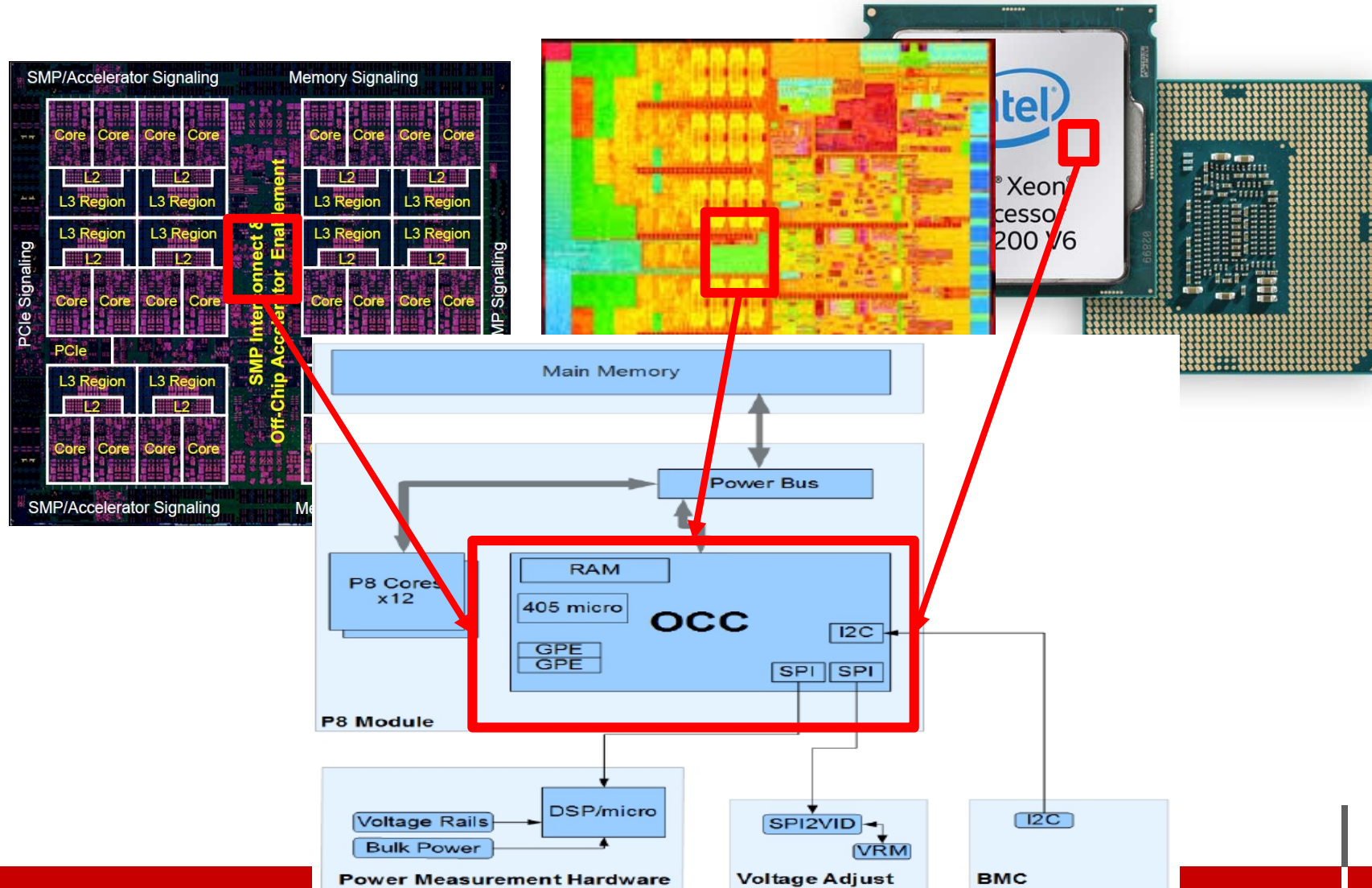


- Used to optimize the datacenter cooling desing





Processors power and thermal control



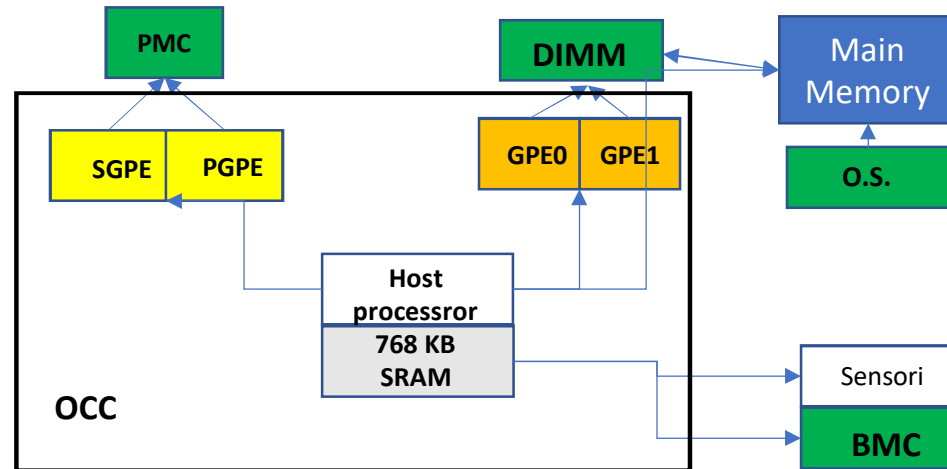
OCC: an heterogeneous embedded system:

Host processor with
768KB private SRAM
SRAM shared with BMC
and sensors

Two general purpose acc:
GPE0, GPE1 for offloading
control and monitoring
tasks.

Two special purpose acc:
SPGE and PGPE.

Interaction with O.S.
through DRAM



Services:

- **Temperature and Power Monitoring**
- Performance management for the **Processor (Frequency) and memory (Bandwidth)**

Functionalities:

- **Power Capping**
- **Thermal Management**
- **Workload optimized Frequency(WOF)**

Can we do it better?

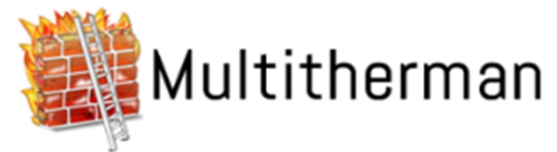


European Processor Initiative

Acknowledge

EEES lab collaborators:

Andrea Borghesi, Francesco Beneventi , Daniele Cesarini, Antonio Libri, Federico Pittino, Christian Conficoni, Roberto Diversi, Andrea Tilli, Michela Milano



Pan-European partners



Scalability allows wide market potential coverage

