May 2026, No 4

The Environmental Impact Metric Publication System (EIMPS)

Welcome to the 4th edition of the GreenDIGIT Newsletter! Following our overview of Key Exploitable Results, this edition deep-dives into one of GreenDIGIT's critical operational artifacts: the ***Environmental Impact Metric Publication System (EIMPS)****. As a comprehensive prototype service, EIMPS facilitates automated publication, validation, normalization, and access to sustainability metrics across highly heterogeneous digital research infrastructures. In this issue, we highlight how EIMPS bridges the gap between hardware telemetry and actionable carbon accounting for High-Throughput Computing (Grid), Federated Cloud deployments, and different Networking topologies.*

Introducing EIMPS: A Unified Metrics Framework

The Environmental Impact Metric Publication System (EIMPS) represents a monumental leap forward in achieving operational transparency for European Research Infrastructures (RIs). Developed as a foundational prototype service within the GreenDIGIT project, EIMPS resolves a long-standing challenge: monitoring, assessing, and reducing energy consumption across heavily fragmented digital environments. By providing an end-to-end telemetry and publishing pipeline, EIMPS unifies disparate performance streams from Grid, Cloud, and Network environments. It maps them to a harmonized metadata schema via a robust Common Information Model (CIM) and stores them systematically in a centralized Metrics Database (Metrics DB). The overview of the proposed system, as well as all the major imvolved components are depicted in Figure 1.

Through an authorized RESTful Publication API, distributed resource providers automatically submit records conforming to the GreenDIGIT Execution Unit Record (EUR) structure. The system seamlessly enriches raw energy data with real-time site parameters, using a standardized carbon tracking formulation:

$$\mathrm{CO_2}\ [\mathrm{kg}] = \mathrm{Energy}\ [\mathrm{kWh}] \times \mathrm{CO_2}\ \mathrm{intensity}\ [\mathrm{kg/kWh}]$$

By factoring in local grid carbon intensity (via *T6.5 - Wattnet*) and facility infrastructure performance indices like Power Usage Effectiveness (PUE) (via *GOCDB*), EIMPS elevates low-level physical resource data into precise, verifiable Key Performance Indicators (KPIs) tailored for modern sustainability auditing.

The EIMPS Processing Pipeline

The entire GreenDIGIT metric publication lifecycle is structured into five cohesive pipeline stages, balancing local edge measurement with centralized semantic enrichment:

1. **Data Collection:** Specialized edge agents active within distributed facilities measure high-granularity energy footprints and compute performance dimensions.
2. **Record Creation:** Raw telemetry streams are formatted into JSON-based objects following the standardized EUR schemas.
3. **Transmission to CIM:** Edge modules securely push records via HTTPS POST requests using robust OAuth2 Bearer token authentication.
4. **Validation and Insertion:** The CIM service verifies structure conformity, fetches location metadata,

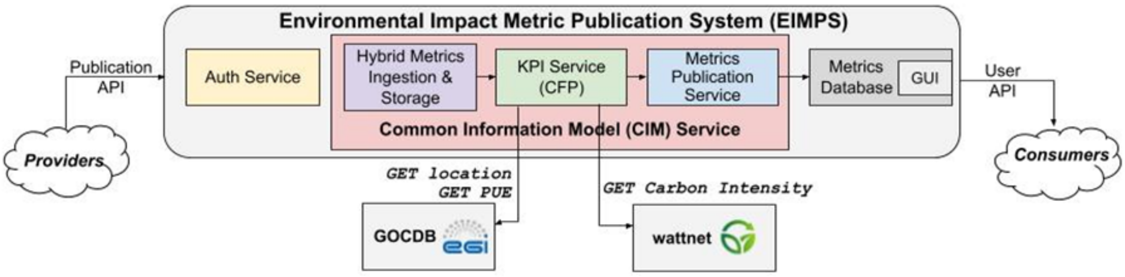

Figure 1: EIMPS architecture and components.

dynamically injects computed carbon payloads, and commits records to database layers.

5. **Metric Publication:** Fully qualified data fields become instantly accessible to dashboards, workflow managers, and policy reporting platforms.

Diversified Telemetry for Heterogeneous Workloads

Because Grid, Cloud, and Network environments operate on fundamentally different resource virtualization paradigms, a single, monolithic data collection approach would be insufficient. High-Throughput Computing relies on batch job logging and model-based hardware estimations, Cloud environments require fine-grained, hypervisor-level sampling of isolated virtual instances and complex Network topologies demand a hybrid mix of direct hardware telemetry and synthetic scaling factors across wired and wireless systems. Consequently, several distinctly specialized energy reporting workflows had to be engineered to extract, normalize, and deliver accurate sustainability metrics to the EIMPS core without disrupting local provider operations. The subsections below, summarize some key details and characteristics for each one of the examined approaches.

1. High-Throughput Compute (Grid) Workflows

High-Throughput Computing (HTC) configurations federated across the EGI ecosystem execute massive volumes of batch processing jobs. Managing environmental observability in this context requires seamless tracking without imposing overhead onto administrative staff or deploying privileged agents on worker nodes. To accomplish this, the GreenDIGIT project introduced **GreenDIRAC**, an energy-aware expansion of the widely adopted DIRAC Workload Manager. GreenDIRAC operates natively through two newly designed internal daemons:

- **GreenReportingAgent:** This daemon scans historical records of terminated computational tasks and calculates CPU energy usage. It bypasses restricted hardware power sensors (e.g., Intel RAPL) by referencing a read-only hardware configuration database mapping CPU model numbers directly to their baseline Thermal Design Power (TDP) specifications.
- **GreenSiteDirector:** A forward-looking, intelligent scheduling component designed to dynamically guide pending batch configurations toward distributed resource centers exhibiting superior environmental profiles or lower instantaneous carbon intensity scores.

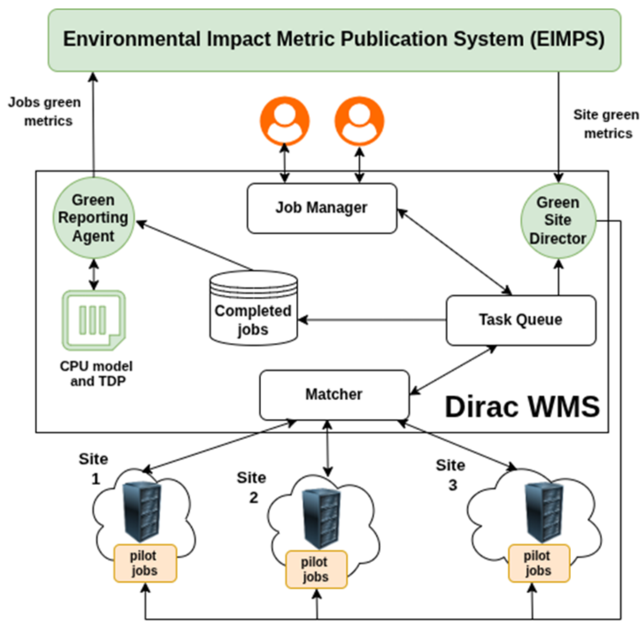

Figure 2: GreenDIRAC Monitoring.

2. Cloud Infrastructure Reporting Workflows

Observability across Infrastructure-as-a-Service (IaaS) configurations (primarily OpenStack platforms running KVM/QEMU hypervisors) relies on precise, non-intrusive hypervisor-side power monitoring.

The GreenDIGIT framework pairs an open-source metrology tool, **Scaphandre**, with a customized account-

Research Infrastructure	Type	Country	Granularity	Volume (Records)	First EUR Timestamp
EGI*	HTC	FR, CZ, UK, NL, IT, DE	Real time	8,074,676	October 2025
EGI	Cloud	FR, CZ, ES	6-hours	55,133	November 2025
SLICES*	Network	GR, PL	6-hours	46,799	January 2026
SoBigData*	Cloud	IT	48-hours	52	March 2026
KM3NeT	HTC	NL, FR	Real time	381	February 2026
EMPHASIS	Cloud	CZ	6-hours	820	March 2026
OPERAS	Cloud	FR	6-hours	115	March 2026
ELIXIR	Cloud	FR	6-hours	46	March 2026
Instruct-ERIC	Cloud	FR	6-hours	253	March 2026
EMSO-ERIC	Cloud	ES	6-hours	7,083	November 2025
Euro-Bio-Imaging ERIC	Cloud	CZ	6-hours	1,558	March 2026

Table 1: Summary Table of Integrated RIs on EIMPS and Ingestion Statistics.

ing extractor named **cASO**, with its details to be summarized below:

1. **Hypervisor Collection:** A Scaphandre systemd module runs continuously with root privileges on individual physical hypervisor hosts to interrogate processor RAPL interfaces directly. It exposes high-resolution consumption data via an embedded Prometheus exporter port.
2. **UUID Correlation:** Instantaneous microwatt utilization fields are captured along with virtual machine KVM domain labels. Prometheus rules extract unique OpenStack UUID metadata strings from process command lines to cleanly link hardware usage to user instances.
3. **Accounting Extraction:** Operating at defined cycle frequencies, the expanded *cASO* extractor communicates with OpenStack endpoints to fetch execution windows and scrapes the Prometheus time-series metrics. It then applies an algorithmic *energy_normalization* factor to accurately compute total facility scale overheads beyond pure processor draw.

3. Diverse Network Activity Tracking

Unlike rigid computing hardware, tracking data transport layers entails mapping heterogeneous wired, wireless, and edge-computing networking segments across European topologies.

Because conventional routing switches, radio gear, and edge points rarely feature standard hardware energy accounting registers, GreenDIGIT developed specialized estimation frameworks tailored to three primary networking deployment tracks:

- **IoT Resource Elements:** Focuses on low-power architectures like wireless Wi-Fi HaLow adapters(IEEE 802.11ah)). Physical power profiles are collected using inline hardware wall-plug adapters, which are correlated alongside a dedicated software estimation library (**HERMIS**) to evaluate micro-scale CPU and wireless radio consumption.
- **Wired Enterprise Infrastructure:** Involves core switches and wide-area link infrastructure. Customized device agents query active interface counters via standard protocols (such as NETCONF, SNMP, or vendor interfaces). They project energy overhead by mapping real-time data throughput rates directly against vendor baseline power parameters.
- **5G Core & Access Lanes (RAN):** Evaluates virtualized core routers and complex radio arrays. Power distribution unit (PDU) logs are correlated alongside Scaphandre agents to map the energy footprint of virtual functions (e.g., UPFs, DUs) across high-velocity 5G traffic profiles.

EIMPS summary & Current Status of Integrated RIs

Beyond direct sustainability tracking, the standardized metrics published by the EIMPS serve as a vital data foundation for advanced analytics and cross-WP (WP6 / WP7) automation within the project. The centralized data system provides the historical and high-granularity data entries necessary to train and validate machine learning (ML) models for predictive energy forecasting. Integrated as an exogenous data source for T6.2, it provides valuable information to further feed and extract decisions in T6.3 (brokering system). Finally, EIMPS integrates seamlessly with various validation technologies and infrastructure orchestration platforms that will be evaluated in WP7 use-cases.

At this point, 10 RIs have already been integrated and report environmental sustainability metrics since EIMPS was passed in the "production" level. Each one's detail are given in Table 1. There, those marked with "*" be-long to GreenDIGIT's consortium, while the rest of them composed of other ESFRI RIs.

Stay Connected

- Check further details of EIMPS (documentation & dashboard) here.
- All public deliverables and materials are available here.
- Follow our updates on LinkedIn and X.