Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

# MULTITHERMAN:
## Out-of-band High-Resolution HPC Power and Performance Monitoring Support for Big-Data Analysis

**HPC Summit Week, Ljubljana**
30.05.2018

Antonio Libri
A. Bartolini, F. Beneventi, A. Borghesi and L. Benini

EU H2020 FETHPC project ANTAREX (g.a. 671623)

EU FP7 ERC Project MULTITHERMAN (g.a.291125)

# Team & Research Activity

A. Bartolini, and L. Benini

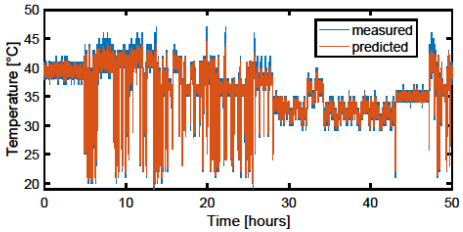
Multitherman

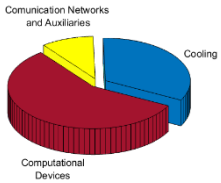
EX 10^18

EXAMON

Topics:

1. Examon: holistic and scalable real-time monitoring for HPC centers – open source, beta release Q2/Q3 2018 (F. Beneventi, https://github.com/fbeneventi/examon)
2. Large-scale multiscale Power and Thermal modeling, and control (F. Pittino, C. Conficoni)
   a. Development of power and thermal models at core level for monitoring and control of large HPC clusters in production
   b. Datacentre cooling modeling and optimization

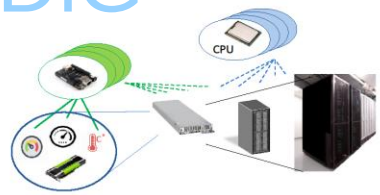


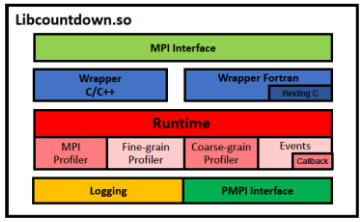



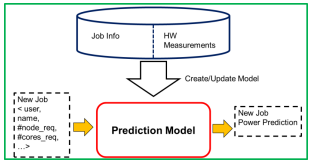
# Team & Research Activity

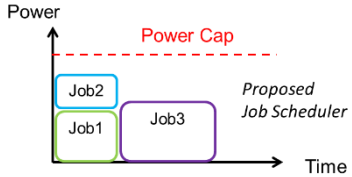
3. Fine grain monitoring and management (A. Libri, D. Cesarini)
   a. DiG: multi-platform and production ready solution, based on low-cost open-source HW, already integrated with E4 technology
   b. Application aware power and thermal management: open source, beta release Q3 2018
4. Power and performance aware Job scheduler (A. Borghesi)
   a. System level power capping and optimization
5. Datacentre automation (A. Borghesi, A. Libri)
   a. Big-data and deep learning based automated power optimization
   b. Big-data and deep learning based automated anomaly detection

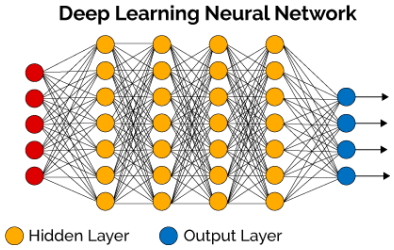
DiG



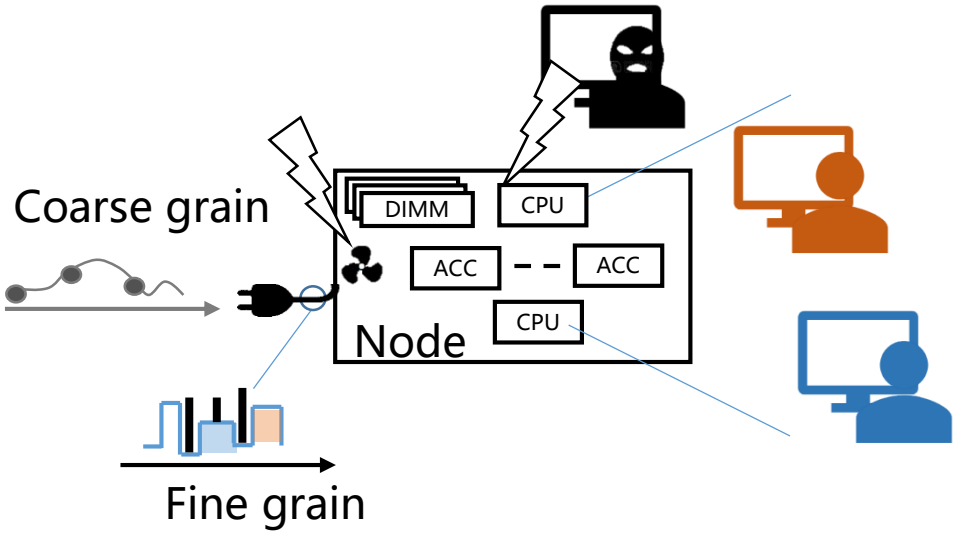



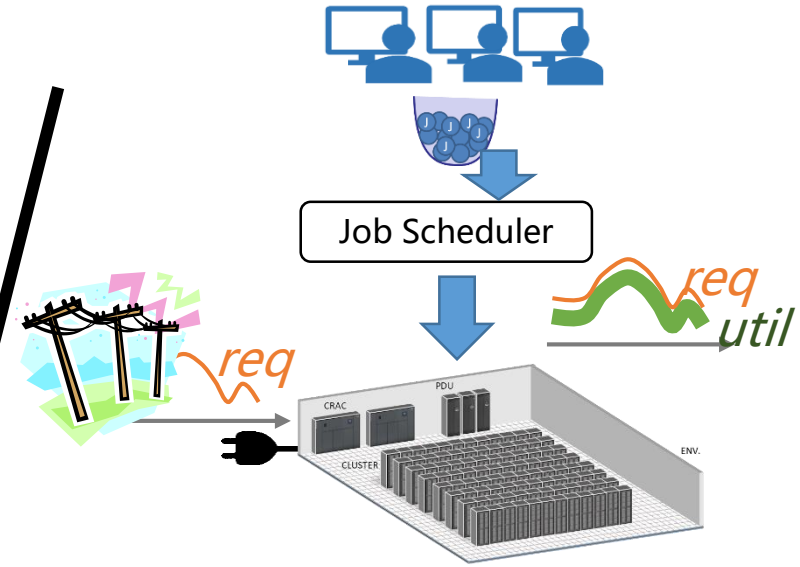



# Outline

- Motivation & D.A.V.I.D.E. Overview
- DiG: High Res Out-of-Band Pow & Perf Mon
- ExaMon: Scalable Data Collection and Analytics
- Monitored Metrics and Data Visualization
- Smart Job Scheduler – SLURM Plugin

# Motivation – Innovative Scenarios




Fine Grain Power and Performance Measurements:

- Verify and classify node performance (e.g. In spec / out of spec behaviour, aging and wear out)
- Predictive maintenance
- Per user - Energy / Performance – accounting

System Power Capping (bound from grid):

- Job scheduler decides the job that enters in the system
- Ensures operating power below a maximum power consumption level
- Reduce Power budget during new Installations and Power Shortage

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

E4 COMPUTER ENGINEERING

D.A.V.I.D.E. SUPERCOMPUTER
(**D**evelopment of an **A**dded **V**alue **I**nfrastructure **D**esigned in **E**urope)

PRACE

TOP 500 The List. THE GREEN 500

OpenPOWER

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

## (Development of an Added Value Infrastructure Designed in Europe)

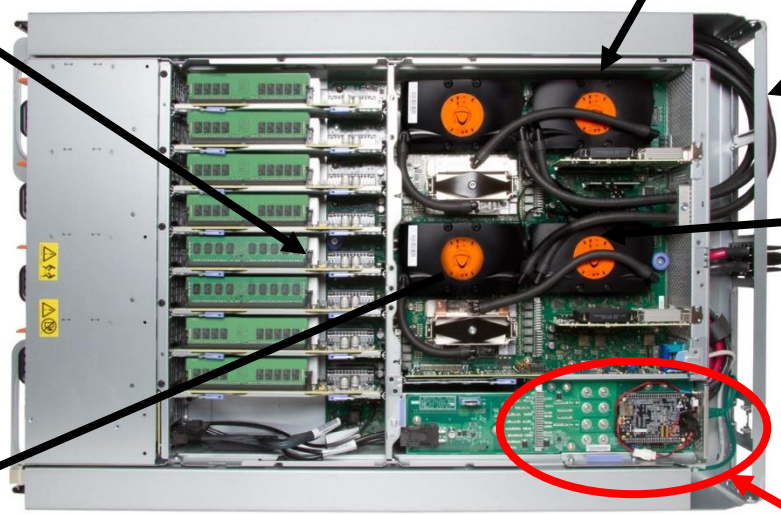
OCP form factor compute node based on IBM Minsky

4x NVIDIA Tesla P100 HSMX2

2 x IBM POWER8 with NVLink

2xIB EDR

BusBar



PRACE

LIQUID COOLING

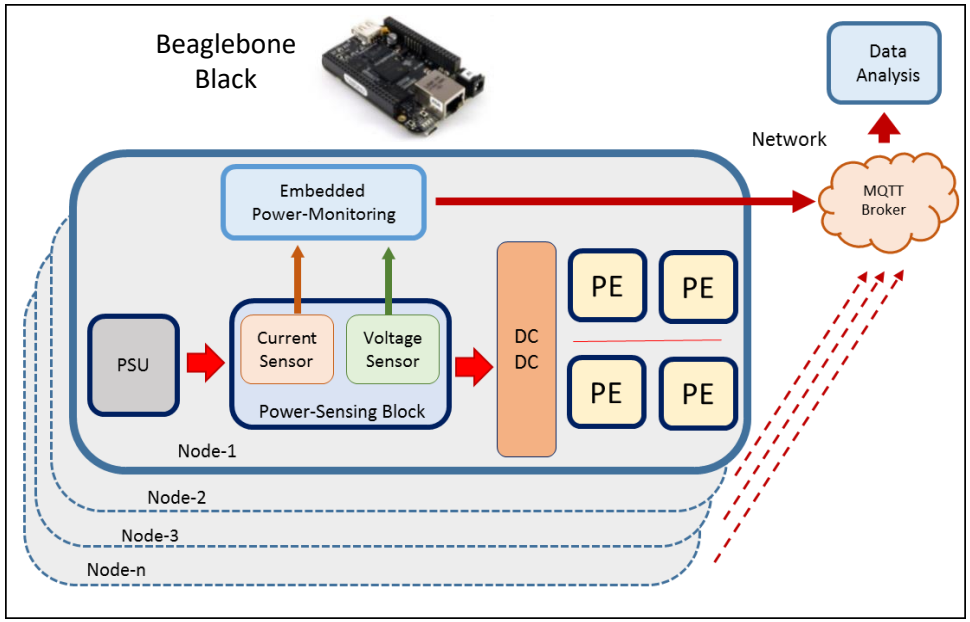
DiG

ETH Zurich / University of Bologna
OUT-OF-BAND
HIGH RESOLUTION POWER AND PERFORMANCE MONITORING

# Outline

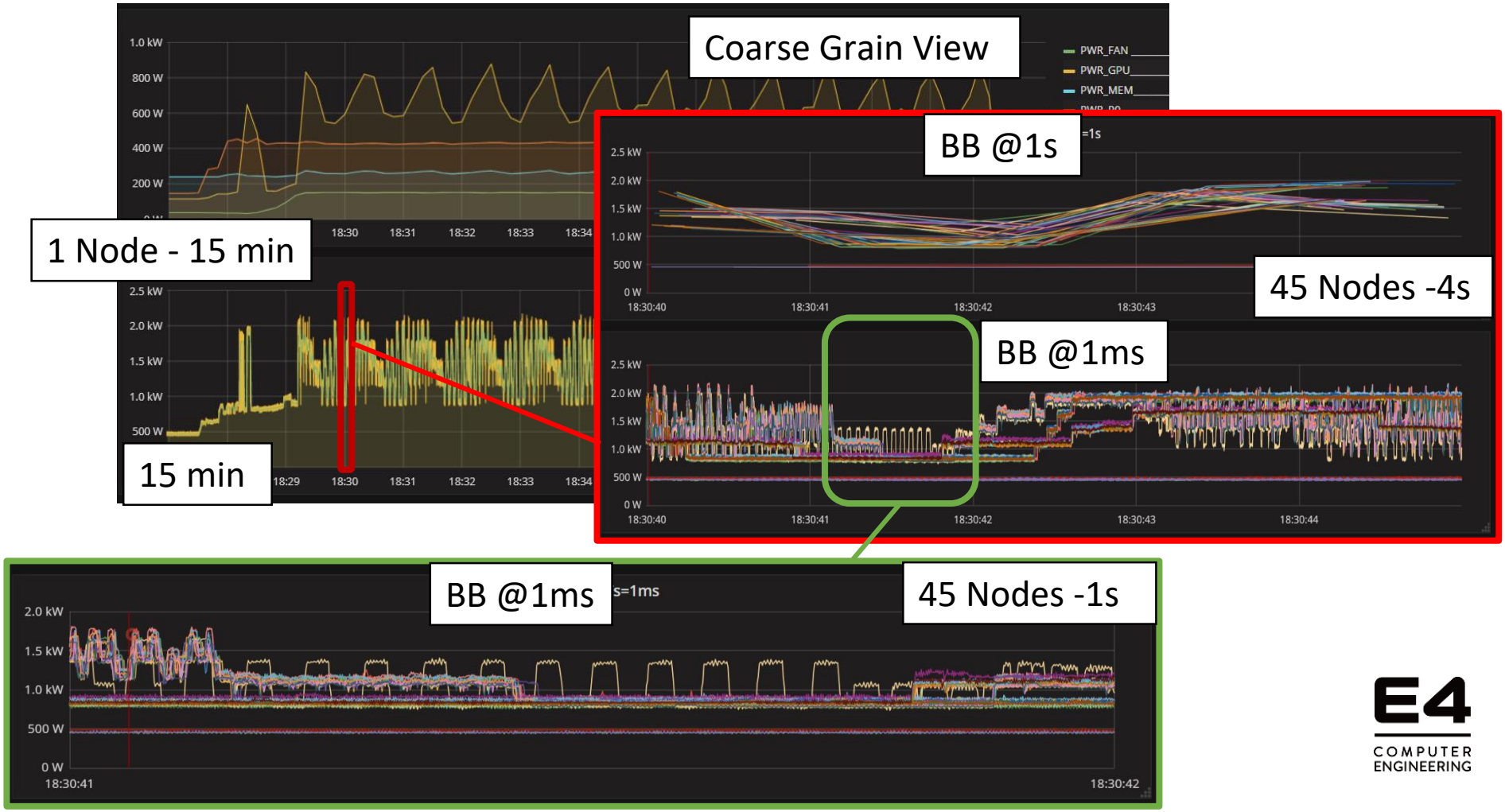
- Motivation & D.A.V.I.D.E. Overview
- DiG: High Res Out-of-Band Pow & Perf Mon
- ExaMon: Scalable Data Collection and Analytics
- Monitored Metrics and Data Visualization
- Smart Job Scheduler – SLURM Plugin

# DiG: High Resolution Out-of-band Power Monitoring

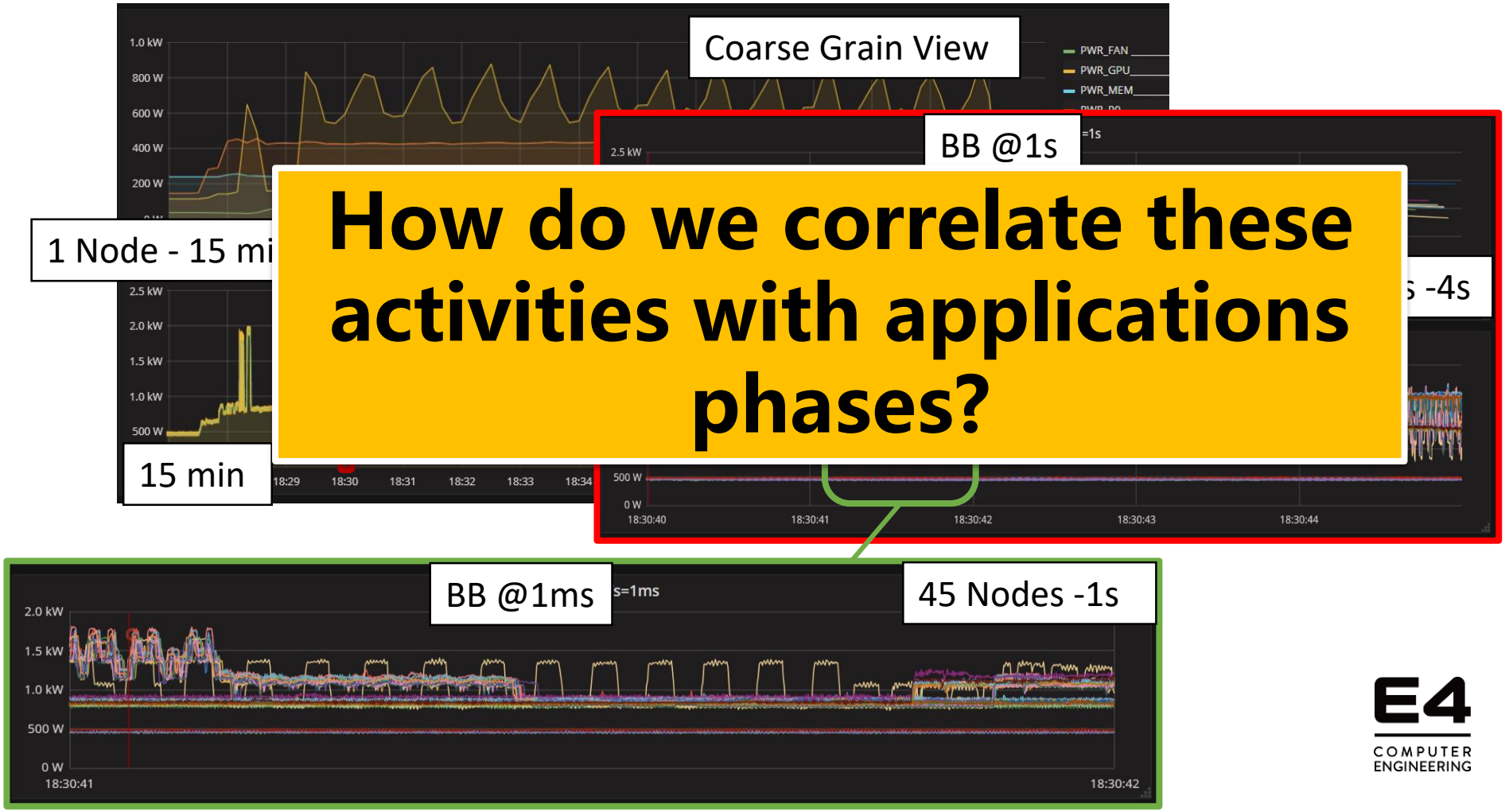



- Out-of-band → Zero overhead
- Collect more than 1.5 kS/s, 7/7d, 24/24h, for all users
- Architecture independent (i.e. tested on Intel, ARM and IBM)
- Fine grain → down to ms scale (sampling @800 kS/s + avg)
- IoT communication technology (MQTT) → scalable
- Time synchronous (NTP, PTP)

# DiG: Example of fine grain monitoring



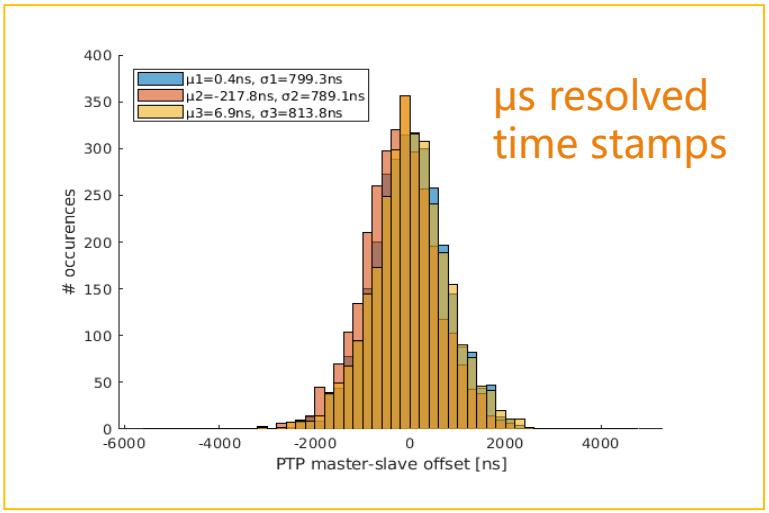
# DiG: Example of fine grain monitoring

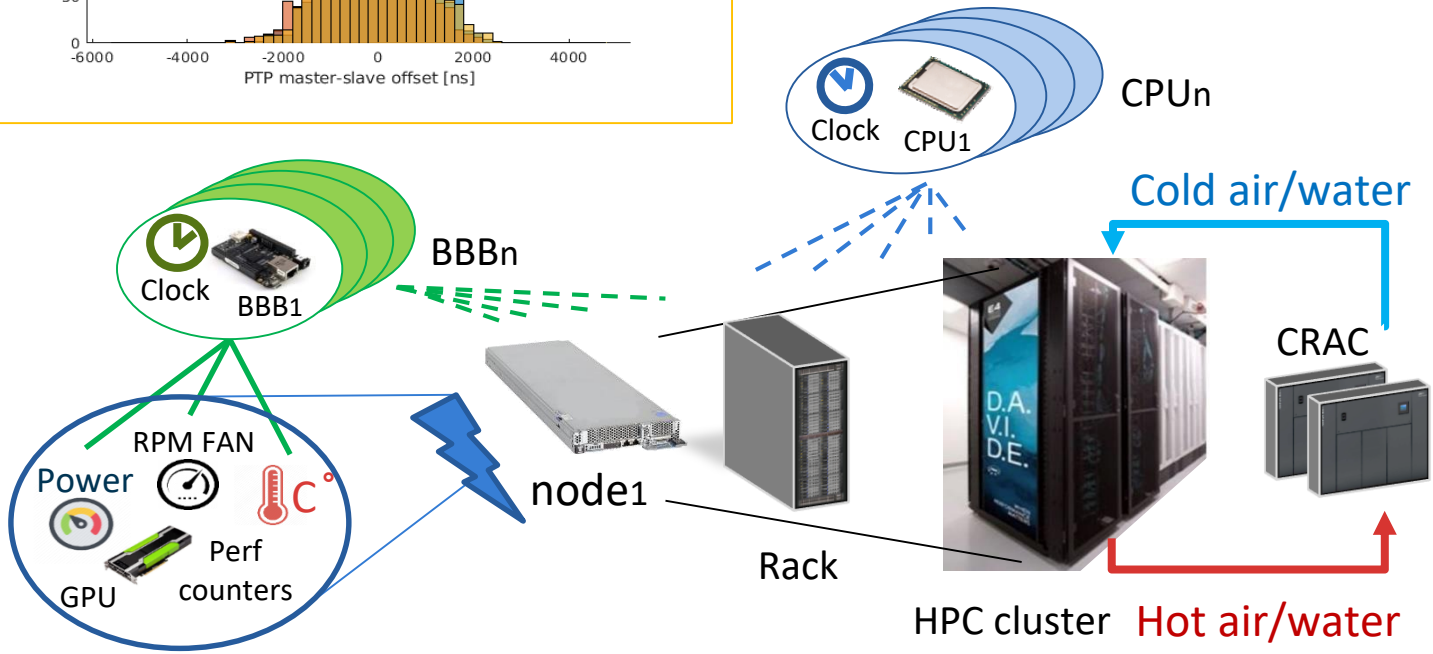



**How do we correlate these activities with applications phases?**

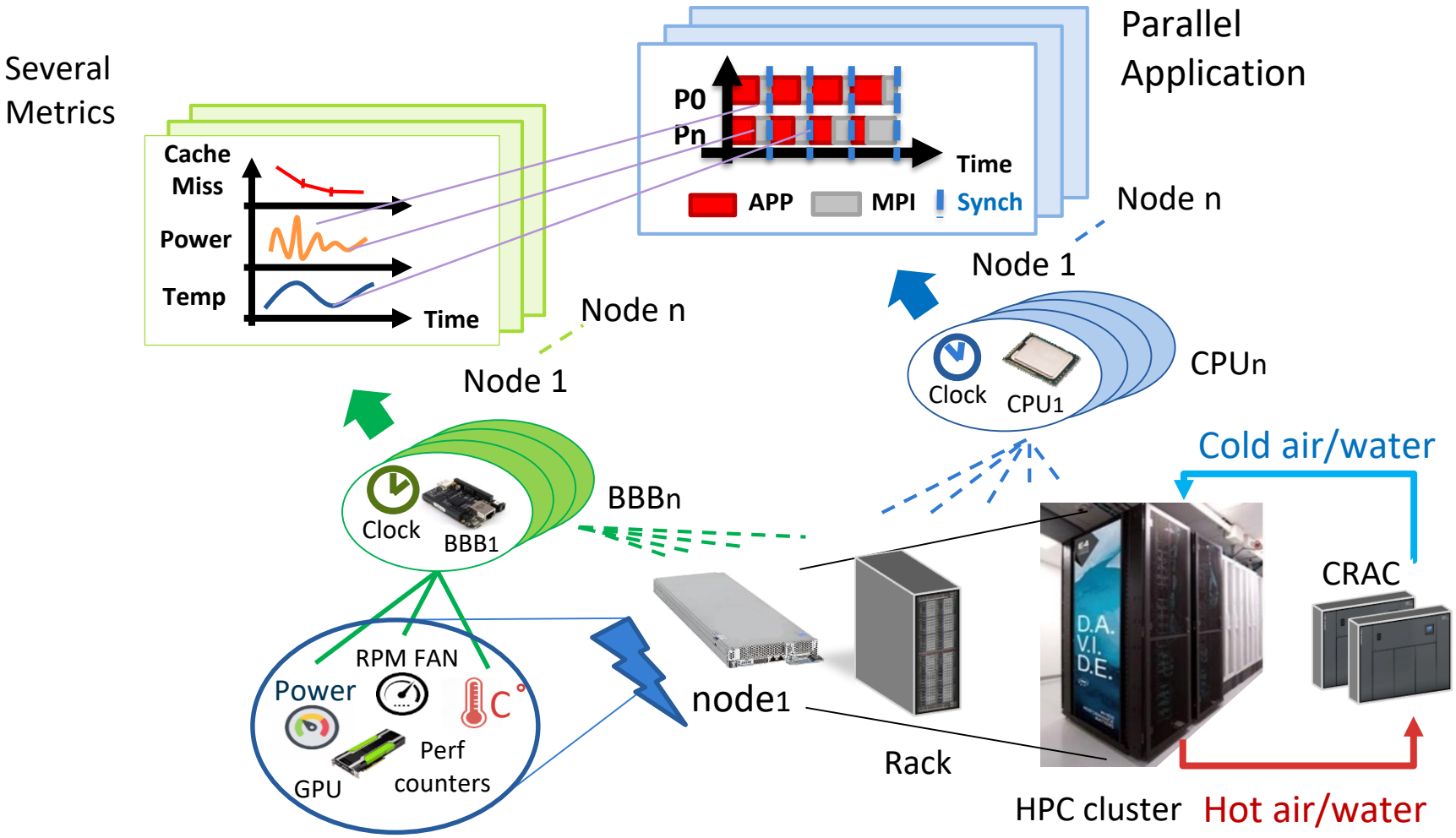
E4
COMPUTER ENGINEERING

# DiG: Fine grain correlation with jobs' activities



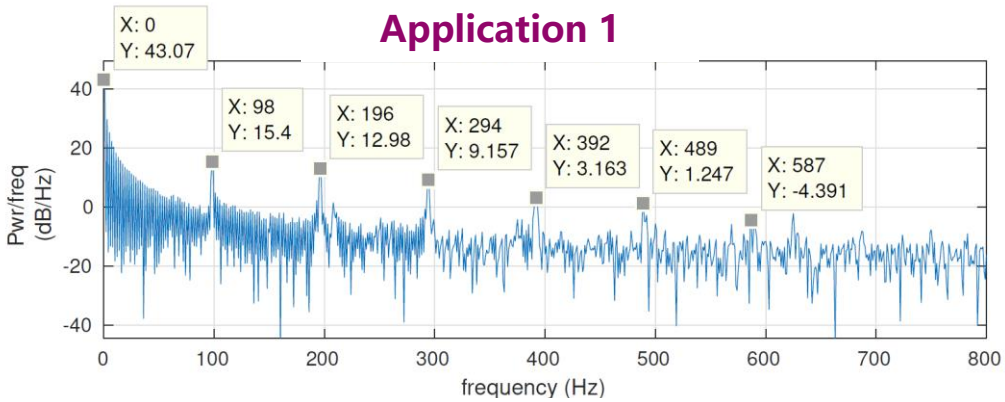



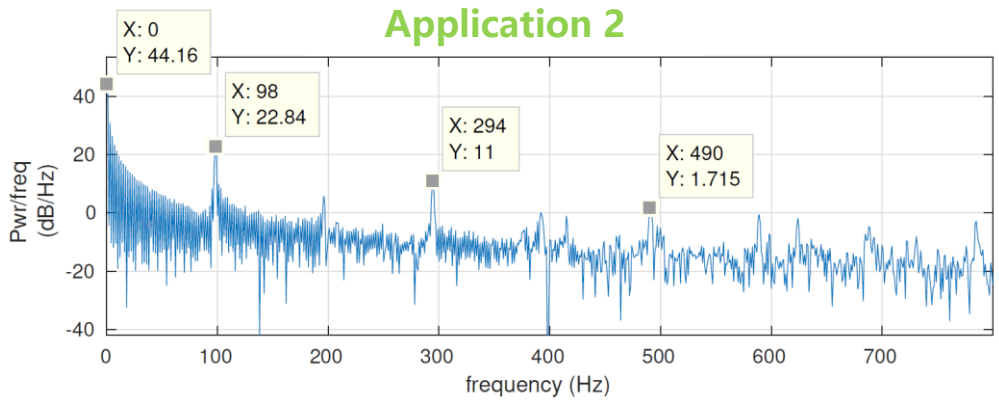
# DiG: Fine grain correlation with jobs' activities



# DiG: live FFT on the power traces

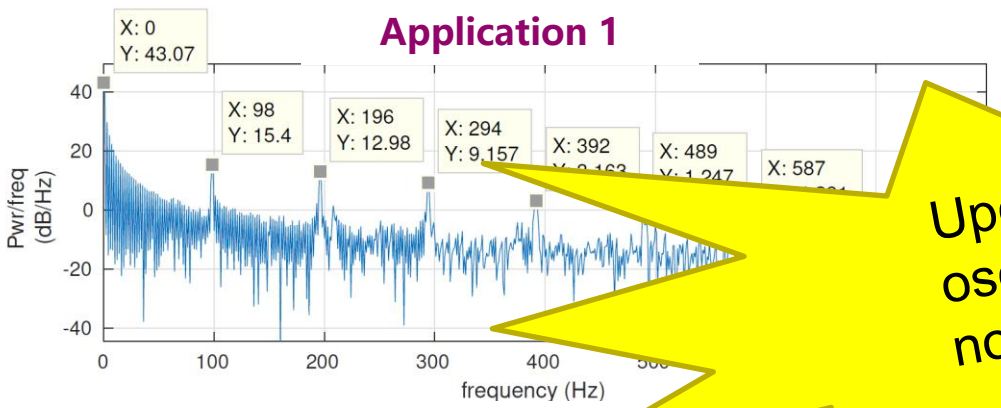
Real-time Frequency analysis on power supply and more…a live oscilloscope

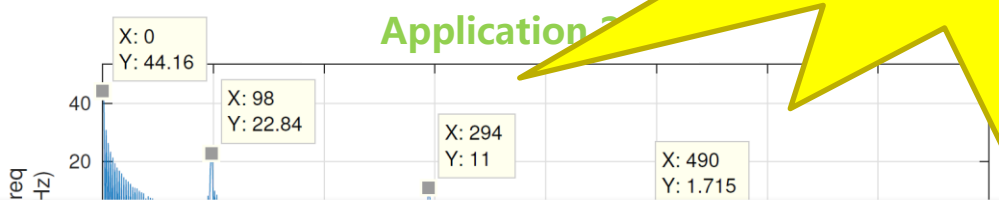


# DiG: live FFT on the power traces

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**

# Outline

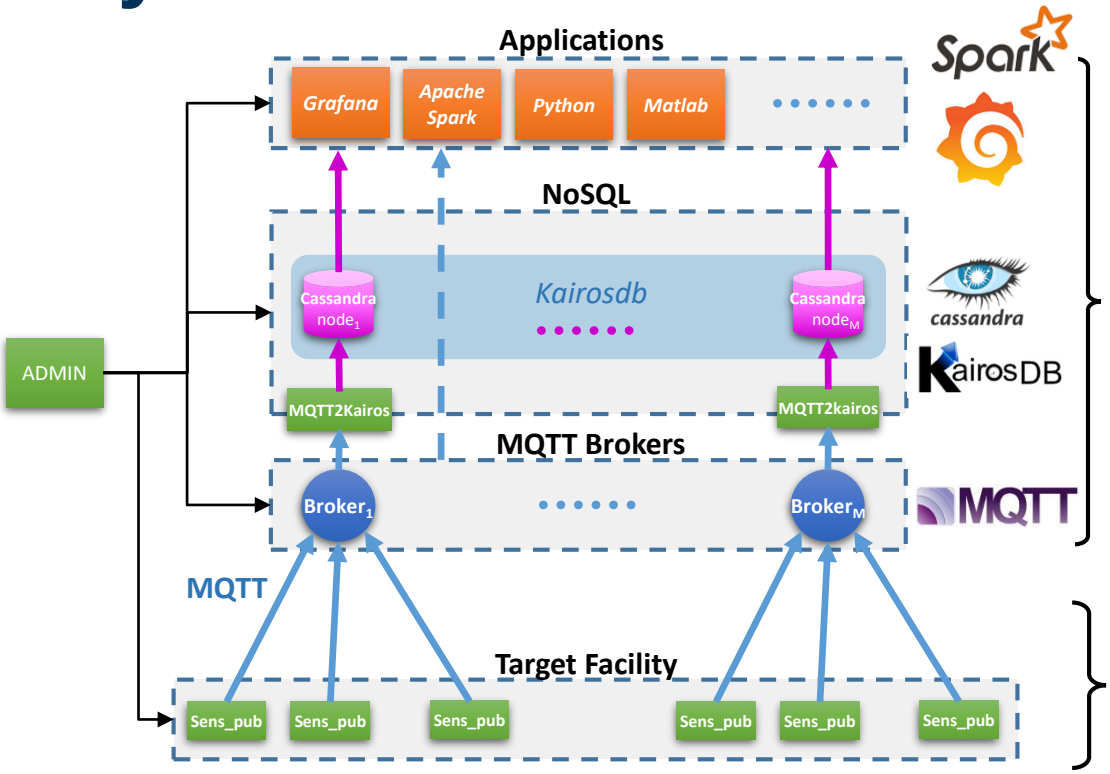
- Motivation & D.A.V.I.D.E. Overview
- DiG: High Res Out-of-Band Pow & Perf Mon
- ExaMon: Scalable Data Collection and Analytics
- Monitored Metrics and Data Visualization
- Smart Job Scheduler – SLURM Plugin

# ExaMon: Scalable Data Collection and Analytics

- Aggregates job's power and performance in real-time and at a fine granularity for Big Data analysis
- Based on opensource SW
- Store, Process, Visualize and Analyze Monitored Data
  - Historical Buffer (i.e. 2 Weeks)
  - Perpetual per Job
- Web frontend

EXAMON

# ExaMon: Scalable Data Collection and Analytics




## Front-end

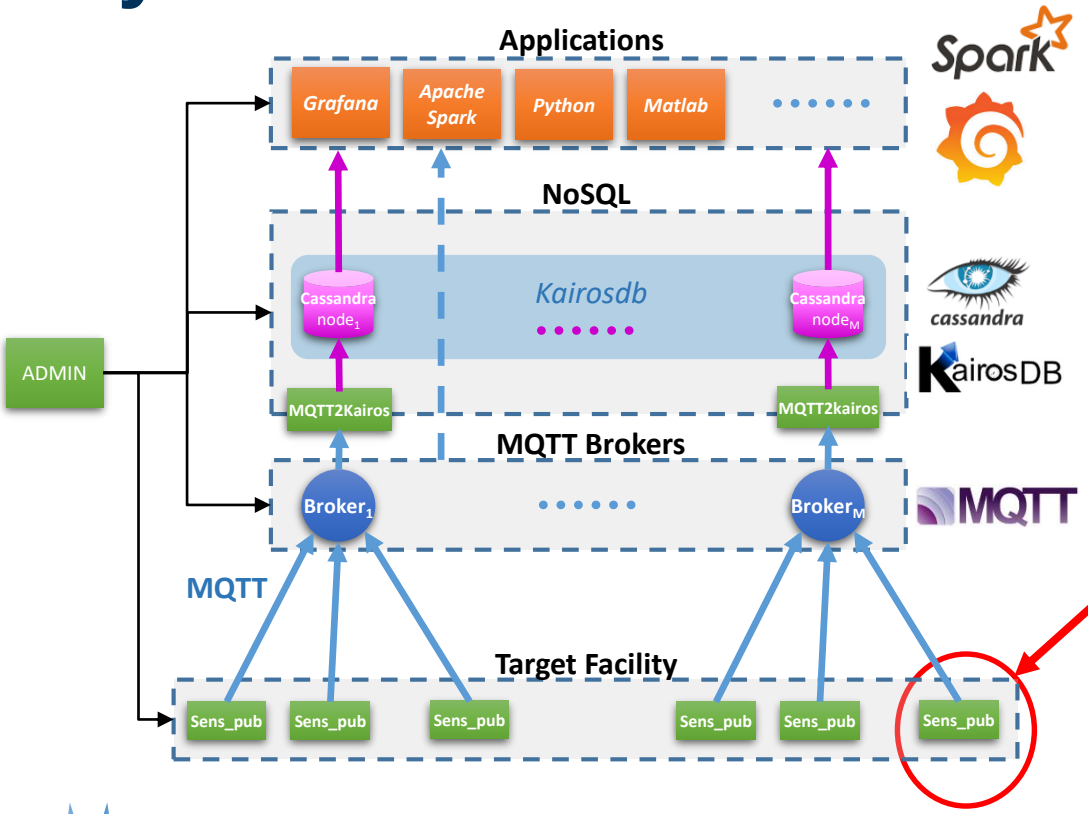
- Host: davidefe01
- **Docker containers**
- **~45KS/s**

## Back-end

- **MQTT**–enabled monitoring agents (e.g. Dig)

EXAMON

# ExaMon: Scalable Data Collection and Analytics

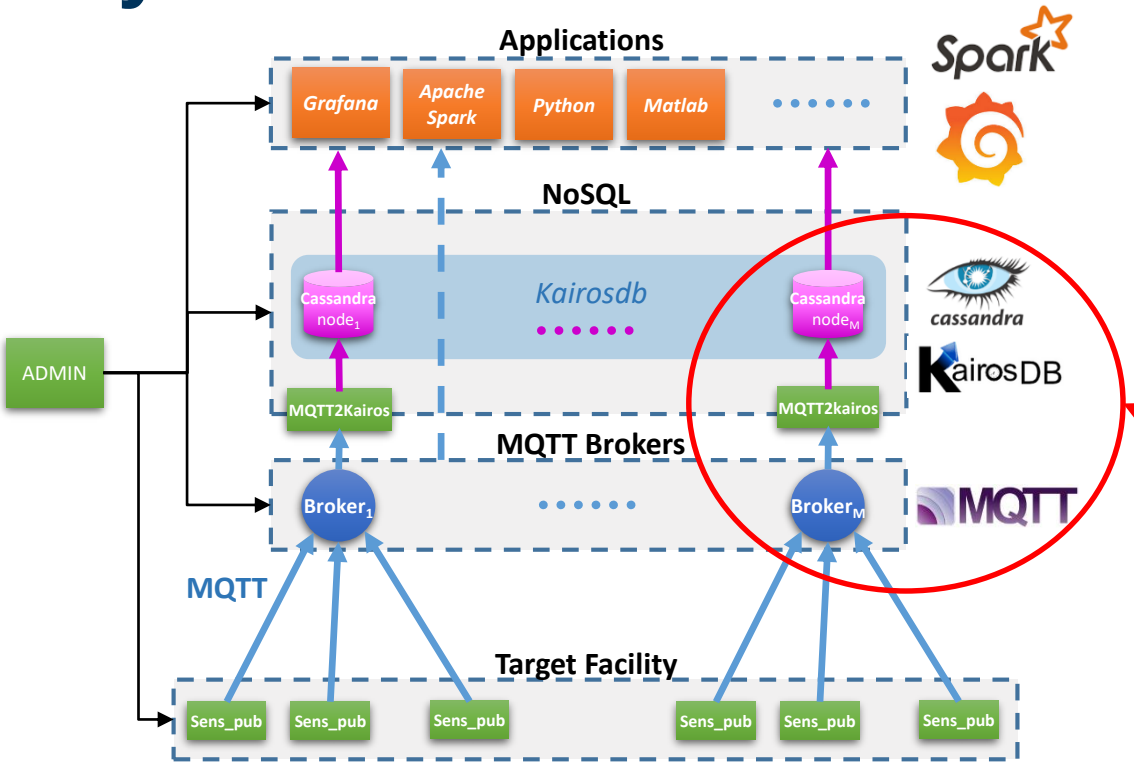



**Monitoring Agents:**

- Send Power and Performance measurements to the FE

EXAMON

# ExaMon: Scalable Data Collection and Analytics



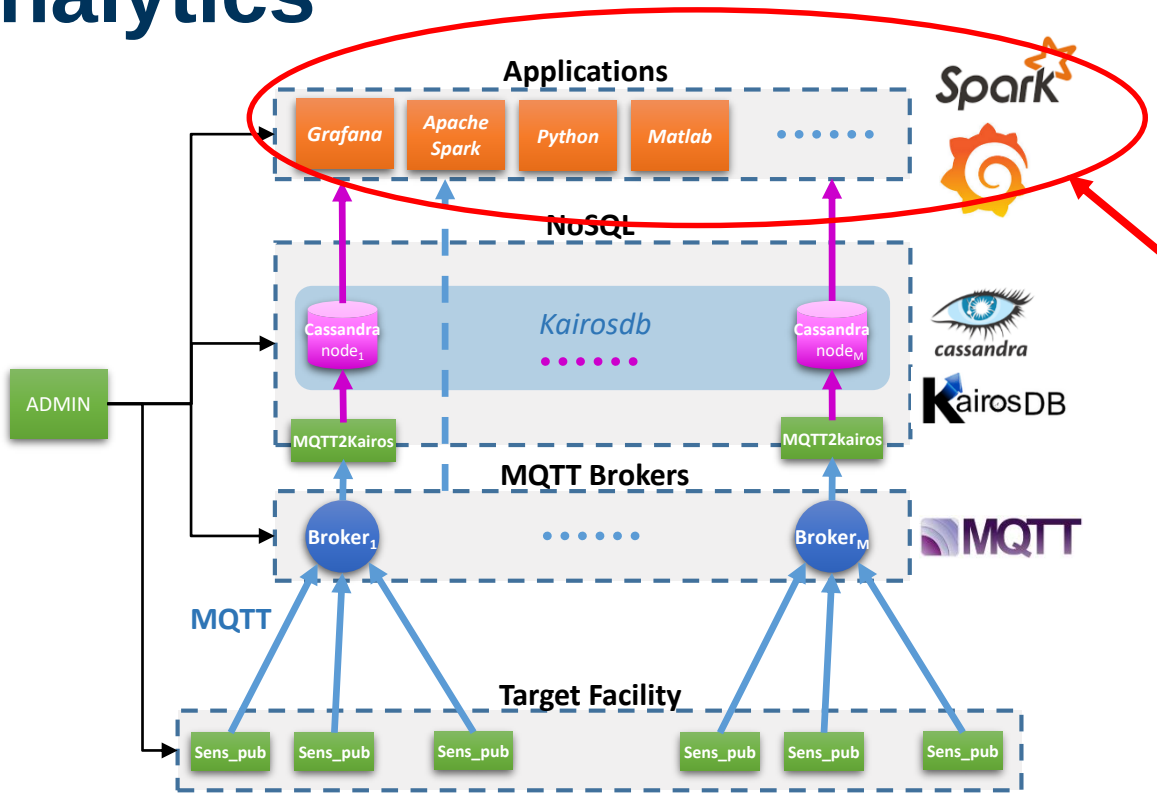

**Broker:**
- Forward data to the listeners (e.g. kairosDB)

**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

# ExaMon: Scalable Data Collection and Analytics

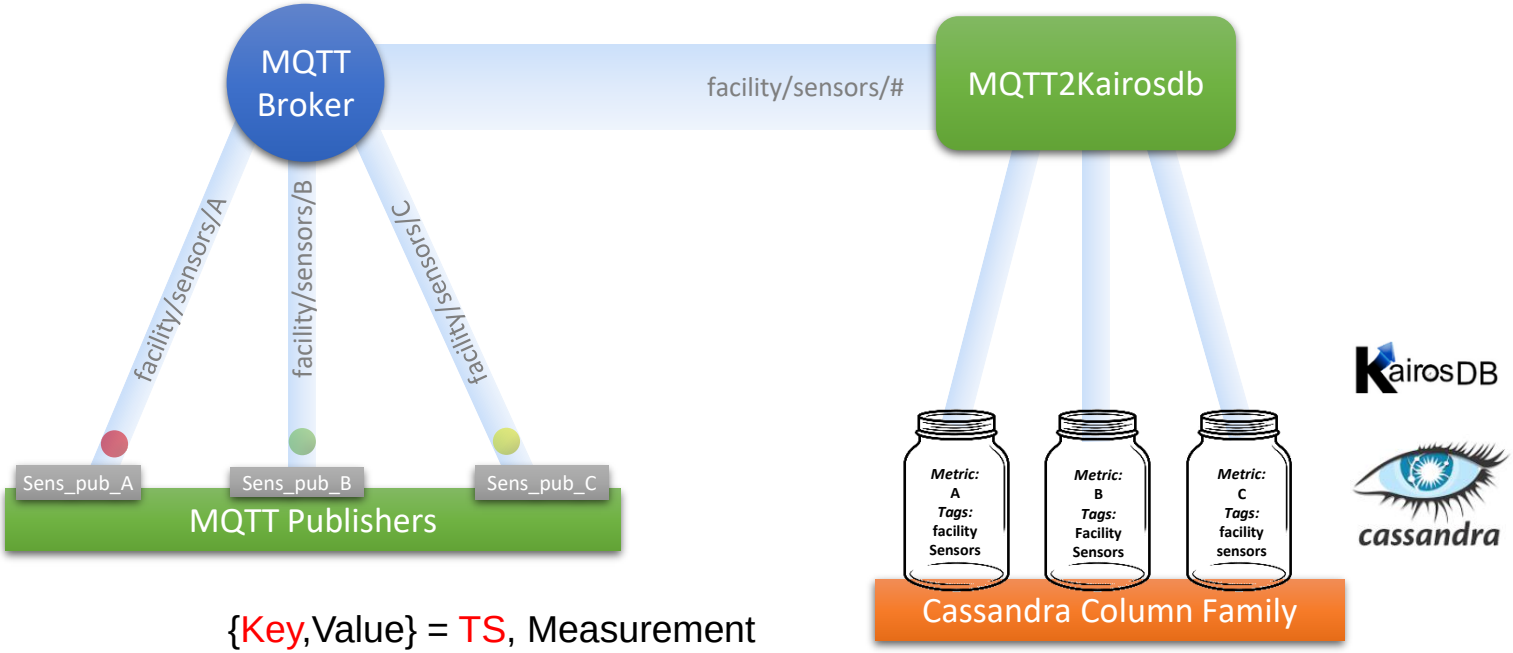



**Application layer:**

- Grafana, Apache Spark, etc …
- Aggregate metrics for Data Visualization, ML Analysis, Post Processing, etc …

EXAMON

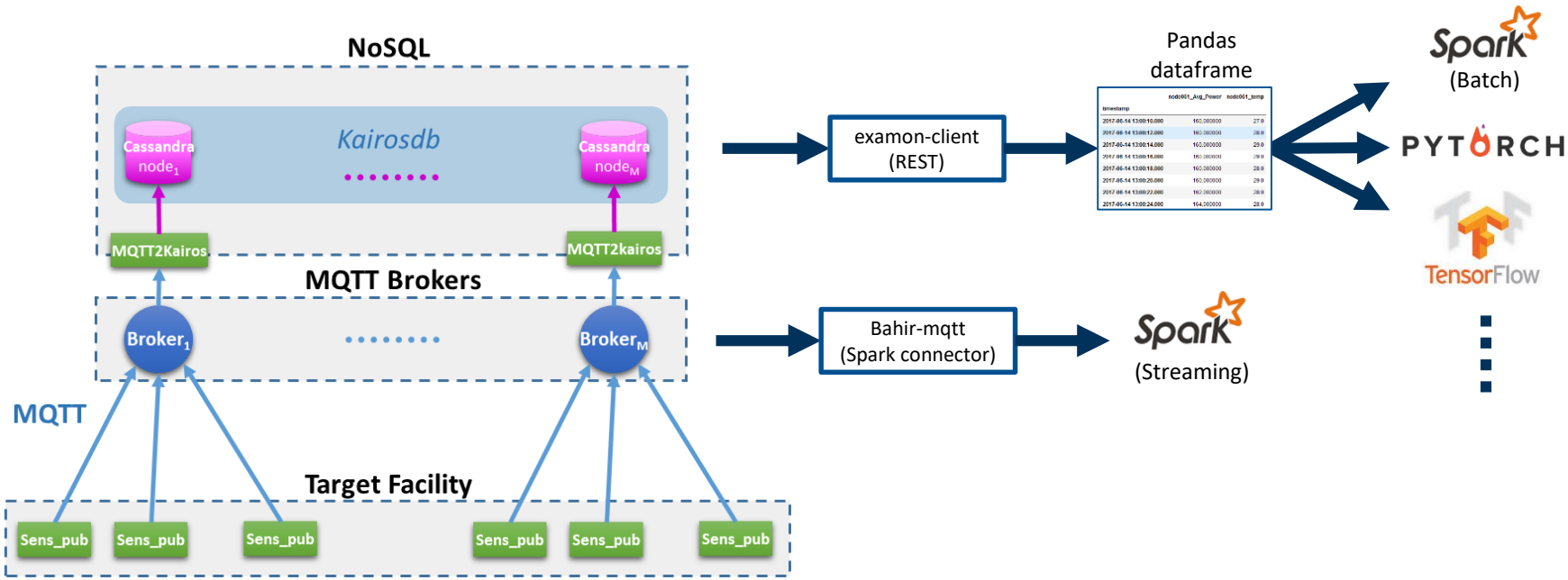
# ExaMon: Scalable Data Collection and Analytics




{Key,Value} = TS, Measurement
Topic = /davide/node1/Metric

EXAMON

# ExaMon: Batch & Streaming

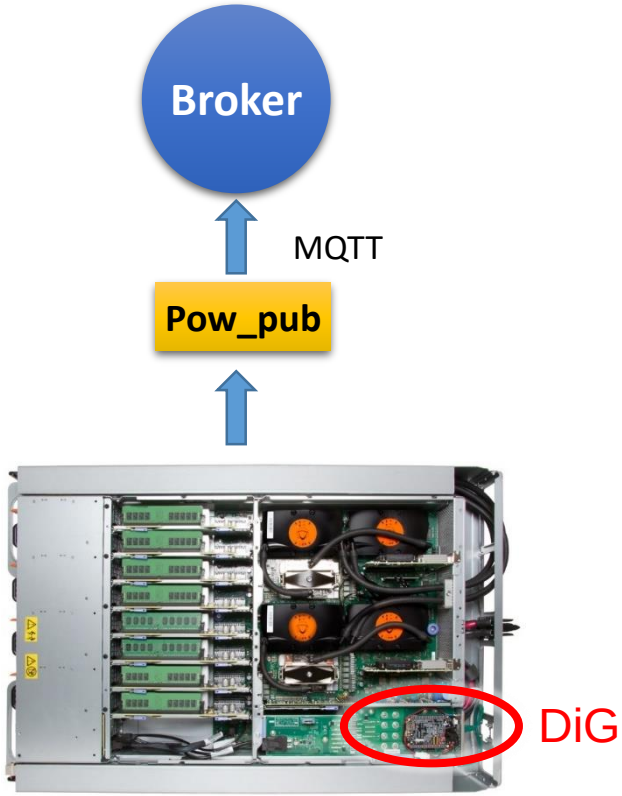



EXAMON

# Outline

- Motivation & D.A.V.I.D.E. Overview
- DiG: High Res Out-of-Band Pow & Perf Mon
- ExaMon: Scalable Data Collection and Analytics
- Monitored Metrics and Data Visualization
- Smart Job Scheduler – SLURM Plugin

# D.A.V.I.D.E. – Fine Grain Power Metric

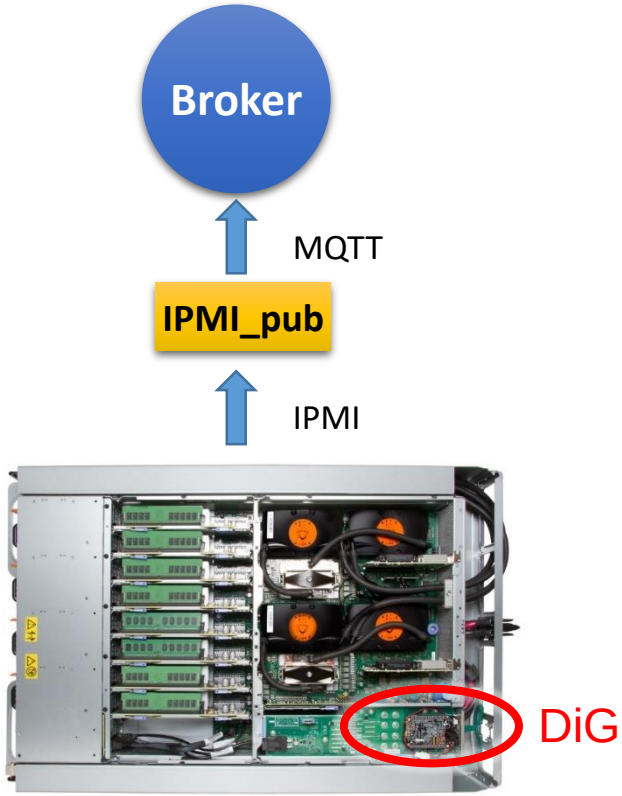



| Metric Name | Ts | Description | Unit |
|---|---|---|---|
| **power** | 1s,1ms | Power consumption at node power plug | W |

EXAMON

Power: ~45 kS/s

# D.A.V.I.D.E. – IPMI Metrics

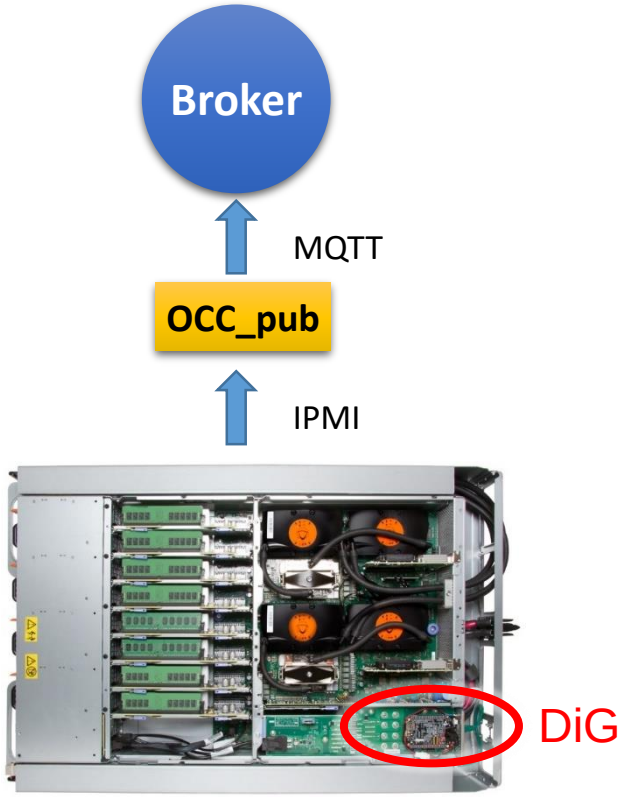



| Metric Name | Description | Unit |
|---|---|---|
| Ambient_Temp | Node ambient temperature | °C |
| CPU_Core_Temp_1,...,CPU_Core_Temp_24 | Core temperature | °C |
| CPU_Diode_1, CPU_Diode_2 | Package temperature (Diode) | °C |
| CPU1_Temp, CPU2_Temp | Package temperature | °C |
| DIMM1_Temp,...,DIMM32_Temp | DIMMs temperature | °C |
| GPU_Temp_1,...,GPU_Temp_4 | GPU temperature | °C |
| Mem_Buf_Temp_1,...,Mem_Buf_Temp_8 | Memory temperature (Centaur) | °C |
| CPU_VDD_Curr | CPU current | A |
| Fan_1,...,Fan_4 | Fan speed | RPM |
| CPU_VDD_Volt | CPU Voltage | V |
| Fan_Power | Fan power | W |
| GPU_Power | GPU power | W |
| Mem_Cache_Power | Memory power (Centaur) | W |
| Mem_Proc0_Pwr, Mem_Proc1_Pwr | DIMMs power | W |
| PCIE_Proc0_Pwr, PCIE_Proc1_Power | PCIExpress power | W |
| Proc0_Power, Proc1_Power | CPU Power | W |
| System_Power | Node total power | W |

EXAMON

IPMI: 89 metrics per node every 5s

# D.A.V.I.D.E. – OCC Metrics (1/2)

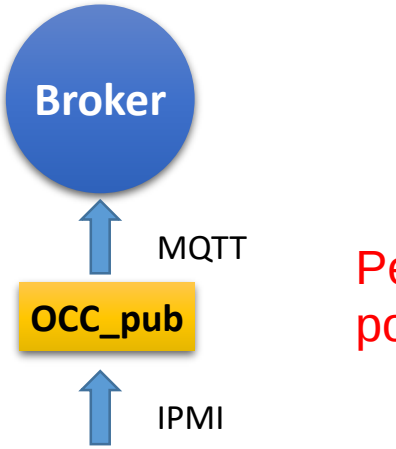



| Metric Name | Occ | Cmp | Id | Description | Unit |
|---|---|---|---|---|---|
| CUR_VDD0 | 1,2 | PROC | 0 | Current consumption at Processor's Vdd Regulator Output | A |
| TEMP_P0 | 1,2 | PROC | 0 | Vector sensor that is the average of all core temperatures for this Processor | C |
| TEMP_P0PEAK | 1,2 | PROC | 0 | Vector sensor that is the peak of all core temperatures for this Processor | C |
| TEMP_P0 | 1,2 | CORE | 0,..,11 | Average temperature of DTS sensors for Processor's Core <id>: 100% corresponds to nominal frequency | C |
| NOTFIN_P0 | 1,2 | CORE | 0,..,11 | Not Finished (stall) cycles counter for core <id> on this Processor | cyc |
| NOTBZE_P0 | 1,2 | CORE | 0,..,11 | Not Busy (stall) cycles counter for core <id> on this Processor | cyc |
| M4RD_MEM | 1,2 | MEM | 0,..,11 | Memory cached (L4) read requests per sec for Processor's MC <occ>, Centaur <id> (MBA01/MBA23) | GBs |
| M4WR_MEM | 1,2 | MEM | 0,..,11 | Memory cached (L4) write requests per sec for Processor's MC <occ>, Centaur <id> (MBA01/MBA23) | GBs |
| CMBW_P0 | 1,2 | CORE | 0,..,11 | Average Memory Band width for core <id> on this processor | GBs |
| MRD_P0 | 1,2 | MEM | 0,1,4,5 | Memory read requests per sec for Processor's Memory Controller <id> | GBs |
| MWR_P0 | 1,2 | MEM | 0,1,4,5 | Memory write requests per sec for Processor's Memory Controller <id> | GBs |
| FREQ_P0 | 1,2 | PROC | 0 | Average of all core frequencies for Processor | MHz |
| FREQA_P0 | 1,2 | CORE | 0,..,11 | Average/actual frequency for this processor, Core <id> based on OCA data | MHz |
| FREQ_P0 | 1,2 | CORE | 0,..,11 | Requested frequency from OCC for Core <id> | MHz |
| IPS_P0 | 1 | PROC | 0 | Vector sensor that takes the average of all the cores in Processor | MIP |
| IPS_P0 | 1,2 | CORE | 0,..,11 | Instructions per second for core <id> on this Processor | MIP |
| VOLT_V0 | 1,2 | VRM | 0 | Voltage request for this Processor's Vdd voltage rail | mV |
| VOLT_V1 | 1,2 | VRM | 0 | Voltage request for this Processor's Vcs voltage rail | mV |

EXAMON

OCC: 242 metrics per node every 10s

# D.A.V.I.D.E. – OCC Metrics (2/2)





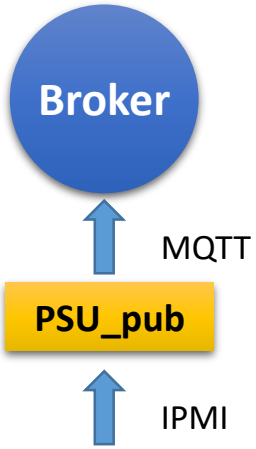
Per-component power

| Metric Name | Occ | Cmp | Id | Description | Unit |
|---|---|---|---|---|---|
| WINKCNT_P0 | 1,2 | PROC | 0 | Count the number of cores that are winkled in this processor | Num |
| SLEEPCNT_P0 | 1,2 | PROC | 0 | Count the number of cores that are sleep in this processor | Num |
| UTIL_P0 | 1,2 | CORE | 0,..,11 | Utilization of this Processor's Core <id> (where 100% = fully utilized): NOTE: per thread HW counters are combined as appropriate to give this core level utilization sensor | Per |
| UTIL_P0 | 1,2 | PROC | 0 | Average of all core utilizations for this Processor (where 100% = fully utilized) | Per |
| PWR_VDD0 | 1,2 | PROC | 0 | Power consumption for this Processor's Vdd Regulator Input (12Volt) | W |
| PWRPX_P0 | 1,2 | CORE | 0,..,11 | Power Proxy sensor for Core <id> on this processor | W |
| PWR_null | 1 | SYS | 0 | Bulk power of the system - Master only sensor | W |
| PWR_P0 | 1,2 | PROC | 0 | Power consumption for this Processor | W |
| PWR_MEM | 1,2 | PROC | 0 | Power consumption for Memory for this Processor | W |
| PWR_STORE | 1 | SYS | 0 | Power consumption of the storage subsystem (storage 12Volt rail) - Master only sensor | W |
| PWR_FAN | 1 | SYS | 0 | Power consumption of the system fans - Master only sensor | W |
| PWR_IO | 1 | SYS | 0 | Power consumption of the IO subsystem (including storage digital 5Volt or less rail) - Master only sensor | W |
| PWR_GPU | 1 | SYS | 0 | Power consumption of the GPU | W |
| PWR_VCS0 | 1,2 | PROC | 0 | Power consumption for this Processor's Vcs Regulator Input (12Volt) | W |
| PWR_APSS | 1 | SYS | 1,..,15 | Power Provided by APSS channel <id> | W |

EXAMON

OCC: 242 metrics per node every 10s

# D.A.V.I.D.E. – PSU Metrics

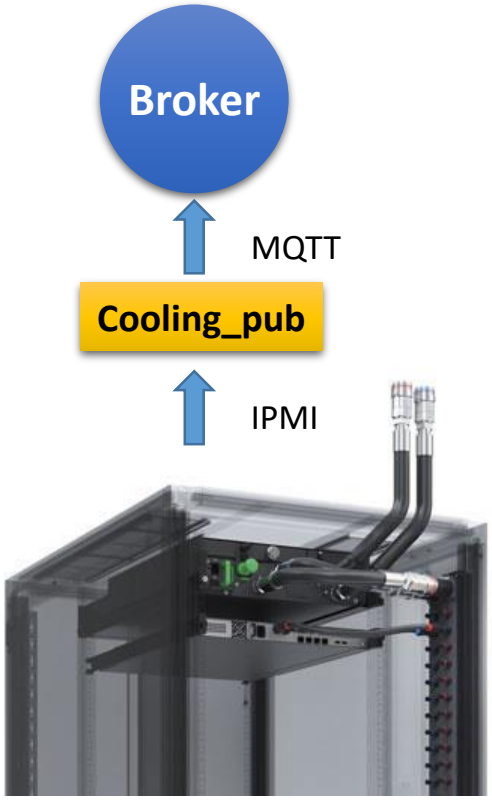



- Running in the front-end
- Full Rack info

| Metric Name | Description | Unit |
|---|---|---|
| LITEON-PSC-MIB::pscPSOutputLoad_0 | Power Supply Output load | % |
| LITEON-PSC-MIB::pscPSOutputEfficiency_0 | Power Supply Output efficiency | % |
| LITEON-PSC-MIB::pscBatteryCapacity_0 | Battery capacity | % |
| LITEON-PSC-MIB::pscBatteryTemp_0 | Battery temperature | °C |
| LITEON-PSC-MIB::pscPSPhaseRInputCurr_1 | Power Supply phase R Input current | A |
| LITEON-PSC-MIB::pscPSPhaseSInputCurr_1 | Power Supply phase S Input current | A |
| LITEON-PSC-MIB::pscPSPhaseTInputCurr_1 | Power Supply phase T Input current | A |
| LITEON-PSC-MIB::pscPSOutputCurr_0 | Power Supply Output current | A |
| LITEON-PSC-MIB::pscPSPhaseRInputFreq_1 | Power Supply phase R Input frequency | Hz |
| LITEON-PSC-MIB::pscPSPhaseSInputFreq_1 | Power Supply phase S Input frequency | Hz |
| LITEON-PSC-MIB::pscPSPhaseTInputFreq_1 | Power Supply phase T Input frequency | Hz |
| LITEON-PSC-MIB::pscPSOutputVolt_0 | Power Supply Output voltage | V |
| LITEON-PSC-MIB::pscPSPhaseRInputVolt_1 | Power Supply phase R Input voltage | V |
| LITEON-PSC-MIB::pscPSPhaseSInputVolt_1 | Power Supply phase S Input voltage | V |
| LITEON-PSC-MIB::pscPSPhaseTInputVolt_1 | Power Supply phase T Input voltage | V |
| LITEON-PSC-MIB::pscBatteryVolt_0 | Battery voltage | V |
| LITEON-PSC-MIB::pscPSOutputPower_0 | Power Supply Output power | W |
| LITEON-PSC-MIB::pscPSPhaseRInputPower_1 | Power Supply phase R Input power | W |
| LITEON-PSC-MIB::pscPSPhaseSInputPower_1 | Power Supply phase S Input power | W |
| LITEON-PSC-MIB::pscPSPhaseTInputPower_1 | Power Supply phase T Input power | W |
| LITEON-PSC-MIB::pscBatteryPower_0 | Battery power | W |

EXAMON

# D.A.V.I.D.E. – Liquid Cooling Metrics

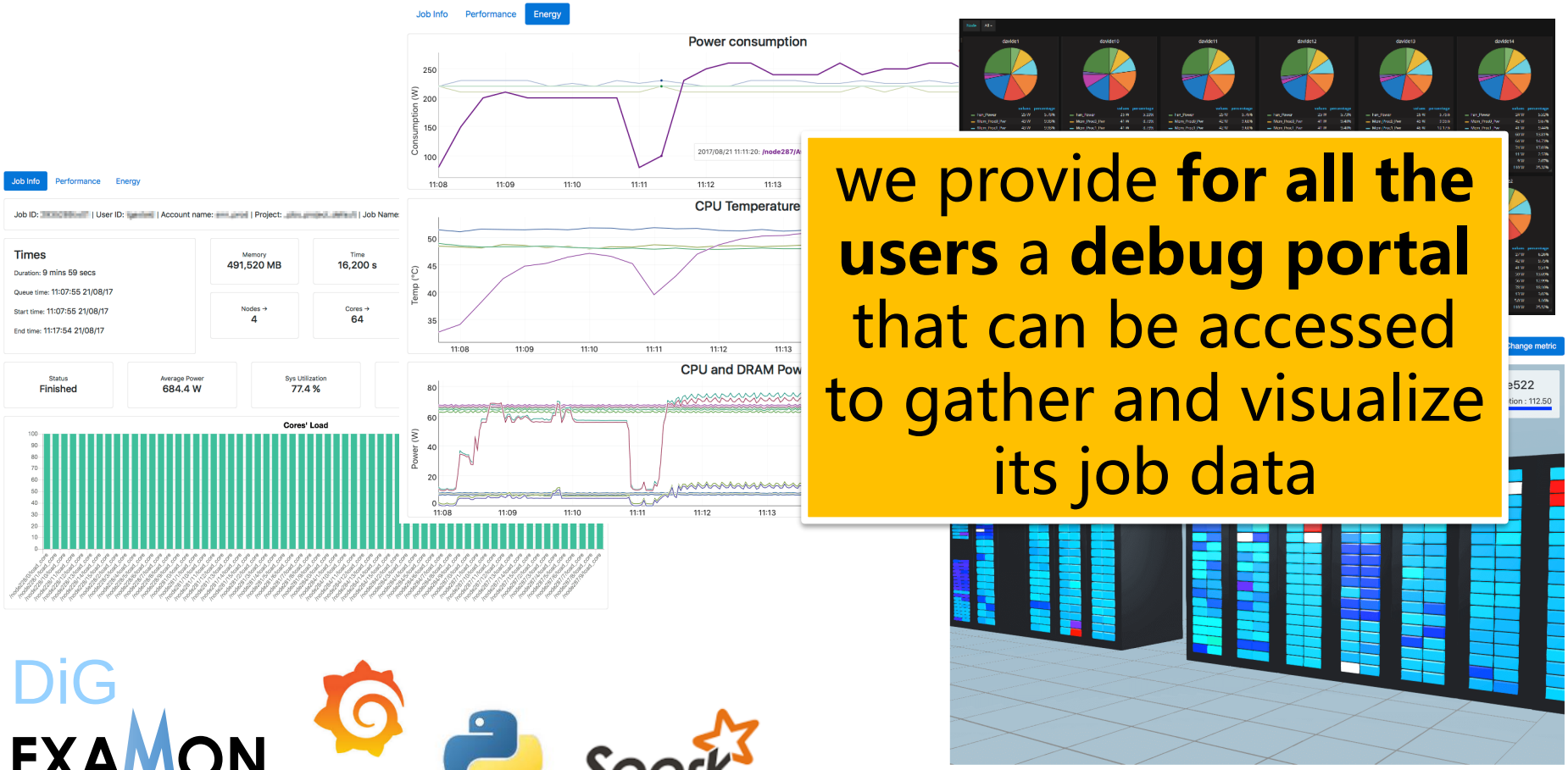



| Metric name | Description | Unit |
|---|---|---|
| ASETEK-RACKCDU-SMI-V2-MIB-V17::temperatureFacilityOut_0 | Temperature facility out | °C |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::temperatureServerOut_0 | Temperature server out | °C |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::temperatureFacilityIn_0 | Temperature facility in | °C |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::temperatureServerIn_0 | Temperature server in | °C |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::temperatureAmbient_0 | Temperature ambient | °C |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::pressureFacility_0 | Pressure facility | mbar |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::pressureServer_0 | Pressure server | mbar |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::flowFacility_0 | Flow facility | ml/s |
| ASETEK-RACKCDU-SMI-V2-MIB-V17::heatload_0 | Heat load | W |

- Running in the front-end
- Per Rack Cooling info

EXAMON

# Real-Time Monitoring, Processing, Visualization

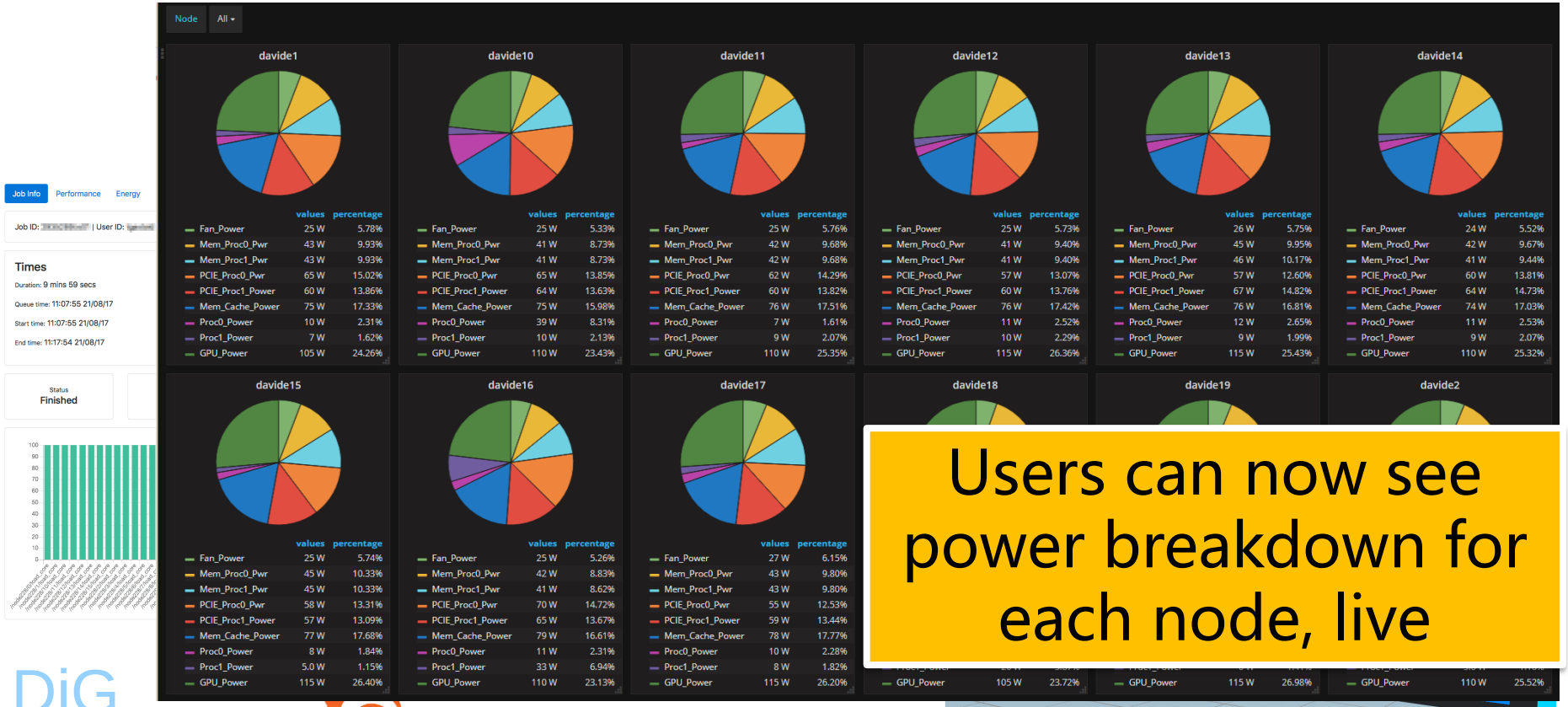



we provide **for all the users** a **debug portal** that can be accessed to gather and visualize its job data

DiG EXAMON

Spark

# Real-Time Monitoring, Processing, Visualization




Users can now see power breakdown for each node, live

DiG EXAMON

Spark

# Outline

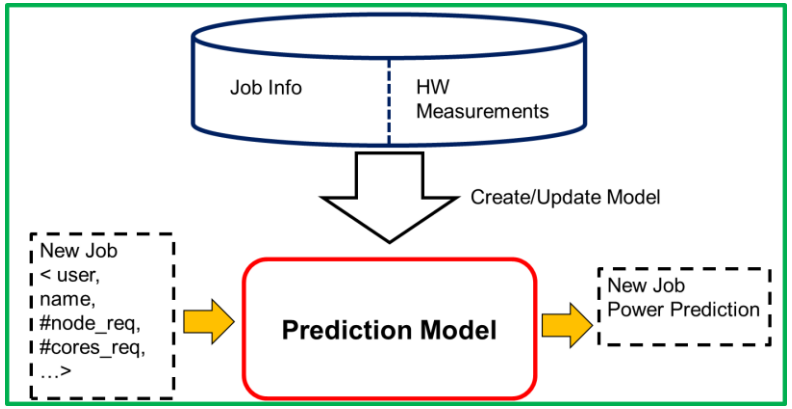
- Motivation & D.A.V.I.D.E. Overview
- DiG: High Res Out-of-Band Pow & Perf Mon
- ExaMon: Scalable Data Collection and Analytics
- Monitored Metrics and Data Visualization
- Smart Job Scheduler – SLURM Plugin

# Smart Job Scheduler – SLURM Plugin

- Monitors Job submission and scheduling
- Estimates Job Power Consumption
- Controls Power Consumption under system level power capping

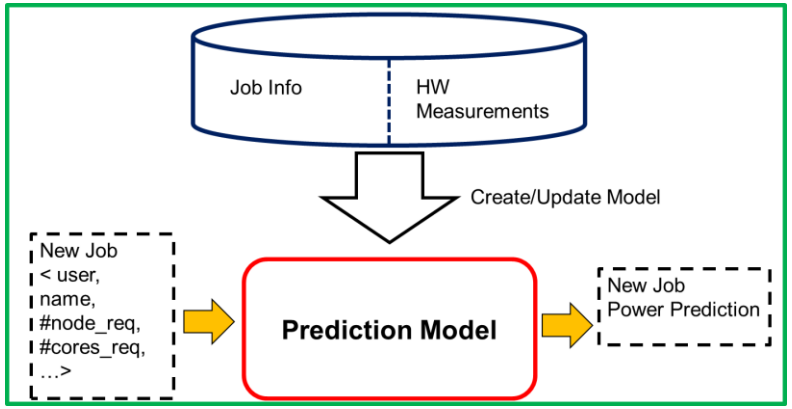
# Smart Job Scheduler – SLURM Plugin

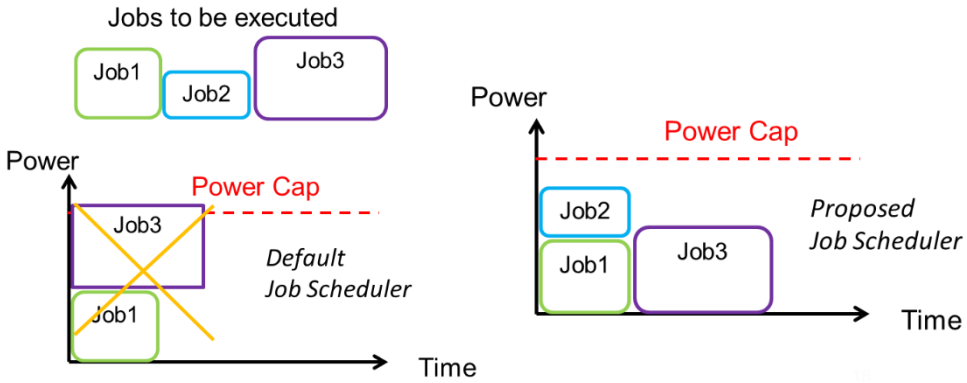
1. ML models to predict power consumption of HPC applications
2. SLURM Custom Extensions to schedule jobs based on their power
3. Interaction with power management
   - Frequency scaling/RAPL-like mechanism



# Smart Job Scheduler – SLURM Plugin

- **Different** from previous power capping approaches
- **We do not perturb the duration** of the application, thus it doesn't change the way users pay the resources



# Conclusion

- With D.A.V.I.D.E. and its monitoring infrastructure we provide real-time analysis on computing nodes
- high resolution power and performance measurements up to ms scale
- upcoming FFT live analysis
- Data Storage, Processing and Visualization, along with Big Data Analysis support
- Power Aware Job Scheduling

ETH zürich

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

# Thanks for your interest.

Co-funded by
*EU FP7 ERC Project Multitherman (No 291125)*

Multitherman

This project has received funding from the European Union's *Horizon 2020* research and innovation programme under grant agreement No 671623 (*ANTAREX*)

Contact:

- Antonio Libri, a.libri@iis.ee.ethz.ch

Unibo/ETH:
A. Bartolini, F. Beneventi,
A. Borghesi, and L. Benini