universität freiburg

Physikalisches Institut
Albert-Ludwigs-Universität Freiburg

Advancing the Environmental Sustainability of Scientific Computing for ErUM

SUSFECIT

Sustainable Federated Compute Infrastructures

Markus Schumacher
on behalf of the proposed SUSFECIT Research Network
(Aachen University, Bonn University, DESY, Freiburg University, Göttingen University, KIT, Öko-Institut e.V.)

2nd ErUM-Data Workshop on Sustainability
Aachen, 29.7. 2025

Hopefully
With funding from the:
Federal Ministry of Research, Technology and Space

The proposed Consortium

Partners

RWTH AACHEN UNIVERSITY

Achim Stahl

DESY.

Martin Gasthuber

UNIVERSITÄT BONN

Philip Bechtle

UNIVERSITÄT GÖTTINGEN

Arnulf Quadt

KIT Karlsruher Institut für Technologie

Manuel Giffels

Öko-Institut e.V. Institut für angewandte Ökologie Institute for Applied Ecology

Jens Gröger

universität freiburg

Markus Schumacher

Associated Partners

CERN

Simone Campana

universität freiburg

Marc Schumann
Dirk von Suchodoletz

universität freiburg

KIT Karlsruher Institut für Technologie

Erik Bründermann
Achim Streit
Kathrin Valerius

UNIVERSITÄT SIEGEN

Christian Gutt

BERGISCHE UNIVERSITÄT WUPPERTAL

Christian Zeitnitz

Energy Consumption, CO_2 footprint of Compute Centres in D in 2024

- Electricity demand of compute centres is increasing.
- In Germany ~ 20 billion kilowatt hours in 2024, corresponds to ~ 3.9 % of total German electricity consumption.
- We expect this to double in the next 4 to 5 years (Felix, Behrens, Öko-Institut)

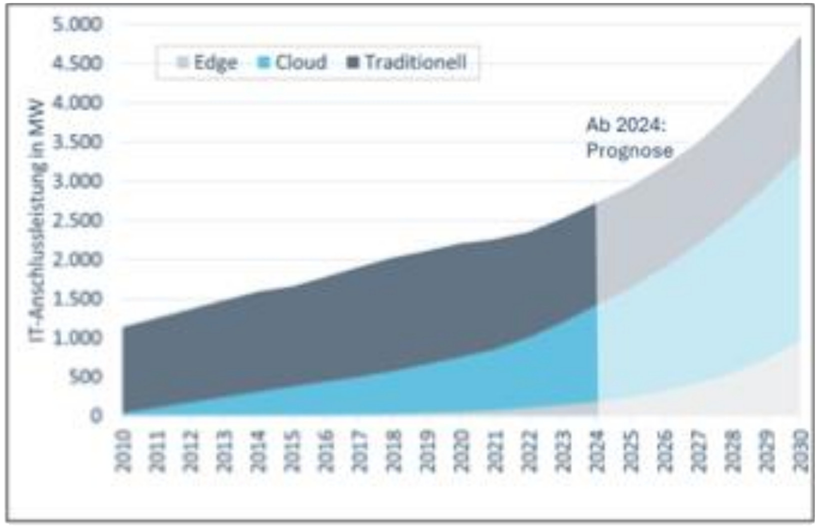
- In 2024, an average of 363 gr of CO_2 were emitted per kWH of electricity consumed in Germany during generation (386 (433) gr/kWh in 2023 (2022)) (Source: Umweltbundesamt)

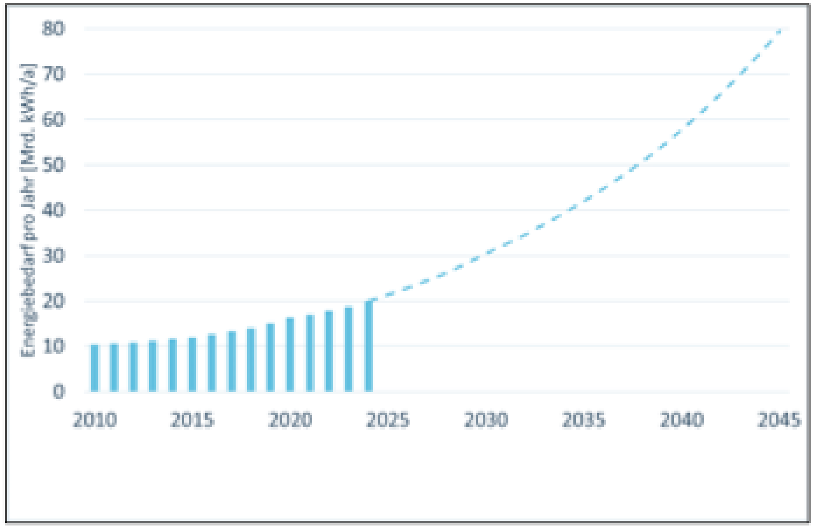
→ CO_2 footprint by compute centres: 7.28 Mt CO_2 in 2024

- Compare:
  Germany in total 649 Mt $(CO_2)_{eq}$ (572 Mt CO_2, energy industry 177 Mt CO_2) (Umweltbundesamt)

- Goal of CO_2 neutrality in 2045 in Germany and 2050 in EU

→ Advance environmental sustainability of scientific computing





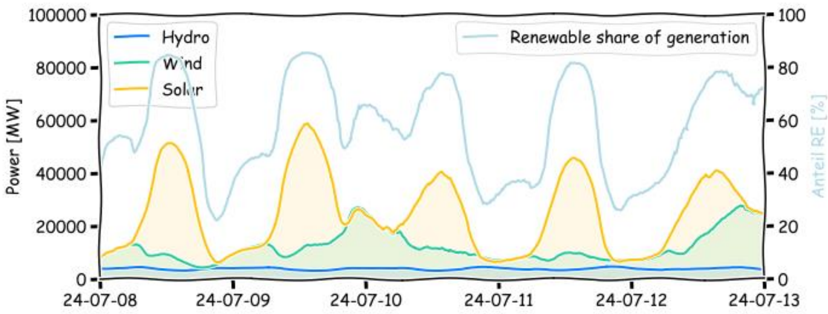
Stand und Entwicklung des Rechenzentrumsstandorts Deutschland ,Gutachten im Auftrag des BMWK
BMWK-Projekt-Nr.: 115/21-45

The Basic Idea of SUSFECIT

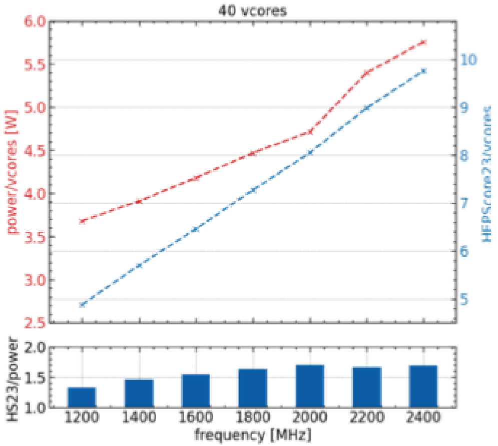
Sloppy: „Perform calculations when and where sun is shining or wind is blowing“

SUS FECIT
Sustainable Federated Compute Infrastructures

- A bit more precise:
  Adapt the amount of calculations and energy consumption to the temporal and local availability of renewable energies (RE) to be used in compute centres e.g. exploit that compute jobs are “easily“ dispatchable in time and space



- Means of adaption:
  - shut down (parts) of cluster
  - reduce clock frequency of CPUs (GPUs)
  - transfer jobs to sites where enough RE is available



- Requirements on hardware:
  - overprovisioning of resources to compensate for volatility of RE and dark doldrums
  - include energy storage (in/close-by compute centres) in concept to compensate for volatility of RE and dark doldrum (→ see presentation by Achim Stahl)

- Requirements on strategy of operation/ software ecosystems:
  - Forecast of available RE with good temporal and local resolution (and requested compute time)
  - Orchestration of jobs in time and on compound of sites according to available RE and compute demands (incl. check pointing for transfer of jobs between sites)
  - Accounting of elapsed time, consumed energy, produced CO_2 footprint (incl. LCA, used water, …)

The „House“ of SUSFECIT

Development of overall strategy
Links btw. three ecosystems

Forecasting

Orchestration

Accounting

Digital twin for above interlinked ecosystems
- Simulation of individual sites and compound of site
- Comparison to measured results from sites
- Optimisation of hyper parameters

Goals:
- Increase the environmentally sustainable use and operation of compute resources.
- Raise awareness of the impact of scientific computing on the CO_2 footprint
- Train talented early career scientists in cutting-edge technologies

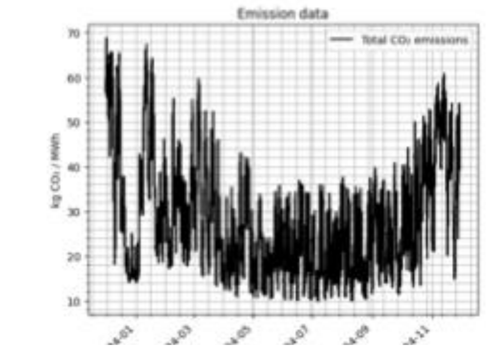
Dissemination:
- First developed and deployed for ATLAS, CMS and ET
- Later:
  - KATRIN, XENONnT, LXZD (neutrino and astro-particle physics)
  - Experiments at PETRA-III, EU-XFL, ESRF (synchrotron radiation physics)
- Perhaps: Further use cases identified

Links to PUNCH4NFDI, EOSC, WLCG via partners and associated partners
Associated partner will act as critical reviewers, alpha-users and multiplicators

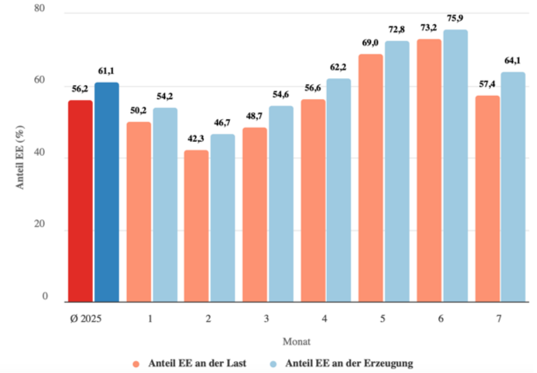
Forecasting+Optimization (lead & coordinated by Bonn, P. Bechtle; partners: AC, DESY, Öko)

First studies toward optimized operation of compute infrastructure performed.

- Taking into account temporal availability of renewable energies (RE), time dependent energy price, network costs, embedded CO_2 in hardware, ….
- First optimized values for replacement of hardware, thresholds for shutdown / throttleing of hardware w.r.t. minimal total costs or minimal total CO_2 footprint.



Energetisch korrigierte Werte - bis 28.07.2025, 09:30 MESZ

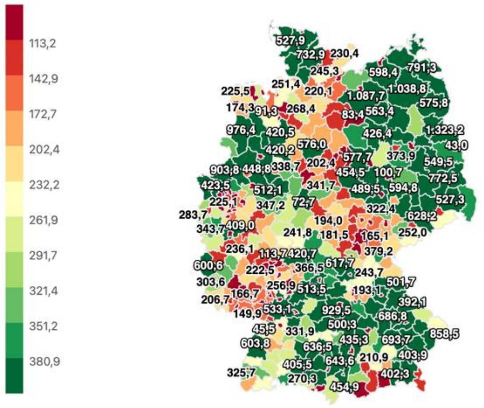


Forecasting:

decide whether in upcoming time window energy consumed by scheduled compute jobs can be covered by sufficient RE (from variable sources such as wind and sunshine) and how expensive energy is (depends on the residual load)

Installierte Solarleistung in Deutschland in 2025

Werte in MWp – Gesamtsumme: 108.352 MWp



Energy-Charts.info; Letztes Update: 27.07.2025, 12:27 MESZ

Optimization for overall CO_2 footprint (and/or lowest total investment):

- operation of a given set of compute centres to be optimized for the given amount of RE
- overprovisioning of resources to compensate for down times due to low RE
- full LCA of machines and infrastructure (storage, local photovoltaic, …)

Forecasting+Optimization (lead & coordinated by Bonn, P. Bechtle; partners: AC, DESY, Öko)

Steps of implementation:

- Creation of data base for available RE (initially w/ coarse local and temporal resolution)
  e.g. using energy-charts.info, and information from hardware suppliers and first model for forecasting

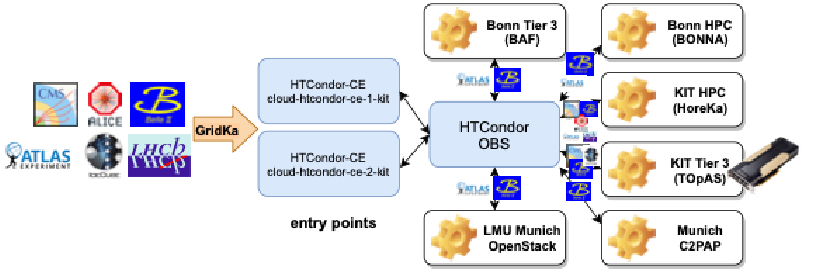
- Initial operationality and interfacing to AUDITOR, the digital twin
  - Develop algorithms for automatically deciding how to dispatch resources at any given time, taking into account the load of requested and queued compute jobs and the availability of RE
  - Decision transferred to COBalD/TARDIS and AUDITOR and incorporation of developed tools in „Digital twin"
  - Simulation of one year of running on federated resources
  - Use evaluation and feedback-loops to continuously optimize set-up

- Extension to further sources of CO_2
  Based on eco-digit project (Öko-Institut et al.) development and optimization of the forecasting methodology taking into account e.g. embedded CO_2 in the hardware, procurement and disposal, etc. (→ see talk by Martin Möller)

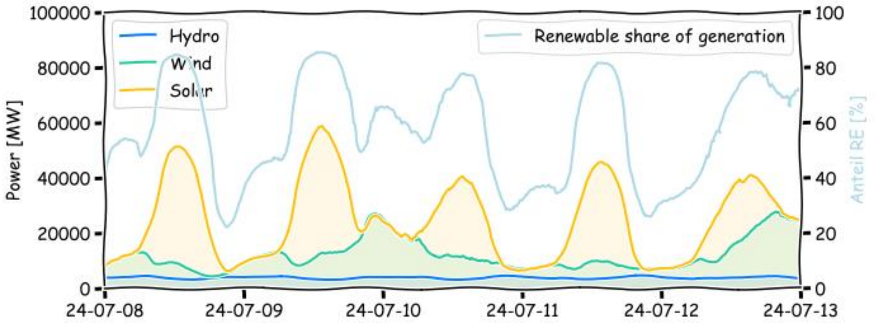
- Towards CO_2-neutral operation with local energy storage (for ET) (→ see talk by Achim Stahl)
  Evaluation of the potential to reach a fully CO_2-neutral operation based on local energy storage and local RE production
  Breathing centre powered by wind & solar energy plus use battery and hydrogen storage to cover the base load of the centre during periods of low availability of RE

Orchestration (lead &coordinated by KIT, M. Giffels; partners: BN, DESY, GÖ, FR)

Meta-scheduler COBalD/TARDIS used for years
for efficient use of distributed compute resources
(at KIT, BN, GÖ, FR, …. for integration of
BAF, CLAIX, Emmy, HoreKa, Nemo, …)

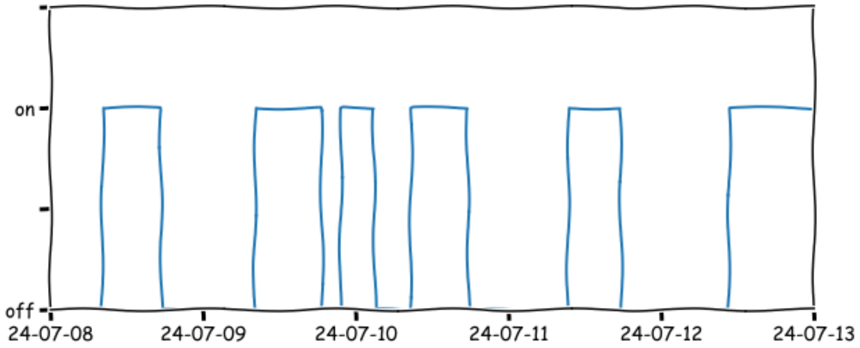


Extension of CObalD/TARDIS and its interfaces
to provide resources depending on the forecasted availability of RE
(from energy grid or local photovoltaic as prototype) → next slide



Development and Implementation of checkpoint-restore for compute jobs
need to stop, transfer and restore jobs from one site to another
(e.g. explore usage of CRIU with container technologies)

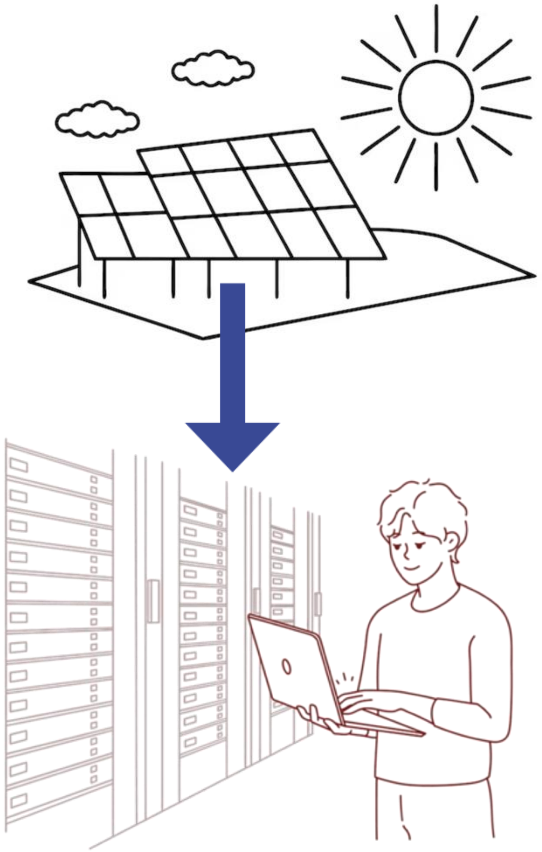
Implementation of dynamic resource limit in COBald/TARDIS
to optimize booking of resources vs. RE availability / compute request, …



Optimization of power efficiency upon tuning clock frequencies

- perform measurement of energy consumption and HS23 performance on clock frequency for various architectures at different sites
- development of strategy and system daemon for how to optimally adjust clock frequency to RE availability
- extend COBalD to provide dynamical clock frequency to AUDITOR

Breathing Compute Centre (BCC) / Sun Cluster (coordinated by KIT , M. Giffels, partners: BN, DESY, GÖ, FR)



Provide opportunistically resources depending on the forecasted availability of RE from photovoltaic at the sites itself (or via an exceeding RE from the electricity provider).

In the future: balancing aspects of the power grid by turning on and off huge power consumers on demand will gain more importance → additional use case of BCC

Steps:

- Extend COBalD/TARDIS to use external input from "Forecasting and Optimization" to dynamically scale compute resources depending on the green energy forecast.
- Define dedicated interface between „Forecast“ and COBalD balancing daemon for resource scheduling and its implementation into the COBalD balancing
- Identify interfaces for the resource life cycle management of bare metal machines and develop corresponding python interfaces (e.g. Redfish (Redfish Scalable Platforms Management API), RedHat Satellite System Management Platform
- Extend TARDIS to effectuate the scheduling decision from COBalD based upon "Forecast" service the by booting or draining/shutting down bare metal servers

500 kW peal @ KIT SCC
15+44 kW peak @ FR Physics
xyz? kW peak @ GÖ
xyz ?kW peak @ DESY

Perform tests

- to determine the potential impact of BCC on strategy for the hardware lifetime,
- to determine the upper bounds of floating servers
- to optimize turn-around times to allow smooth operation

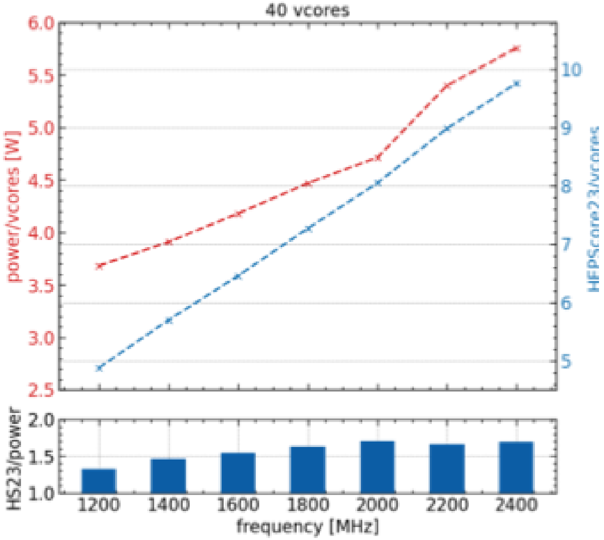
Accounting (lead and coordinated by Freiburg, M. Böhler; partners KIT, Öko-Institut)

Flexible accounting ecosystem AUDITOR in use at CERN, DESY, KIT, FR, WU, … and dissemination to all WLCG sites in progress

Extension of AUDITOR in various ways foreseen
with final goal to deliver a full accounting of the CO_2 footprint.

AUDITOR

New AUDITOR plugin is to be developed and implemented
to determine energy consumption and CO_2 footprint for compute jobs from the runtimes and used local energy mix (from interface to „Forecasting“)



Evaluation. adaption and extension of set of key performance indicators (KPI)

- developed in previous project by Öko-Institut on green computing
- newly defined for the specific processes relevant in data handling, job distribution and data analysis in the ErUM community

Extension of AUDITOR to collect data for

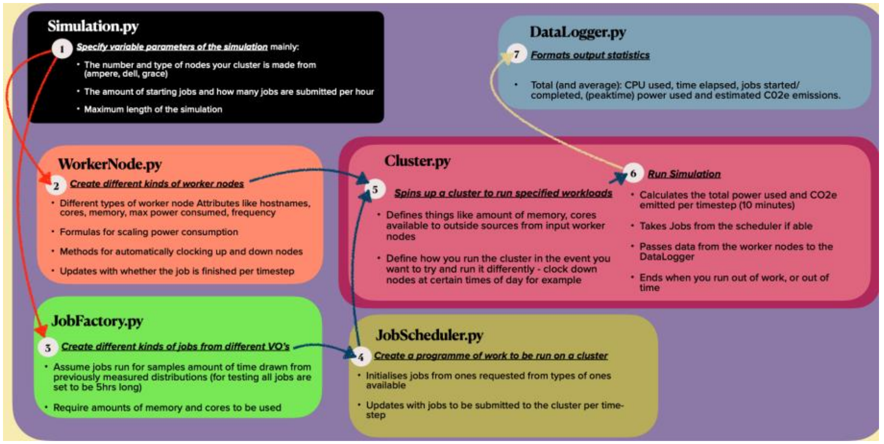
- data concerning water usage and abiotic resources depletion related to specific hardware components
- energy consumption and CO_2 equivalent due to data transfer and network
- embedded emissions in production and disposal

Development of optimization strategy and implementation of feedback loop
into „Forecasting“, “Orchestration“ and „Digital twin“

Extension for GPU accounting
Extension for time dependent clock rate
(Extension for analysis facilities)

Digital Twin **(lead and coordinated by DESY, D. Spiteri; partners BN)**

Simulation and optimization framework prototype
developed at DESY for single site
using as input information: site inventory (CPU type, nr. of cores, memory, …),
machine-dependent power specifications (idle, full load, freq. dep)
and carbon intensity of the grid, job mix, ….
delivering an estimate of the total power used and its CO_2 emissions
for different ways of how to perform a specified workload



Goal: Simulation of geographical federated IT infrastructure, which commits to pledges and help reducing energy impact, primarily by having the functionality to transfer jobs to different sites.

Steps:

- Partner data-centre integration: collect information about site inventory, metrics (PUE, HS23-per-core as funnction of clock frequency) of sites, local grid intensity at sites in common data base (in cooperation with data curation in „Forecasting“)
- Improved estimate of CO_2 footprint: include embedded CO_2, storage, network, … by adapting to databases/ accounting models developed in work packages „Forecasting“ and „accounting“
- Deployment of full simulation infrastructure for various sites including realistic job orchestration a la COBald/TARDIS including job transfers from one site to another when running
- Development and deployment of feedback-loops from/ to other work packages
- Validation and optimization: compare simulation to measured data and optimization of simulation hyperparameters

Education & dissemination, role of associated partners, …. (all)

Thorough documentation, development of tutorials and conduction of educational workshops
in order to foster new use cases and to disseminate the developed concepts and ecosystems to a wider audience
Workshops shall be advertised and organized in cooperation with the ErUM-Data Hub.

Feedback by and roll out to WLCG via S. Campana (WLCG project leader)

Connection link, alpha-users and multiplicators in respective community:
KAT community via K. Valerius (KIT) and M. Schumann (Freiburg)
KFS community via C. Gutt (Siegen) and KFB community via E. Bründermann (KIT)

Heads of the HPC centres at Freiburg (D. von Suchodoletz) and at the SCC at KIT (A. Streit)
provide feedback from a computer science and resource provider point of view

Links to other BMFTR-funded research networks such as "Federated storage FUSE", to PUNCH4NFDI, EOSC …
via individuals from (associated) partner institutes, who are active members (of boards) in this initiatives

The ErUM-Data Hub kindly agreed that it will provide support for dissemination of developed concepts and technologies and for conducting workshops and tutorials to train scientists in all ErUM communities.

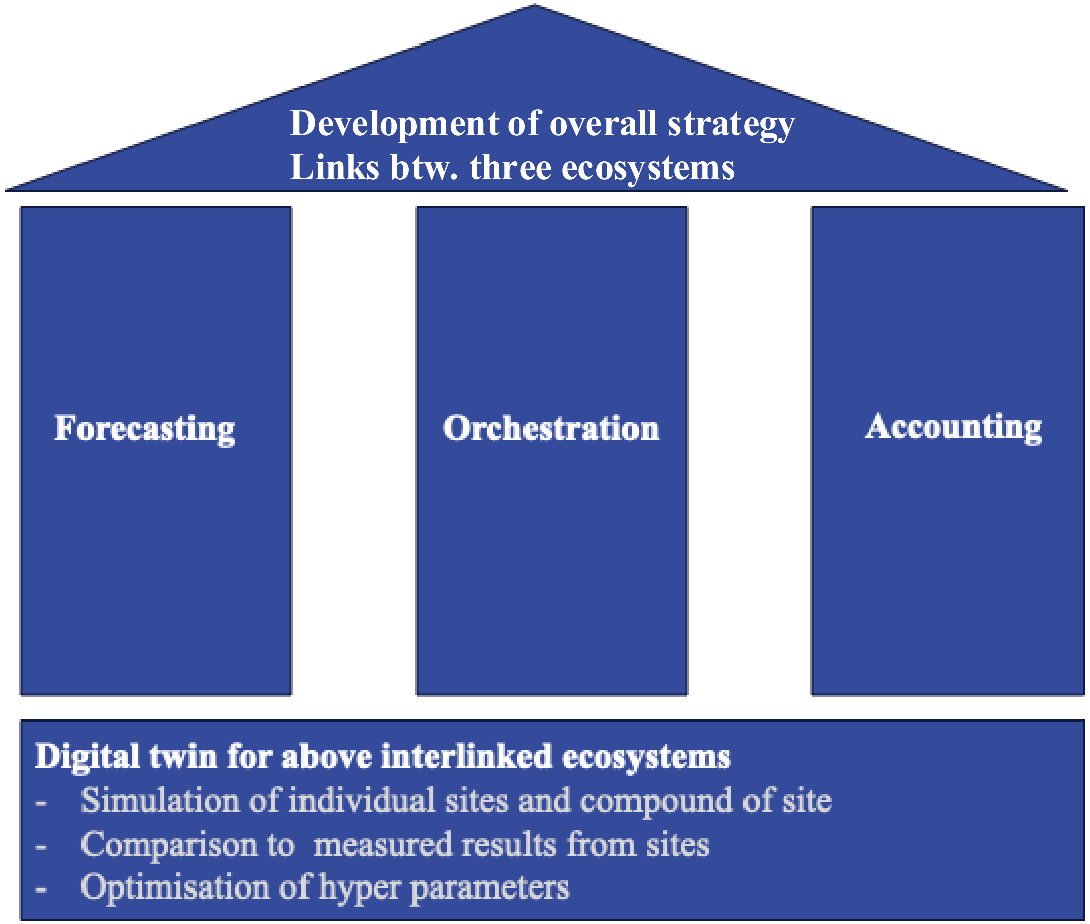
Conclusions

- 2024: electricity consumption of compute centres in D 20 billion kWh in 2024 (4% of total), 7.3 t CO_2 footprint
- Next 5 years: energy consumption expected to double ←→ goal of CO_2 neutrality in D in 2045

→ Advance environmental sustainability of scientific computing in SUFECIT project

SUSFECIT

Sustainable **Fe**derated **C**ompute **I**nfrastructures



Goals:

- increase the environmentally sustainable use and operation of federated compute infrastructures
- raise awareness of the impact of scientific computing on CO_2 footprint from ErUM communities
- train talented early career scientists in cutting-edge technologies

Hope:

- SUSFECIT will provide tools to help reach this goal and dissemination to ErUM communities and beyond will be successful