Instituto de Física de Cantabria | CSIC Consejo Superior de Investigaciones Científicas | UC Universidad de Cantabria

GreenDIGIT | WP4/WP5/WP6 Design Workshop
Amsterdam Science Park
January 23rd/24th 2025

GreenDIGIT WP4: Metrics & Infrastructure

Environmental metrics collection experiences and possibilities at sites

Jaime Iglesias Blanco

iglesias@ifca.unican.es

Advanced Computing and e-Science Group
https://advancedcomputing.ifca.es

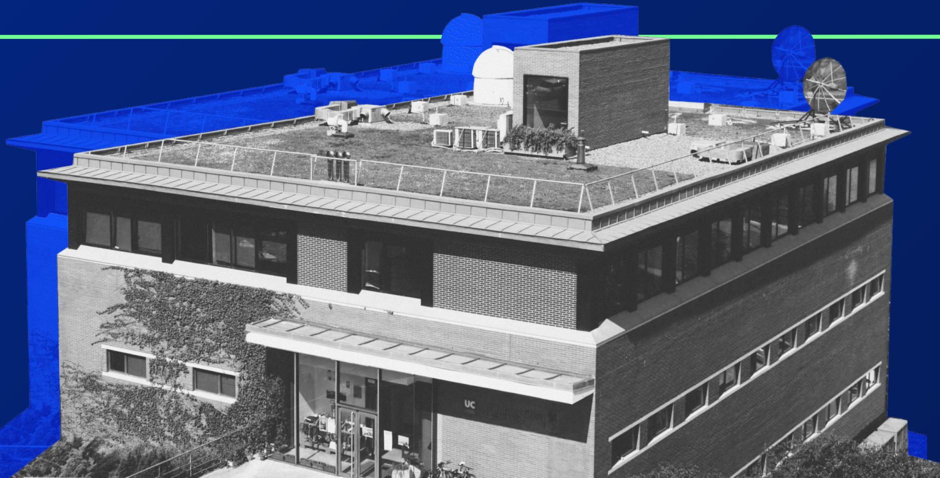


Table of contents:

1. Introduction
2. Power consumption metrics
3. Gathering power consumption measurements
   3.1. At technical room level
   3.2. At machine level
   3.3. At virtual machine level
   3.4. Other IT systems metrics
4. Electricity suppliers metrics
5. Final picture

Introduction

IFCA Site

IFCA Site Description:

Dedicated to advanced computing R&D as well as offering HPC, Grid and Cloud computing services to scientific researchers.

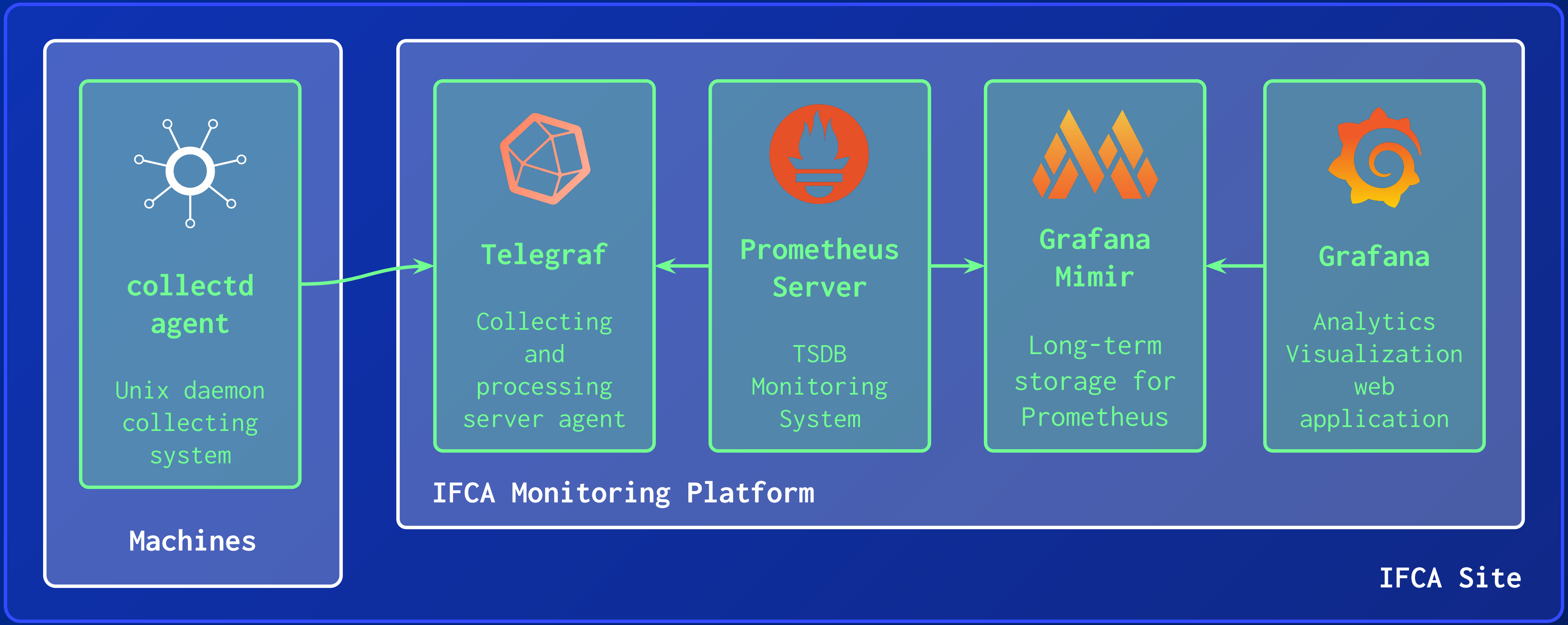
Relevant IT Equipment:

➔ **12 racks** (42U each):
- 300 servers: From many manufacturers | From 2009 to 2024 (All being monitored)
  - Currently, 491 virtual machines are hosted (All being monitored)
- 20 network switches, 7 drives enclosures (2.5PB) and 1 tape library (Static)

Relevant Non-IT Equipment:

➔ 3 Coolers (2/3 being monitored, soon the last one)
➔ 4 Fans for free-cooling (heat extraction) (All being monitored)
➔ 2 Uninterruptible power supply units (UPS) - 75 kW each (Not monitored yet)
➔ 2 Electrical power panels (Only 1 being monitored, soon the other)
➔ 1 Diesel power generator unit (Not used except in case of power failure)

New IFCA Site Monitoring Platform:



IFCA Site Roadmap:

1. Define **basic metrics** (or those suggested in the WP3 survey).
2. Search for metrics most reliable **data sources**.
3. **Tooling up** the facilities and equipment for data collection.
4. Integrate the **new tools** into the existing platform.
5. Wait and collect the **measurements**.
6. **Compute** the metrics.
7. Verify if the metric values are **realistic** (fair).
8. **Publish** relevant metrics into GOCDB.

Power consumption metrics

Now available at IFCA Site

Power consumption metrics:

Metrics as defined on
ISO/IEC 30134-2 / EN 50600-4-2

Measured as defined on
CEN/CENELEC EN 50600-4-2

TUE = PUE x ITUE

- → Total Site Power Consumption (Watts)
- → Site carbon footprint (gCO2eq)
- → Power Consumption per machine (Watts)
  - In some cases it is possible by PSU, CPU, GPU, Memory…
- → Power Consumption per virtual machine (Watts)
- → Power Usage Effectiveness (PUE): *Total Facility Energy (Power) / IT Equipment Energy (Power)*
  - Another way: *Total / Total - Non IT* (Simpler to measure on our site)
- → IT Usage Effectiveness (ITUE): *IT Equipment Energy (Power) / Energy (Power) into Compute Components*
- → Total Usage Effectiveness (TUE): *Total Facility Energy (Power) / Energy (Power) into Compute Components*

… and more combinations between the different measurements (ratios and partial carbon footprints)

Source: Patterson et al. (2011) & https://knowledge.sdialliance.org/archive/data-center-metrics

Relevance of these metrics

- **ITUE** metric on compute equipment represents the energy efficiency of the servers:
  - Therefore it depends on the manufacturer.
    - If it can be obtained per machine, it can be used to determine the most efficient server in a scheduling process.

- **TUE** metric can be useful to estimate a realistic carbon footprint of user workloads:

1. User can **monitor** VMs or jobs power consumption using a variety tools.
   a. These normally **only measures the CPU/GPU** power consumption (Compute Components).
2. Obviously there is more necessary equipment consuming power to run these user jobs.
3. We can **scale the user workload consumption value** to a real average value by multiplying it by the **TUE** metric of the site where it is running.
4. Afterwards, we can apply them the **CO2 emission factor** (by country) to get the footprint.

Gathering power consumption measurements

IFCA Site Roadmap

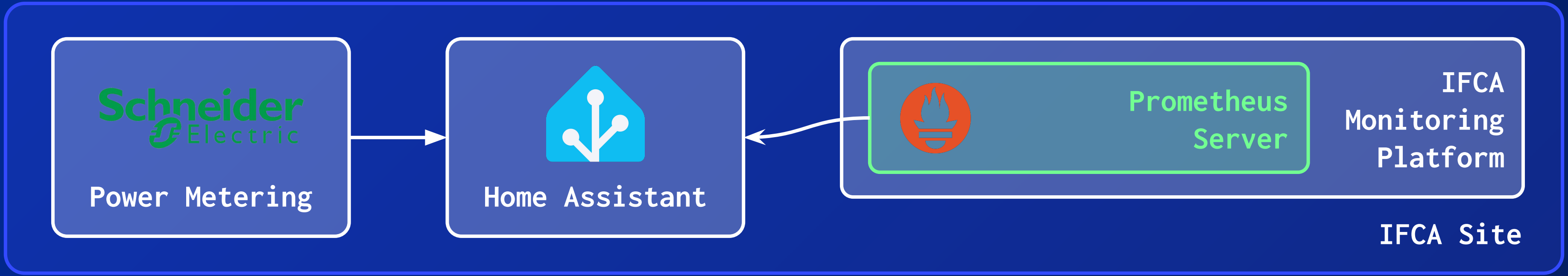
At technical room level:

→ We have placed **Schneider Electric Power Sensors** for power metering:

- On 1/2 electrical panels (The other one soon, after room renovation)
- 17 sensors distributed between some racks and coolers + fans.

→ Sensors provide a selection of electrical parameters (Amps, Volts, Watts, W/h)

→ These measurements can be exported by network, through the sensor hub.



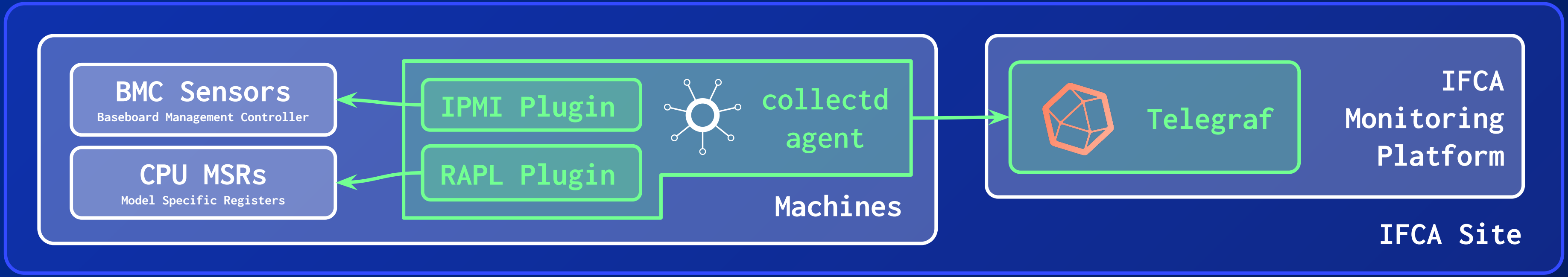
→ Using a **custom HA integration**, we can connect the sensors hub to Home Assistant

At technical room level:

Live demo at:
https://monitor.ifca.es/grafana/

At machine level:

- We select **2 data sources**, as some machines have only one of them available.
  - On some machines the measurements are more detailed on certain components and on others more high-level.
- Each machine has installed collectd agent to scrape system metrics.
- We add 2 more plug-ins (as is plug-in based) to get:
  - BMC sensors measurements, available at motherboard, via **IPMI.**
  - CPU MSRs measurements via **RAPL** Interface.



At machine level:

Live demo at:
https://monitor.ifca.es/grafana/

Nº Machines	Power Consumption	Nº Measures		Power Consumption by Measure	
Count 216	Max 75.687 kW	IPMI 196	RAPL 162	IPMI 73.535 kW	RAPL 29.354 kW

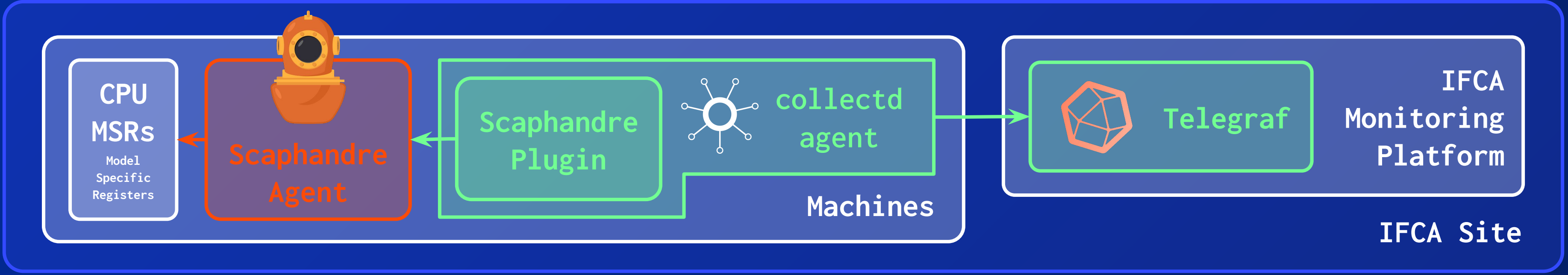
Machine	IPMI	RAPL
aitken	490 W	
aitken02	510 W	
bacula	110 W	
cephosd11	535 W	92.8 W
cephosd12	495 W	84.6 W
cephosd13	455 W	45.6 W
cephosd21	365 W	115 W
cephosd22	385 W	111 W
cephosd23	375 W	117 W
cmsfj01	80 W	43.3 W
cmsfj03	76 W	42.8 W
cmsfj04	160 W	128 W
cmsfj05	232 W	177 W
cmsfj08	156 W	114 W
cmsfj09	188 W	139 W
cmsfj12	264 W	192 W
cmsfj13	80 W	49.8 W
cmsfj14	328 W	233 W
cmsfj16	76 W	41.5 W
cmsfj17	84 W	45.9 W
cmsgiga01	422 W	
cmsgiga02	426 W	
cmsgiga03	422 W	
cmsgiga04	422 W	
cmsgiga05	260 W	
cmsgiga06	268 W	
cmsgiga07	264 W	
cmsgiga08	264 W	
cmsgpu01	904 W	131 W
cmsln01	370 W	320 W
cmsln02	370 W	319 W
cmsln03	370 W	320 W
cmsln04	60 W	35.2 W
cmsln05	370 W	321 W
cmsln06	100 W	67.5 W
cmsln07	60 W	36.5 W
cmsln08	380 W	320 W
cmsln09	370 W	319 W
cmsln10	90 W	67.7 W
cmsln11	90 W	60.4 W
cmsln13	60 W	32.5 W
cmsln14	60 W	39.2 W
cmsln15	60 W	35.7 W
cmsln16		320 W
cmsln17	90 W	61.7 W
cmsln18	380 W	320 W
cmsln19	100 W	34.4 W
cmsln20	380 W	321 W
db01	102 W	46.7 W
db02	102 W	58.6 W
db03	102 W	43.7 W
dss01		294 W
dss02		300 W
fcloud01	100 W	34.9 W
fcloud02	370 W	273 W
fcloud03		57.8 W
fcloud04	105 W	40.3 W
fcloud05	90 W	31.4 W
geoffrey	120 W	
gpumad01		70.0 W
gpumad02	295 W	71.5 W
gpumad03		67.4 W
gpumad04	265 W	56.5 W
gpumad05	280 W	66.7 W
gpumad06	335 W	108 W
gpumad07	265 W	64.2 W
gpumad08	270 W	66.2 W
gpumad09	280 W	66.7 W
gpumad10	285 W	70.2 W
gpumad11	265 W	65.0 W
gpumad12	275 W	74.0 W
gpumad13	290 W	69.9 W

At machine level:

Live demo at:
https://monitor.ifca.es/grafana/

At virtual machine level:

➔ The Scaphandre agent installed on physical machines provides the power consumption of the virtual machines (KVM-QEMU) hosted on it.



➔ We add into collectd agent a new metric set with all hosted VMs (KVM-QEMU) power consumption measurements in a physical machine.

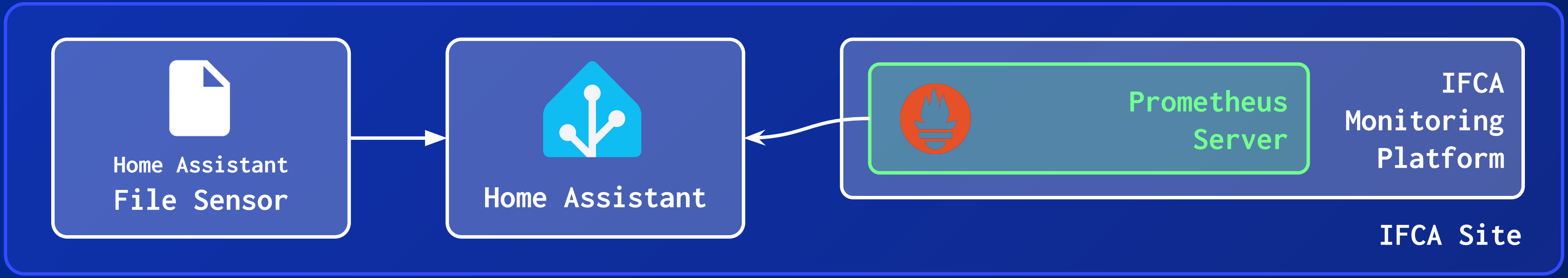
- For now, the power consumption of the VM is not shared with the VM itself.
  - See *IFCA + AI4EOSC Power Consumption & Environmental Impact Evaluation* for more details.

At virtual machine level:

Live demo at:
https://monitor.ifca.es/grafana/

Other IT systems metrics:

- Some IT equipment could not be measure from the system itself.
  - These are: **Network Switches and Attached Storage Drives Enclosures**.
    - It could be done through smart power meters or monitoreable PDUs.
      - This is not possible yet
- We add the power consumption statically from product datasheets.



- Using Home Assistant File Sensor we introduce a static value into the platform.

Following the formulas…

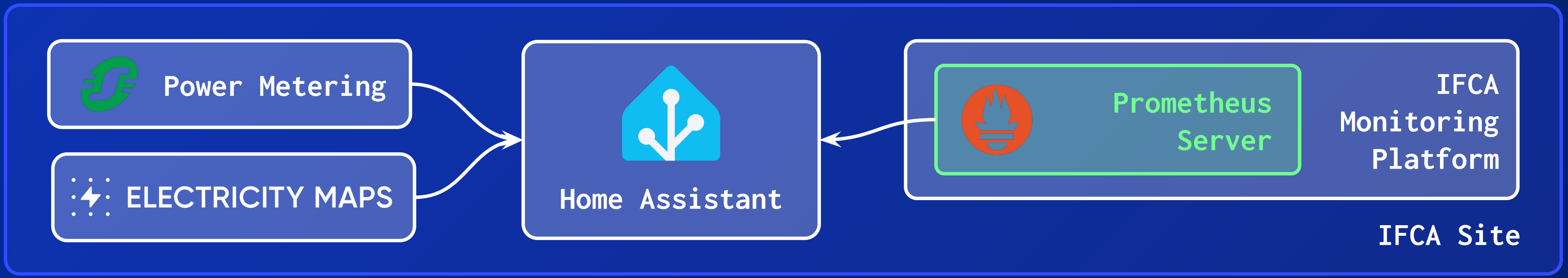
Live demo at:
https://monitor.ifca.es/grafana/

Electricity suppliers metrics

Carbon intensity

Electricity Maps

→ We connect **Electricity Maps API** via HomeAssistant Plug-in.

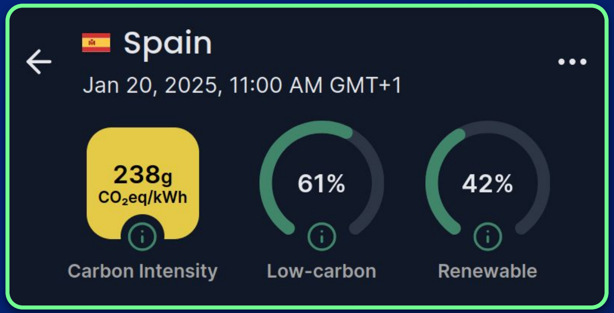


→ We now have available and stored on our site:

- CO2 Intensity from our zone (Spain)
- Grid fossil fuel percentage of our zone

→ Metrics are at our hourly resolution

→ We may need to obtain a few more metrics, e.g. water footprint, from hydropower generation.



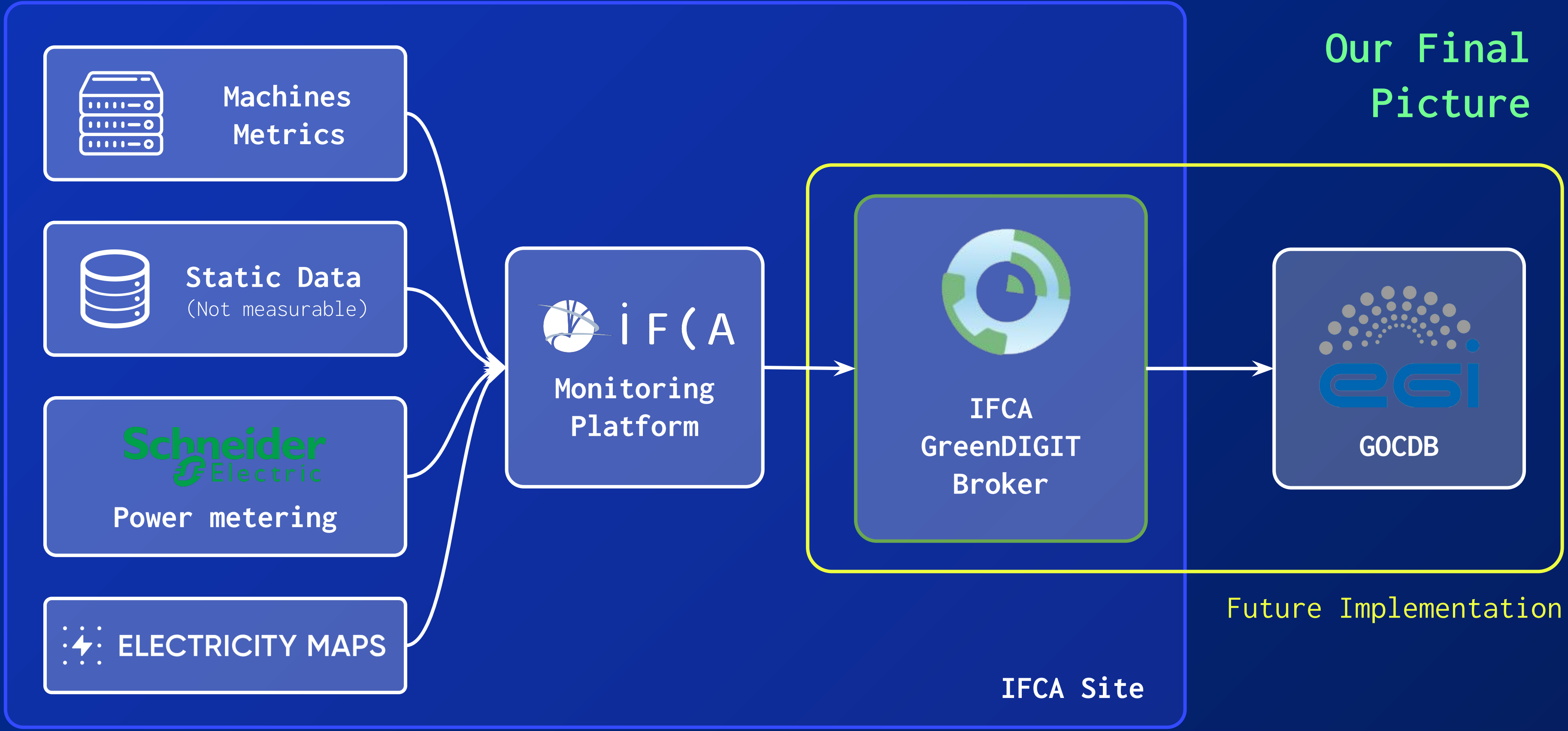
Partial IFCA Carbon Footprint

Live demo at:
https://monitor.ifca.es/grafana/

Final Picture

To sum up and define the next step

Our Final Picture



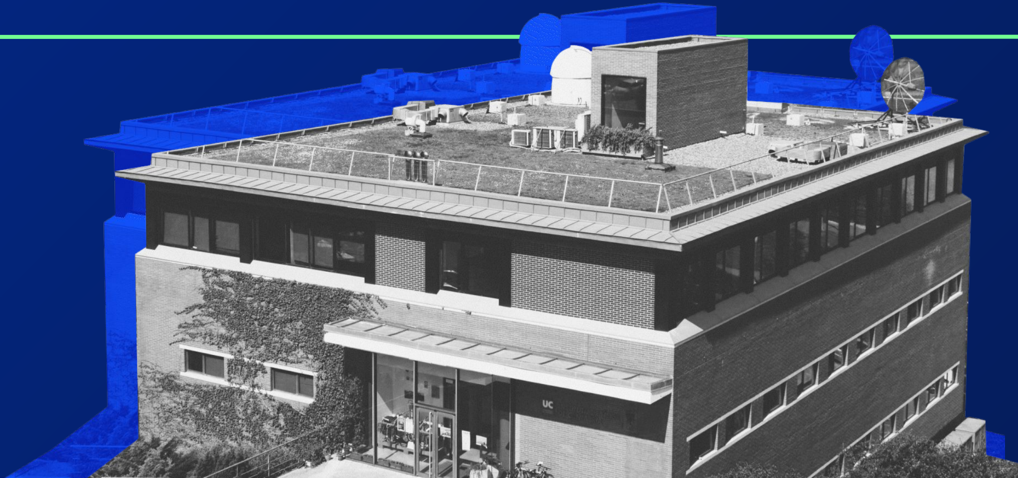
Instituto de Física de Cantabria
CSIC — Consejo Superior de Investigaciones Científicas
UC Universidad de Cantabria

GreenDIGIT | WP4/WP5/WP6 Design Workshop
Amsterdam Science Park
January 23rd/24th 2025

Thanks for your attention

Jaime Iglesias Blanco
iglesias@ifca.unican.es

Advanced Computing and e-Science Group
https://advancedcomputing.ifca.es



Instituto de Física de Cantabria

CSIC
Consejo Superior de Investigaciones Científicas

UC Universidad de Cantabria

GreenDIGIT | WP4/WP5/WP6 Design Workshop
Amsterdam Science Park
January 23rd/24th 2025

GreenDIGIT WP6: Energy-efficiency aware schedulers

Environmental impact aware schedulers for IFCA and AI4OS

Jaime Iglesias Blanco

iglesias@ifca.unican.es

Advanced Computing and e-Science Group
https://advancedcomputing.ifca.es

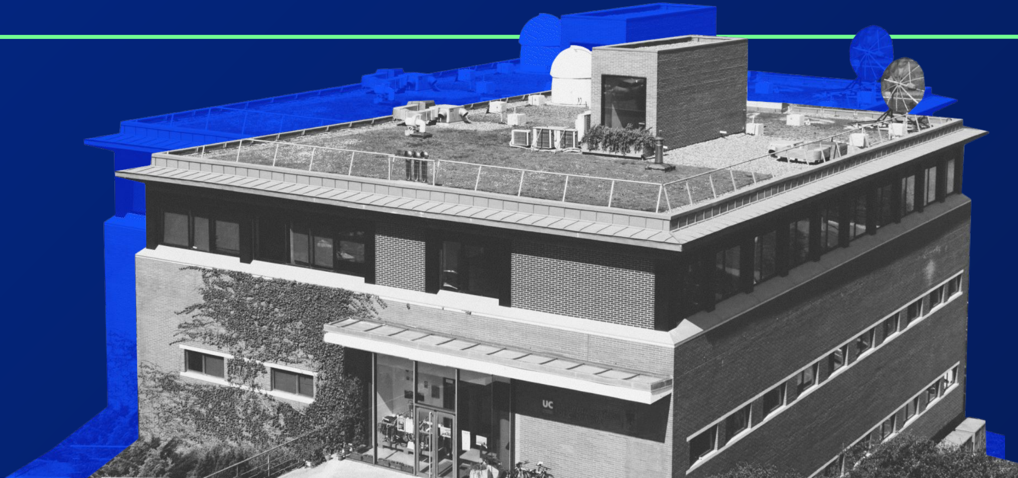


Table of contents:

1. Introduction
2. Schedulers at infrastructure level
   2.1. IFCA Cloud energy-aware scheduler
3. Schedulers at platform level
   3.1. AI4OS energy-aware scheduler

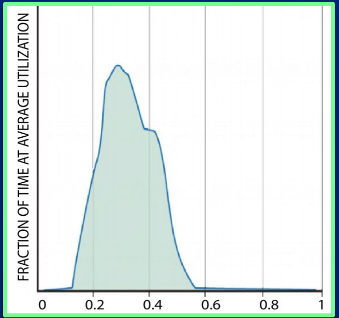
Introduction

Our initial thoughts and why we need energy-aware schedulers

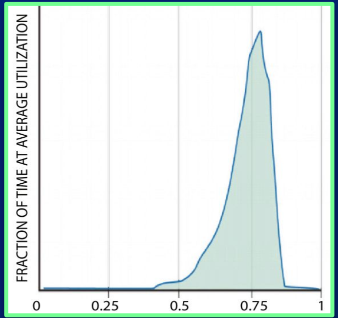
Introduction

As mentioned above, our site offers cloud computing services to users and deploys powerful platforms such as AI4EOSC.

- Over time, by monitoring our services, we observed that our datacenter could be overdimensioned for the usual workload.
  - Not only in our site, it happens in practically all of them.



Current State



Desired State

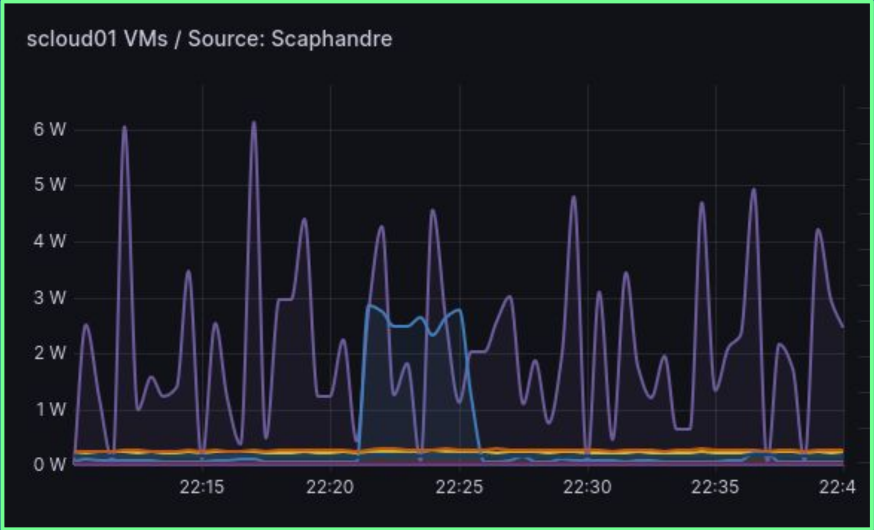
These are our main ideas in this area, but we are still discussing and prototyping them.

So far, we have focused more on the previous part of obtaining metrics to establish what data is available and from there, develop the software, with strong fundamentals.

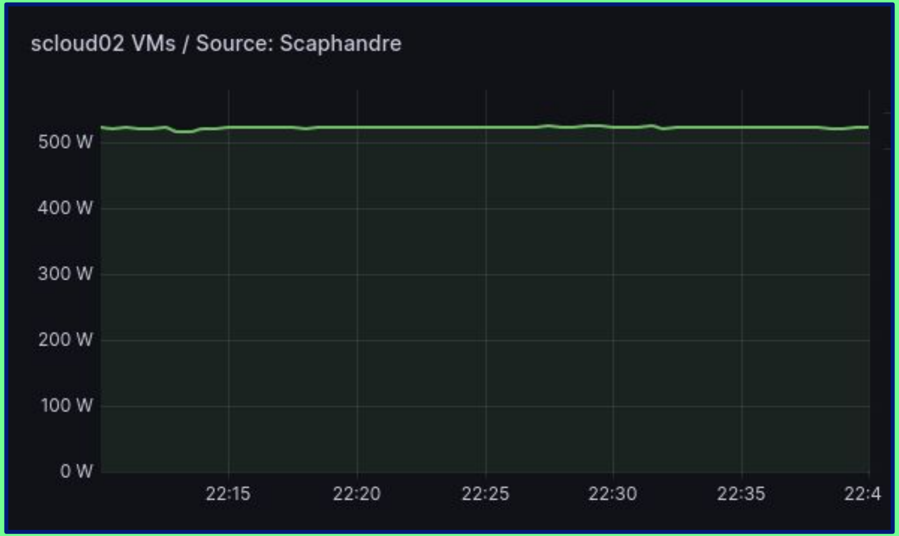
Schedulers at infrastructure level

IFCA Cloud energy-aware scheduler

The main idea:



Live Migration & Consolidation



IFCA Cloud energy-aware scheduler

Description: An energy-aware dynamic resource placement manager for cloud computing infrastructures. It has a extensible and modular design, so it can interoperate with multiple cloud infrastructure managers (in our case: OpenStack) and with a large number of heterogeneous equipment, using various communication and management protocols.

Keypoints:

- ➔ Monitor the infrastructure and the power grid to feed the optimization models with the fair data needed.
  - For example, percentages of resource use, energy consumption, carbon intensity…
- ➔ Develop and train server consolidation models.
- ➔ Perform VMs live migrations (in our case through OpenStack).
- ➔ Develop and train models to estimate demand.
- ➔ Develop and train models for deciding whether to turn servers on/off.
- ➔ Perform dynamic switching on/off of equipment on demand (via IPMI).

Status: Prototyping

Schedulers at platform level

AI4OS energy-aware scheduler

AI4OS energy-aware scheduler

AI4OS stands for Artificial Intelligence for Open Science. It is the collection of software and tools that allows to build:

This cloud platform is orchestrated through the **Nomad** workload management system.

We intend to develop similar strategies as for cloud scheduler but at platform level, adding some energy-aware capabilities to the nomad scheduler in our platform.

GreenDIGIT | WP4/WP5/WP6 Design Workshop
Amsterdam Science Park
January $23^{rd}/24^{th}$ 2025

Thanks for your attention

Jaime Iglesias Blanco
iglesias@ifca.unican.es

Advanced Computing and e-Science Group
https://advancedcomputing.ifca.es