GreenDIGIT: Project and Initiative to Lower Environmental Impact of Future Digital Research Infrastructures

Defining Architecture Elements for RI Sustainability

Yuri Demchenko

GreenDIGIT Project, University of Amsterdam

1 October 2024

GreenDIGIT – EGI2024 Conference Joint Session

Outline

- GreenDIGIT project scope and goals
- Shared Responsibility Model for Sustainability
  - RI Role and actors definition: RI providers/operators – Research applications developers - Researchers
- Sustainability by Design: Infrastructure components and aspects to be addressed
- Ongoing work and expected results:
  - Energy and CO2 impact by scientific workflows
  - Energy optimisation for ML/Data analytics model building and inference
  - DevOps extension with software Sustainability analysis and assessment – Future DevSustOps

Funded by the European Union.
Grant ID: 101131207

GreenDIGIT project (2024-2027) – Objectives

- **O1: Assess status and trends** of low impact computing within 4 DIGIT RIs (EGI, SLICES, SoBigData, EBRAINS) and in the broader digital service provider community of ESFRIs, to produce **recommendations and roadmaps** for providers for during and beyond the project.
- **O2: Provide reference architecture and design principles**, as well as an actionable model for RIs about environmental impact assessment and monitoring, reflecting on the **whole RI lifecycle** and including the digital infrastructure components and their interaction with the broader environment.
- **O3: Develop and validate new and innovative technologies, methods, and tools** for digital service providers within European Research Infrastructures through which they can reduce their energy consumptions and overall environmental impact.
- **O4: Develop and provide for researcher technical tools** that assist them in the design, execution and sharing of environmental impact aware digital applications with reproducibility, Open Science and FAIR data management considerations.
- **O5: Educate and support digital service providers and researchers** in the RI communities about good practices on environmental impact conscious lifecycle management and operation of infrastructures and services.

Energy Efficiency – Decarbonisation – Environmental Impact

Architecture, Design, Recommendations

- **Energy Efficiency in Digital Infrastructures**:
  - **Definition**: This refers to optimising digital infrastructures to consume as little energy as possible for a given workload or service. It's about achieving more computational or storage results with less energy input.

Operation, Monitoring. KPI

- **Decarbonisation of Digital Infrastructures**:
  - **Definition**: This specifically targets the reduction of carbon emissions associated with the operation and maintenance of digital infrastructures.

Lifecycle, Policy, Training

- **Reducing Environmental Impact of Digital Infrastructures**:
  - **Definition**: This is a more comprehensive consideration of the various ways digital infrastructures might affect the environment, going beyond just energy consumption and carbon emissions.

Importance of the Architecture definition

- Architecture is a way to coordinate/synch/unite
  - Developers of Infrastructure and Applications
  - Operators
  - Users/Researchers
  - Policy and decision makers
  - Refer to TOGAF architecture principles (as an approach accepted by the majority of businesses)
- A basis for linking standards and regulations to functional architecture components
  - Ensure compliance of the designed/developed RI and services (including for audit)

GreenDIGIT Architecture Definition Methodology

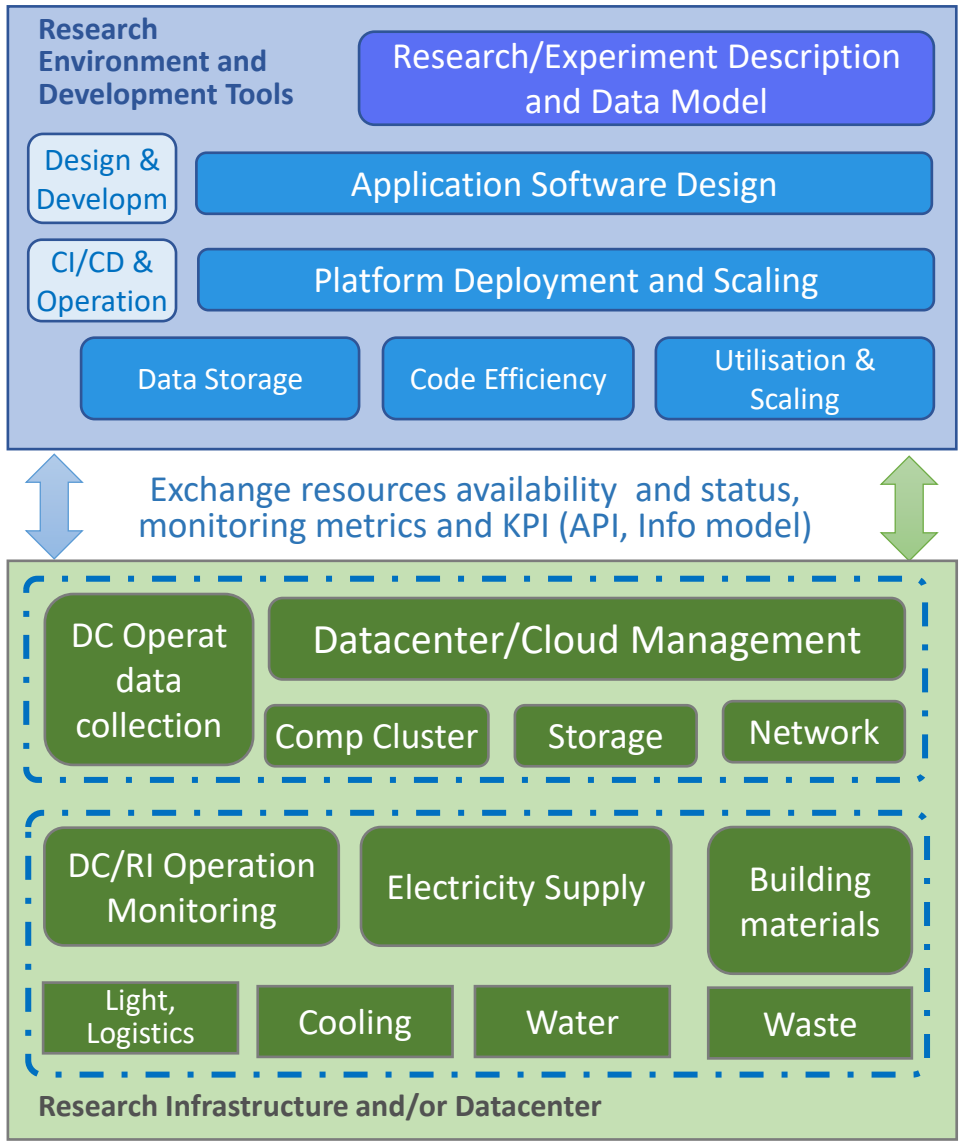
- General view on the RI Ecosystem Optimisation and Green IT
  - Horizontal, vertical, lifecycle
  - RI Operators and Researchers
  - RI continuum: (Research Object) – Sensor - (RAN) – Edge – Cloud – Workflow - Researcher
- Sustainable architecture design principles
  - As a basis for modelling and metrics for RI infrastructure operation and optimisation
  - **Shared Responsibility Model**: RI provider/operator and Researchers/Projects
  - **Sustainability by design** – A novel concept to be introduced addressing different aspects and stages
- Linkage with existing Standards and Regulations to ensure Sustainable Architecture Design principles support compliance with the standards, regulations and audit
  - To provide the opportunity for RI/datacenter operation (and design) optimisation (through the whole lifecycle)
- System Engineering and Design (thinking) approach in Green research and technologies

Shared Responsibility in Sustainability – Reflecting Operational and Management Aspects

GreenDigit

Users responsible for sustainability ***on*** the RI

Providers responsible for the sustainability ***of*** the RI



Standards and regulations
Software Development Quality and Design Patterns

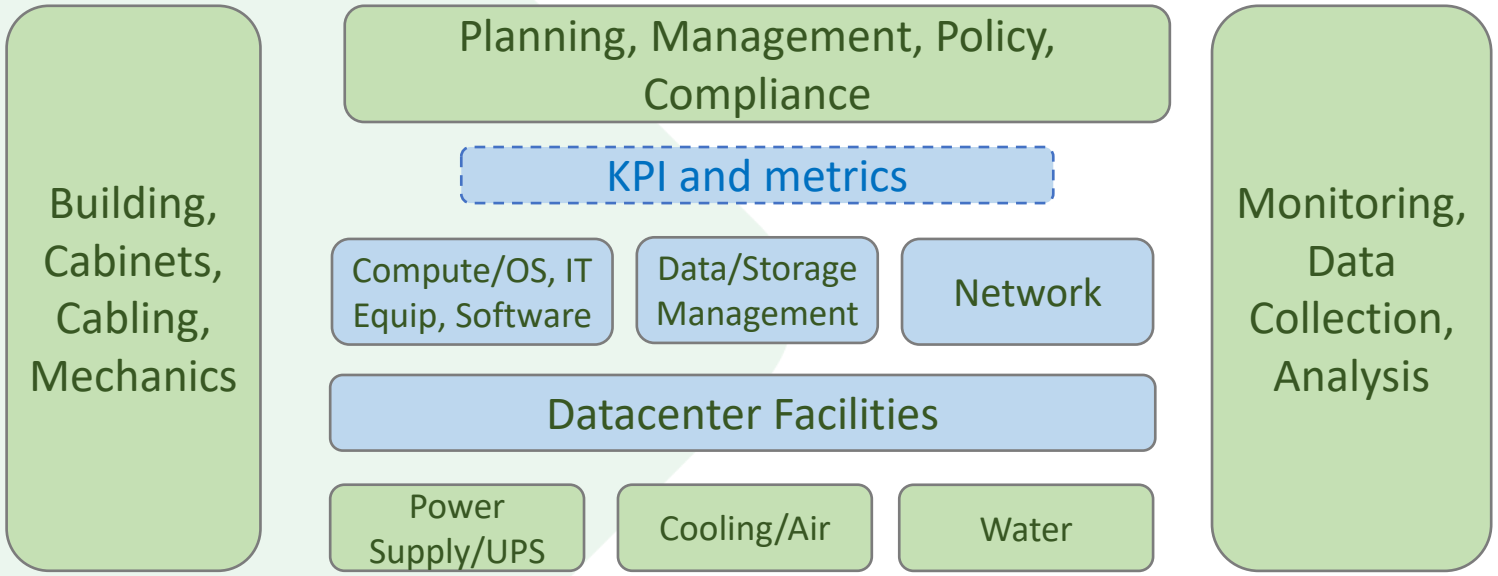
Project/Researcher Responsibility:
Applications Development, Deployment, Operation, Energy usage and KPI monitoring

Provider/Operator Responsibility:
Research Infrastructure or Datacenter, Monitoring Energy and environmental impact metrics and KPI

Standards and regulations
Datacenter and RI Building and Operation

Best Practice for the EU CoC on Data Centre Energy Efficiency: Main Infrastructure components – RI/Datacenter Operator Perspective

Also supported in the EC Delegated Regulation (EU) 2024/1364 of 14 March 2024



Areas of responsibility/ management

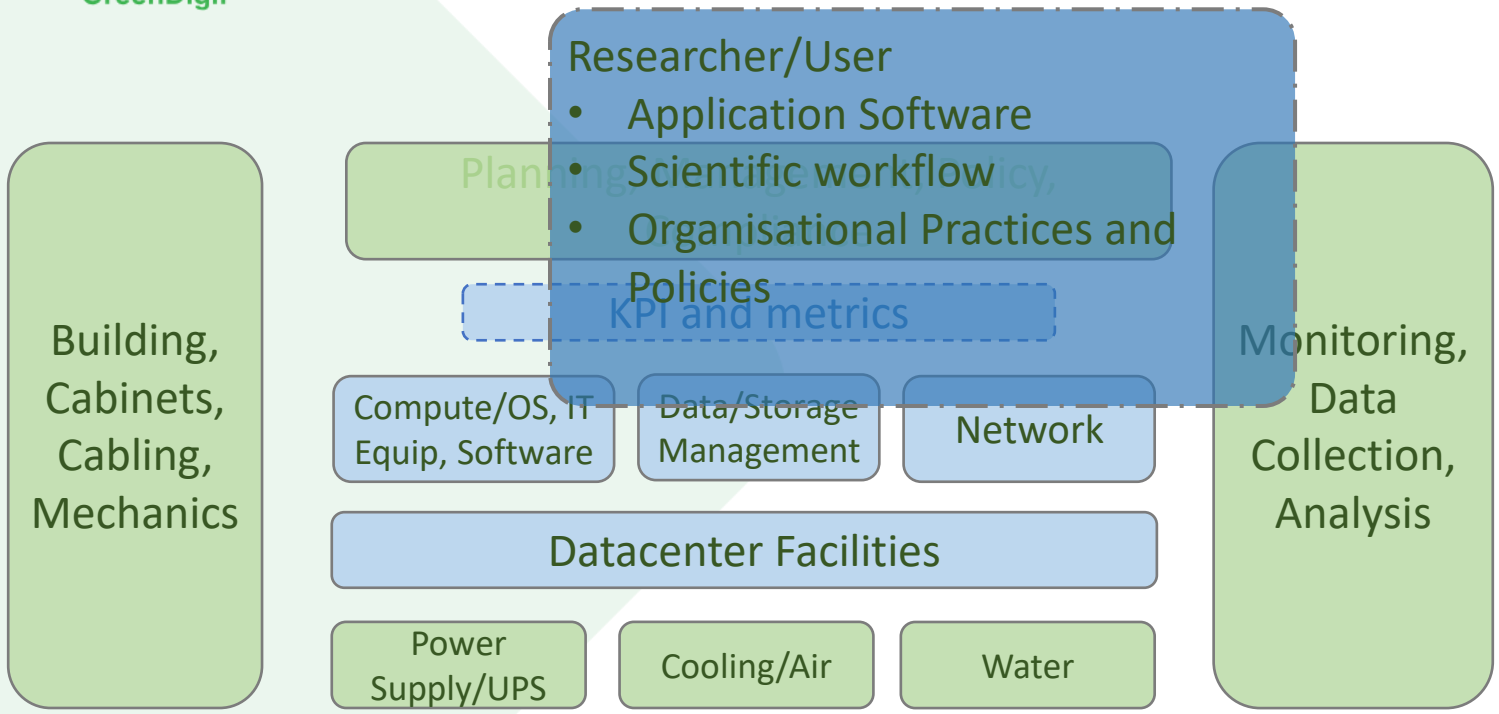
- Physical building
- Mechanical and electrical plant
- Data floor and air floor
- Cabinets and cabinets airflow
- Metrics and operation measurement points (?)
- **IT equipment**
- **Operating systems and virtualisation**
- **Software**
- Business Practices

Different roles of participants

- Operator
- Colocation provider
- Colocation customer
- Managed services provider
- Managed services provider in colocation space

[ref] https://e3p.jrc.ec.europa.eu/sites/default/files/documents/publications/jrc136986_2024_best_practice_guidelines.pdf

Best Practice for the EU CoC on Data Centre Energy Efficiency – Researcher/User Perspective



Areas of responsibility/ management

- Research Development Environment
  - Application Software
  - Scientific workflow
- Control/Management RI/datacenter platform
  - Applications/SW lifecycle
  - IT equipment operation/ control and Optimisation
  - Operating systems and virtualisation
- Organisational Practices and Policies

Different roles of participants

- Managed services provider
- Research organisation, project
- Researcher (scientific workflow, data collection, scheduling)

Also supported in the EC Delegated Regulation (EU) 2024/1364 of 14 March 2024

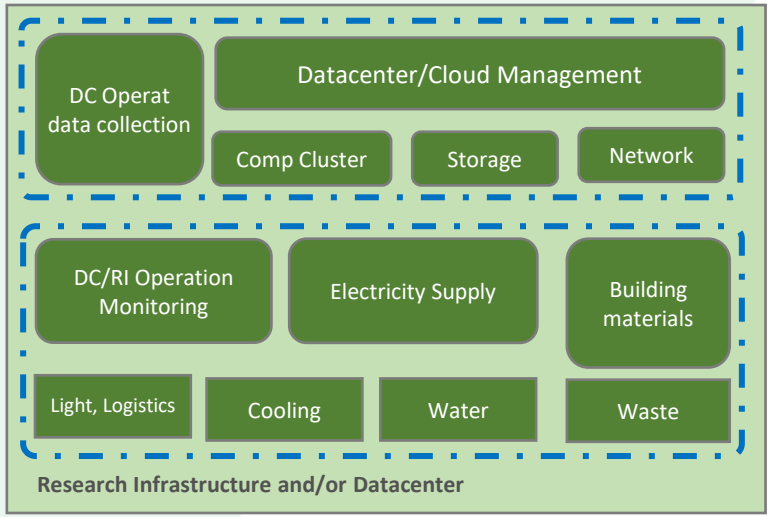
Linking KPI, Metrics and Design Patterns – To be clarified and extended

KPI (Key Performance Indicators)

- Energy consumption (kWh).
- PUE (Power Usage Effectiveness): Ratio total energy used to energy consumed by IT
- Carbon footprint: Amount of CO_2 emissions associated with energy usage.
- Uptime or reliability: How well the infrastructure maintains consistent service
- Resource utilization efficiency: Resources (servers, storage, etc.) usage w/o wasting energy.

Metrics

- CPU usage and load: The utilization levels of computing resources.
- Energy draw per server.
- IT equipment energy consumption
- Temperature: The internal temperature of servers or cooling systems.
- Cooling system efficiency



Design Solutions/Patterns

Under RI provider control/management

- Energy-efficient hardware
- Green energy sources.
- Monitoring and energy efficiency analytics and optimisation at the level of RI provider
- Advanced cooling systems
- Smart grid integration

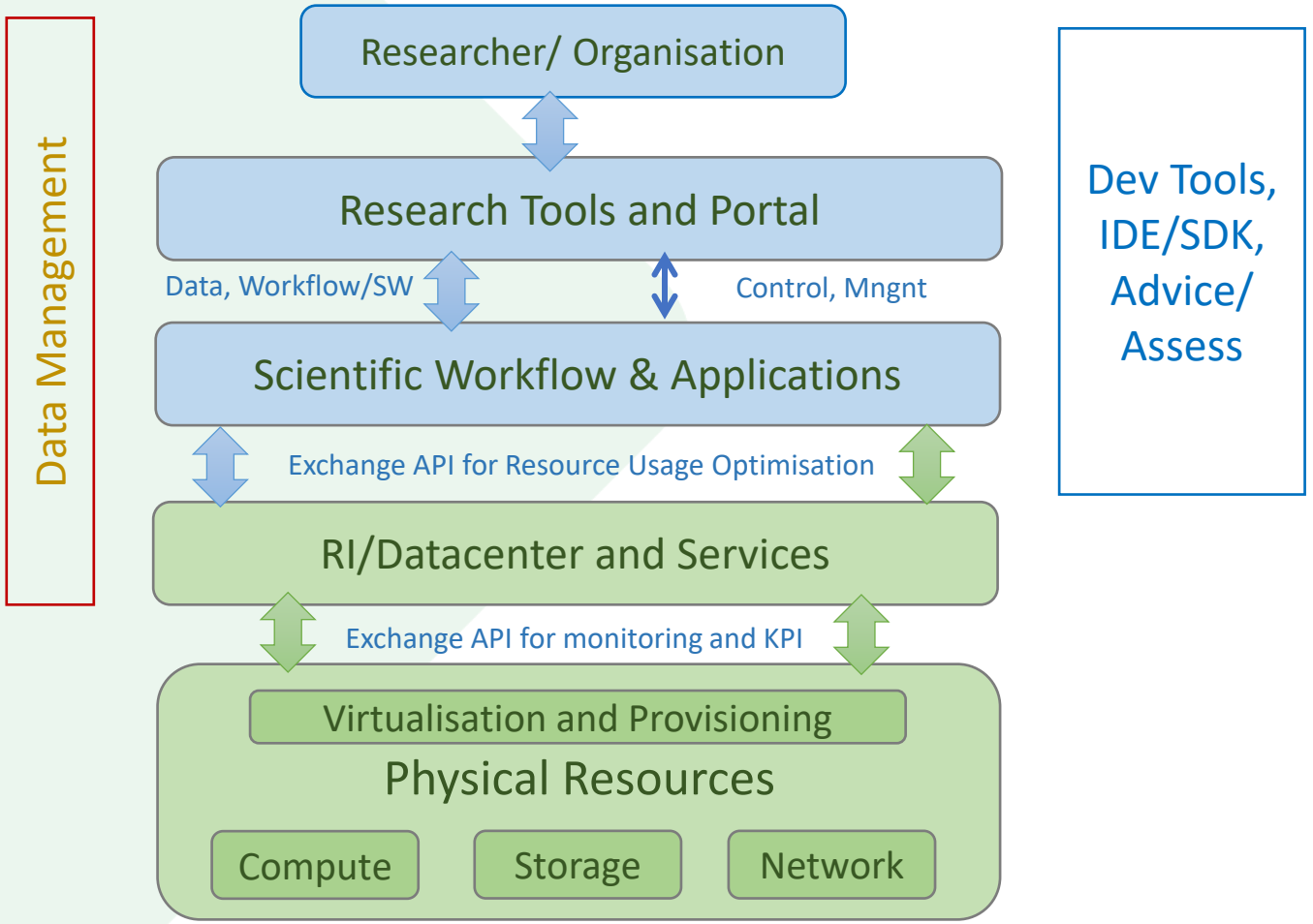
Under user/developer control

- Energy and environment aware applications development
- Modelling, simulation and testing as part of the development process
- Green software practices and templates
- Virtualization and containerization

RI/Systems Sustainability by Design Components/Aspects

- Architecture for Sustainability by Design
  - Functional components, layers, API, Requirements
- Software and application components that can be optimised during design and controlled during operation
  - Corresponding optimisation model to be proposed
  - Supported with VRE and IDE with sustainability awareness
  - Sustainability design patterns
- Link and interaction between components via APIs
  - Green aware API including necessary energy, performance, environment information
- Common information/data model and metadata (naming)
  - Including Requirements, KPI, Metrics
  - Verified with existing standards
  - Compliant with PUE and other KPI
- Relations between the system/RI and software sustainability aspects to be defined
  - System related aspects: Lifecycle, Operation and Governance, metrics on energy and GHG
- RI and applications lifecycle
  - Scientific workflow and research data lifecycle to be aligned with sustainability policy and monitoring

RI/Datacenter Components and Layers



- Researcher
- Research Tools and Portal
- Scientific workflow and Applications
- RI/Datacenter and Services
- Physical resources (compute, storage, network)
  - Controlled, monitored, optimised

Discussion and Questions

- GreenDIGIT project is in active development state
- Active dissemination and outreach activity is planned
- Research community contribution will be solicited

Existing Initiatives and Developments for Green Software and Applications

- Green Software Foundation https://greensoftware.foundation/
  - Carbon Aware SDK, Green Software Patterns
- SustainableIT - https://www.sustainableit.org/
  - IT Standards for Environmental, Social, and Governance Sustainability
- ICT Footprint: European Framework Initiative for Energy & Environmental Efficiency in the ICT Sector - https://www.ictfootprint.eu/
  - Self-assessment Framework (Organisations) – Light Certification Scheme
- Cloud Carbon Footprint - https://www.cloudcarbonfootprint.org/
- RIPE NCC Green IT and Network activity
  - Green Hackathon https://labs.ripe.net/author/becha/announcing-the-green-tech-hackathon/
- AWS Sustainability Pillar and Shared Responsibility *https://docs.aws.amazon.com/wellarchitected/latest/sustainability-pillar/the-shared-responsibility-model.html*
-

Standards and Regulation – Overview and Analysis

- Architecture Requirements based on Standards, BCP and Regulations
  - Mapping standards requirements to infrastructure functional components
- EU Code of Conduct for data centers - Best Practices published annually by JRC
- European standards series EN 50600
- International group of standards ISO 50001, ISO 50002, ISO 14001, ISO 30134

- Full Summary of the standards and regulations analysis will be produced in a separate document – To be open for Research community

Landscape of the Sustainability Research, Development Policy - Goal to define GreenDIGIT place and contribution

General, not technical

- Economy and Financial sustainability
- Social and Community
- Cultural Sustainability
- Governance and Policy
- Sustainability in Supply Chain

Research, technology, circular operation

- Technology
- Manufacturing
- Environmental Research and Monitoring
- Research and Research Infrastructure/Facility
- Education and Competences

Place for GreenDIGIT

- Research domains
- Operation
- Education and competences

- In general, to promote the Shared Responsibility Model for Sustainability