



Economical and Ecological Impact of Sector Coupling Applied to Computing Clusters

Philip Bechtle¹, Oliver Freyermuth¹, Maximilian Peter Geffers^{1,2}, Manuel Giffels³, Michael Hübner¹, Florian Kirfel^{1*} , Jochen Kreutz⁴, Stefan Krieg^{4,5}, Sebastian Matberg¹ and Matthias Schnepf³

*Correspondence:
fkirfel@uni-bonn.de

¹ Physikalisches Institut,
University of Bonn, Bonn,
Germany

² Universität Heidelberg,
Heidelberg, Germany

³ Karlsruhe Institute
of Technology, Karlsruhe,
Germany

⁴ Jülich Supercomputing Centre,
Forschungszentrum Jülich,
Jülich, Germany

⁵ Helmholtz-Institut für Strahlen-
und Kernphysik, University
of Bonn, Bonn, Germany

Abstract

The rising share of abundant renewable energy inevitably increases volatility in the electricity production. The concept of sector coupling means that the volatility of electricity production to a large degree can be absorbed by dispatching electricity consumption whenever excess renewable energy is available. A system that is dynamically operated based on this principle can lower its total environmental impact. In addition, operational costs might be reducible as electricity prizes strongly depend on the residual load of the energy system. High-performance computing clusters in the field of science represent an ideal testing ground for such dynamic operation. Short-term delays in computing results due to electricity production being associated with high costs or carbon emissions are often negligible, provided that an overall computing target remains constant over long time periods. This study simulates the simplified operation of computing clusters using publicly available data on electricity production in Germany. The optimal utilisation along with associated carbon emission and cost reductions are determined separately. Hardware acquisition costs and embedded emissions are taken into account. The stability of a fixed computing target given the determined utilisation optima is evaluated in two validation periods. Additional simulations with modified parameters are carried out to estimate potential conditions under which dynamic operation of a computing cluster would continue to enable savings in the future.

Keywords: Environmental impact, high-performance computing (HPC), high-energy physics, sector coupling

Introduction

The global warming induced by carbon emissions, first predicted in 1896 [1], is expected to cause a global increase in average temperatures of more than 1.5 °C by 2030 compared to pre-industrial levels [2]. The consequences of this trend are already evident across Europe in the form of more frequent and more intense extreme weather events. In Germany alone, the annual costs of recorded extreme-weather-related damages exceeds six billion euros during the period from 2015 to 2024 [3].

To at least mitigate these effects in the future, Germany, originally along with 195 other nations, committed itself to the Paris climate agreement with the goal of limiting global warming to well below 2 °C above pre-industrial levels [4]. This requires a

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

drastic reduction in greenhouse gas emissions, primarily achievable through lowering the carbon intensity of the German electricity mix, which amounted to an average of 344 kgCO₂/MWh in 2025 [5], by switching to renewable energy sources. In 2025, the share of renewable energies in Germany electricity consumption was already at 55.1 % [6]. However, its share in the final total energy consumption only amounted to 23.8 % [6] and was thus significantly lower. This discrepancy can be attributed to sectors, such as transport and heating, both of which still rely heavily on fossil fuels.

Naturally, attempting to electrify these areas as well requires more energy to be produced by renewable energy sources, i.e. volatile wind and solar power will become the dominant power sources. Ultimately this increases fluctuations in the energy grid. To control such fluctuations, a strong sector coupling [7] is required, capable of leveraging the flexibility of various sectors. In contrast to large industrial plants, which, to be profitable, must often operate continuously, the focus for sector coupling lies on systems that can be dynamically regulated. An almost perfect use case for this desired dynamic regularisation are computing clusters, as their workload can be flexibly controlled by shifting computing tasks over time. As of 2025 data centres already account for 4 % of Germany's gross electricity consumption [8] and, given the recent developments in the field of artificial intelligence, will account for an ever-increasing share of consumed power. Forecasts assume that, by 2030, computing might be responsible for up to 8 % of the total power consumption [9].

Shifting computing tasks to periods with an oversupply of renewable energies can help at stabilising the energy grid. At the same time, procuring electricity from the spot market offers significant savings in terms of carbon emissions while also enabling economically optimised operation of computing clusters. As an example, negative wholesale energy prices on the German electricity spot market were reported for a total of 573 h in 2025 [10], primarily at weekends or on public holidays during which renewable energy sources led to an oversupply. Targeted ramp-up of a computing cluster at such periods, therefore, offers potential cost savings.

This dynamic operation of computing clusters inevitably leads to delays in computing results on short time scales. However, this is often negligible, especially in the field of science, making computing clusters in research institutions and universities the ideal place to test such an application. The following studies are, therefore, carried out by relying on different computing clusters operated by research facilities. The main focus is placed on the BAF high throughput computing cluster [11], operated at the Physics Institute of the University of Bonn. This cluster is primarily used for computing tasks in the field of high-energy physics, which, with more than 20 % [12], has the greatest demand for computing power in the German scientific sector and thus also represents the largest lever for dynamic control.

Initially, in Sect. “[Background](#)” various options for reducing greenhouse gas emissions in the area of computing clusters are discussed. Particular emphasis is placed on the feasibility of the respective countermeasures on a short time scale. Building on this, Sect. “[Methodology](#)” introduces the simulation of a dynamic cluster operation for data centres by relying on public data on Germany's electricity generation in recent years. Finally, Sect. “[Simulation Results](#)” presents results for concrete case studies based on this simulation.

Background

In general, computing clusters account for a significant share of total emissions in modern research. In particular, large high-energy physics experiments almost all rely on computing clusters. A study evaluating the professional carbon footprint of doctoral researchers at one of the four major LHC experiments shows that computing with 1.91 tCO₂ accounts for 9.3 % of the experiment's total emissions [13].

Fortunately, computing clusters offer many possibilities to reduce their carbon footprint either directly or indirectly. The first key decision for any newly built computing cluster involves selecting an adequate location. Besides the clearly important access to power supply, the local climate also plays a significant role. Depending on a cluster's architecture, current outside temperatures in Germany allow operation without active cooling [14]. This leads to a significant reduction of carbon emissions associated with the cluster operation. The heat, inevitably generated during operation, can be fed into the local heat networks of neighbouring towns. However, the temperature gradient between existing heat network's operating temperatures and the waste heat generated often poses a problem [14].

Regarding a computing cluster's installed hardware, using more modern components usually goes hand in hand with lower carbon emissions. This is either achieved through the components being more energy efficient or simply due to their increased performance, which shortens the overall runtime of computing tasks. Obviously, the embedded emissions associated with producing and transporting hardware have to be factored in.

Ultimately, training end users also offers a great indirect potential for savings. Every end user initially must learn to correctly estimate the cluster resources required for their own computing tasks, e.g. the number of CPUs needed to efficiently complete a certain task. Furthermore, the correct use of an appropriate programming language is crucial to make optimal use of the available hardware. In large scientific collaborations in particular, further savings can be achieved by adopting the FAIR data principle [15]. This principle intends to ensure that data and software are easily findable, accessible, interoperable and maybe most importantly reusable. Such a reusability of results helps at eliminating redundancies and thus lowers unnecessary emissions associated with computationally intensive tasks.

A very comprehensive overview of potential energy and carbon emission savings for computing clusters can be found in Refs. [14, 16]. In order to operate a computing cluster that complies with increasingly stringent emission regulations in Germany [17], progress is needed across all these aspects. Whether by modifying clusters to utilise waste heat, modernising hardware or providing intensive training for end users, progressing toward a carbon-neutral operation inevitably involves considerable effort and expenses.

However, these measures vary significantly in terms of the effort required for implementation. Particularly when targeting short- and medium-term emission reductions, the dynamic operation of a cluster, as introduced in Sect. "Introduction", can represent a promising alternative. Particularly when assuming that an existing cluster is not operating at full capacity, enabling dynamic operation involves significantly less work than, for example, completely rebuilding a cluster in order to benefit of modern cooling technology and hardware.

Given that the installed hardware still has spare capacity, a cluster generally needs to be equipped with a control software to implement dynamic operation. Based on publicly available weather data and information on energy generation, this software can automatically decide whether to accept or continue computing tasks at a given time, or whether the cluster should switch to an idle state and resume work when conditions are more favourable. Such a cutoff in terms of carbon emissions per MWh up to which a operation is deemed favourable could offer significant short-term potential in terms of carbon emission and operation cost reduction. In addition, given some flexibility with respect to key parameters, once implemented such a control software could be re-used for different clusters.

Methodology

The conducted simulation¹ is intended to assess the optimal operation of a computing cluster in the field of high-energy physics. Both the minimisation of carbon emissions and reductions with respect to a computing cluster's operating costs are being targeted. For the sake of simplicity, it is assumed that the cluster is entirely powered by spot market-traded electricity. The data, covering both the cost of consumed electricity in €/MWh as well as the net electricity generation in MWh, was obtained using the available energy data in Germany, publicly provided by the Fraunhofer Institute for Solar Energy Systems ISE² [18]. Their database collects net electricity production broken down by generation type along with the spot market prices for time intervals of 15 min. The corresponding carbon emissions associated with each generation type per year can be accessed separately via ref. [19], allowing the carbon emissions to be calculated in kgCO₂/MWh for each 15 min interval. The resulting price and carbon emission development over a 1-year period are shown exemplary for the calendar year 2024 in Fig. 1a, b.

As already described in Sect. “Background”, a short-term delay when processing computing tasks is acceptable during time periods where electricity is either expensive or associated with high carbon emissions. Nevertheless, the total available computing power for research should remain approximately constant over a longer period of time. This implies that a cluster which, due to high carbon emissions, can only be operated for 50 % of a given time period must, assuming linear scalability, be designed twice as large. This way a constant compute target can be maintained. For the sake of simplicity, a homogeneous cluster consisting of identical CPUs was simulated in this study. This allows the variable cluster size, depending on the possible cluster utilisation to be modelled directly via the number of logical cores of all combined CPUs. Conversely, the estimated embedded carbon emission for every single logical core has to factor in all components required to expand the computing clusters.

Electricity consumption and thus also the operating emissions for day-to-day operations were calculated based on the computing cluster's estimated electricity consumption and the aforementioned data regarding public net electricity production in Germany. Similar to the embedded carbon emissions, the total power consumption of

¹ The complete simulation workspace can be found under https://gitlab.uni-bonn.de/grp_phy/drittmittelprojekte/erum-data/susfecit/eco-clusters-study.

² The website can be found under <https://energy-charts.info>.

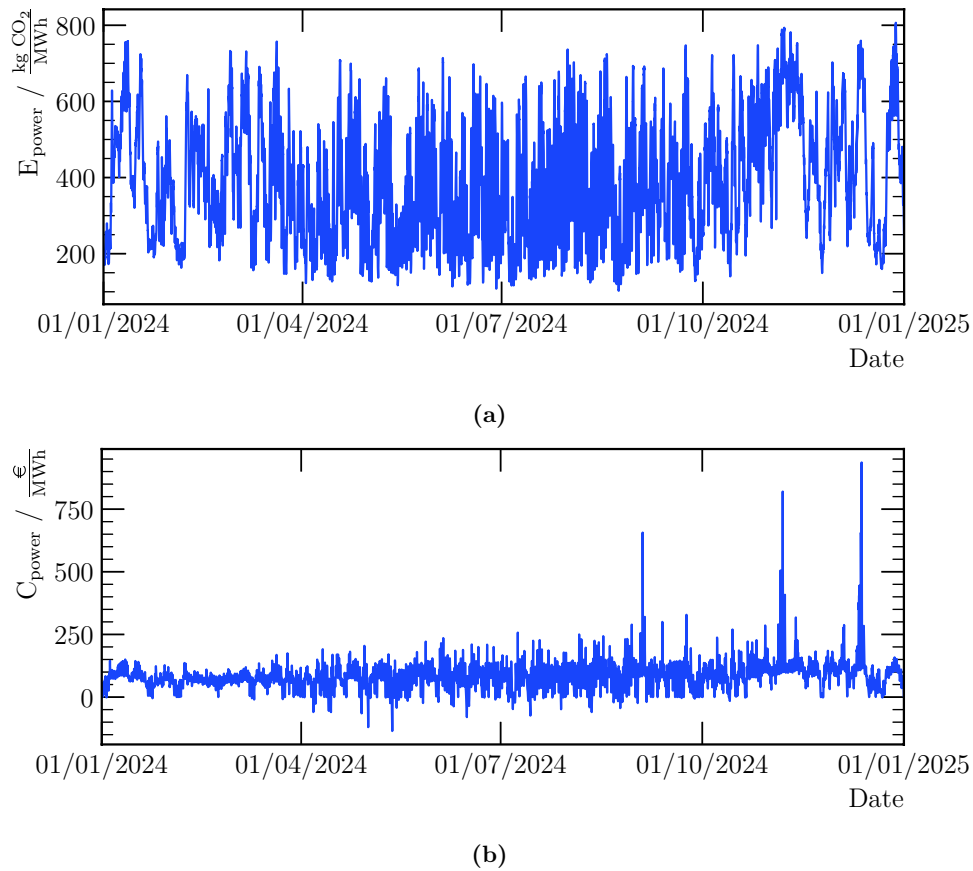


Fig. 1 Development of **a** carbon emissions in kgCO₂/MWh and **b** costs in €/MWh arising from German net electricity generation for the calendar year 2024

the cluster was calculated based on the scaling factor, i.e., the number of logical cores. To allow for a more realistic simulation, different workload scenarios were specified. Each scenario is defined via the respective partial load power consumptions and the relative time fractions during which the computing cluster is operated at each load. This is particularly relevant to avoid neglecting the power consumption of a computing cluster during idle phases. To enable rapid start-up under more favourable conditions and limit hardware damage, complete shutdown was not investigated. Therefore, even in the case of high emissions associated with the net electricity production, i.e. periods in which operation is generally avoided, a non-zero electricity consumption was assumed.

Under these assumptions, the total carbon emissions $E_{\text{total}}(u)$ are calculated through

$$\begin{aligned}
 E_{\text{total}}(u) = & \underbrace{\frac{n_{\text{cores}}}{u} \cdot E_{\text{embedded, CPU hour}} \cdot t_{\text{total}}}_{E_{\text{embedded}}} + \underbrace{\frac{n_{\text{cores}}}{u} \cdot P_{\text{avg.}} \cdot \sum_i (E_{\text{power},i} \cdot t)}_{E_{\text{operation}}} \\
 & + \underbrace{\frac{n_{\text{cores}}}{u} \cdot P_{\text{idle}} \cdot \sum_j (E_{\text{power},j} \cdot t)}_{E_{\text{idle}}},
 \end{aligned}$$

where n_{cores} denotes the number of logical cores, u the relative utilisation within the considered period t_{total} , and t the respective sampling time interval. The carbon emissions of electricity generation in kgCO_2/MWh for each time interval are represented by E_{power} . The index i runs over all intervals in which E_{power} is below the utilisation threshold value X_{emission} . In contrast, the index j covers the time intervals in which the cluster remains in an idle state as E_{power} exceeds X_{emission} . The parameters $P_{\text{avg.}}$ and P_{idle} represent the workload scenario dependent average power consumption per logical core during operation phases and idle phases, respectively. An overview of all parameters used in the simulation along with their definition is given in Table 1.

The optimal threshold value for the emissions associated with the net power generation can be found by minimising $E_{\text{total}}(u)$ as a function of the relative cluster utilisation u . This minimisation is achieved numerically by varying the threshold between the minimum and maximum values of E_{power} . For each threshold value, the corresponding utilisation and thus ultimately the total emissions $E_{\text{total}}(u)$ are derived. Any time the net power production exceeds the optimal threshold associated with the minimum of

Table 1 Simulation parameters and their definitions

Parameter	Unit	Description
n_{cores}	–	Default number of logical cores in a computing cluster
u	–	Relative operating time of the cluster throughout the simulated period ($0 \leq u \leq 1$)
$E_{\text{total}}(u)$	kgCO_2	Total carbon emission throughout the simulated period
$E_{\text{embedded, CPU hour}}$	kgCO_2/h	Hourly embedded emissions per logical core, based on an assumed lifetime of 10 years for all components
E_{embedded}	kgCO_2	Total embedded emission throughout the simulated period
E_{power}	kgCO_2/MWh	Carbon emissions for electricity generated in Germany at each sampled time interval
$E_{\text{operation}}$	kgCO_2	Total carbon emissions from power consumption during cluster operation throughout the simulated period
E_{idle}	kgCO_2	Total carbon emissions from power consumption during idling throughout the simulated period
$C_{\text{total}}(u)$	€	Total cost throughout the simulated period
$C_{\text{acq., CPU hour}}$	€/h	Hourly acquisition cost per logical core, based on an assumed lifetime of 10 years for all components
$C_{\text{acq.}}$	€	Total acquisition cost throughout the simulated period
$C_{\text{yearly demand}}$	€/kW/year	Electricity supplier's capacity-based charge per year
C_{demand}	€	Total demand cost throughout the simulated period
C_{power}	€/MWh	Spot market energy price for electricity generated in Germany at each sampled time interval
$C_{\text{operation}}$	€	Total cost from power consumption during cluster operation in the simulated period
C_{idle}	€	Total cost from power consumption during the cluster's idle state in the simulated period
$P_{\text{avg.}}$	MW	Workload-dependent average power consumption per logical core
P_{max}	MW	Maximum power consumption per logical core
P_{idle}	MW	Power consumption per logical core when in idle mode
t_{total}	h	Total time period used in the simulation
t	h	Sampling interval duration
X_{emission}	kgCO_2/MWh	Emission threshold above which the computing cluster remains in idle state
X_{cost}	€/MWh	Cost threshold above which the computing cluster remains in idle state

$E_{\text{total}}(u)$ it is more favourable to shut down the computing cluster, i.e., remain in idle mode. A minimum far from $u = 1$ implies, that the computing cluster must be extended to fulfil the constraint arising from the constant compute target.

Similarly, an optimisation was carried out with regard to the financial costs of operating a computing cluster. As the constant compute target was again modelled via the number of logical cores n_{cores} , the price for the purchase of a single core $C_{\text{acq.}}$ must reflect all additional acquisition costs for the cluster expansion. Once this price is defined, the total costs $C_{\text{total}}(u)$ can be calculated via

$$C_{\text{total}}(u) = \underbrace{\frac{n_{\text{cores}}}{u} \cdot C_{\text{acq., CPU hour}} \cdot t_{\text{total}}}_{C_{\text{acq.}}} + \underbrace{\frac{n_{\text{cores}}}{u} \cdot P_{\text{max}} \cdot C_{\text{yearly demand}} \cdot t_{\text{total}}}_{C_{\text{demand}}} + \underbrace{\frac{n_{\text{cores}}}{u} \cdot P_{\text{avg.}} \cdot \sum_i (C_{\text{power},i} \cdot t)}_{C_{\text{operation}}} + \underbrace{\frac{n_{\text{cores}}}{u} \cdot P_{\text{idle}} \cdot \sum_j (C_{\text{power},j} \cdot t)}_{C_{\text{idle}}}.$$

Here, P_{max} denotes the maximum power consumption per core, which, when multiplied by the number of cores and the electricity provider's yearly demand charge, constitutes the second term of the total cost. The operating costs can be determined using the price per power produced at each sampling interval C_{power} and the average power consumption $P_{\text{avg.}}$ of a defined workload scenario. Electricity costs during idle periods are determined via the idle power consumption P_{idle} . The used indices are defined analogously to the emissions calculation, i.e., i again covers all intervals in which C_{power} lies below a derived threshold value X_{cost} in €/MWh, while j runs over all intervals exceeding this threshold. The base charge of any electricity contract was disregarded in this model, as it would only add an offset to the overall function. The minimum of $C_{\text{total}}(u)$ is determined analogously to $E_{\text{total}}(u)$.

A definition of all parameters used in the financial cost simulation can be found in Table 1.

In order to evaluate the determined emission threshold's stability over a long period, the optimal threshold was extrapolated to a test period covering the calendar years 2023 and 2025 for specific scenarios. To allow for a realistic comparison, the extrapolated threshold $X_{\text{emission}}^{\text{extra.}}$ factors in the changed share of renewable energy in Germany. This changed share's impact on the emissions was derived via data on the public energy production's average carbon emission per year and the relative share of renewable energies, again provided by the Fraunhofer ISE platform [20]. Assuming a linear relationship over the years from 2002 to 2025, an increase of 1 % in the share of renewable energies was found to correspond to an average emission reduction of 5.4(3) kg. Thus, the extrapolated threshold is given through

$$X_{\text{emission}}^{\text{extra.}} = X_{\text{emission}} \cdot 5.4(3) \text{ kg} \cdot (s_{2024} - s_{\text{val.}}),$$

where s_{2024} and $s_{\text{val.}}$ correspond to the share of renewable energies in the year 2024 and the chosen validation year, respectively. In both years, this extrapolated threshold $X_{\text{emission}}^{\text{extra.}}$ was compared to thresholds explicitly derived for each validation year $X_{\text{emission}}^{\text{val.}}$ by again minimising $E_{\text{total}}(u)$.

Simulation Results

The dynamic operation of computing clusters was simulated based on the 2024 data. The optimisation with regard to the total carbon emissions, discussed in Sect. “[Optimisation with Respect to Carbon Emission](#)”, was carried out for the following five different computing clusters setups:

- BAF_{default}: Default worker nodes of the Bonn Analysis Facility (BAF) high throughput cluster installed at the Physics Institute at the University of Bonn.
- BAF_{modern}: Modern worker nodes of the BAF high throughput cluster installed at the Physics Institute at the University of Bonn.
- DEEP_{CM}: Part of the DEEP modular supercomputer prototype [21] at the Jülich Supercomputing Centre (JSC) [22]. Designed for high single-thread performance and a modest amount of memory.
- DEEP_{DAM}: Part of the DEEP modular supercomputer prototype at the JSC. Designed for data-intensive analytics and machine learning applications. Accelerators are ignored in the simulation.
- GridKa_{ARM}: ARM worker nodes in the WLCG’s Tier 1 Grid Computing Centre Karlsruhe (GridKa) [23].

It is important to note, that the tested cluster setups do not exactly correspond to their real-life counterparts. Instead, a simplified setup, neglecting heterogeneous hardware configurations and accelerators was simulated. The optimisation in terms of total operating costs, discussed in Sect. “[Optimisation with Respect to Operational Costs](#)”, was only carried out for the two BAF cluster scenarios.

Optimisation with Respect to Carbon Emission

An overview of the parameters used for emissions optimisation across all the different scenarios can be found in Table 2. The first difference between the setups concerns the maximum power consumption per logical core $P_{\max}$, which was determined via full load tests. The resulting values naturally depend on the hardware used as well as the specific cluster settings, and vary between 9.2 W (BAF_{default}) and 2.9 W (GridKa_{ARM}). The associated idle values P_{idle} range between 15 % (BAF_{modern}) and 48 % (DEEP_{DAM}) of the respective maximum power consumption.

While the power consumption of each individual setup was directly measured, no precise data were available for the embedded emissions of the installed hardware. Therefore, $E_{\text{embedded, CPU hour}}$ in all five setups was estimated using a life cycle assessment of

Table 2 Parameter values chosen for the simulation of all five cluster scenarios

Parameter	Setup				
	BAF _{default}	BAF _{modern}	DEEP _{CM}	DEEP _{DAM}	GridKa _{ARM}
n_{cores}	7104	2816	2400	1536	2816
$P_{\max} / \text{W}$	9.2	7.6	7.8	6.9	2.9
$P_{\text{idle}} / \text{W}$	2.3	1.1	2.6	3.3	0.9
$E_{\text{embedded}} / \text{kgCO}_2/\text{h}$	1.5×10^{-5}	1.8×10^{-4}	1.8×10^{-4}	1.8×10^{-4}	1.8×10^{-4}

the Dell R740 rack server conducted by the thinkstep AG [24]. The study lists embedded emissions of 4092 kgCO₂ per rack server, taking into account manufacturing costs (4288 kgCO₂), transport costs (3 kgCO₂) and end of life credits (−199 kgCO₂). With approximately 80.3 % the dominant share of manufacturing emissions is accounted for by the SSD production. To express the embedded emissions per logical core, it was assumed that 256 logical cores can be installed per rack server. Assuming a life cycle of 10 years, this resulted in $E_{\text{embedded, CPU hour}} = 1.8 \times 10^{-4}$ kgCO₂/h for all but the BAF_{default} setup.

In contrast to the others, the BAF_{default} setup, solely relies on HDDs as storage media, for which significantly lower embedded emissions had to be assumed. As a reference, the study conducted in [25], which compared life cycle assessments from various manufacturers to derive an average storage emission factor SEF of 0.16 kgCO₂/GB for SSDs and 0.02 kgCO₂/GB for HDDs, was used.

Applying these storage emission factors, the BAF_{default} setup's embedded emissions amounted to 1.5×10^{-5} kgCO₂/h. It should be noted at this point that the embedded emissions evaluated in ref. [25] displayed significant fluctuations, especially when comparing different manufacturers. Thus, the exact value of E_{embedded} is highly dependent on the chosen setup and the installed hardware. All values chosen here are only rough estimations and assumed to be associated with considerable uncertainties.

In order to simulate realistic use cases of a computing cluster, different workload scenarios were considered for all setups. Besides a medium workload and a heavy workload scenario, which are oriented toward studies in ref. [24], a backfilling [26] scenario was defined. Such a scenario is frequently used in high-energy physics and allows free resources to be allocated more efficiently by prioritising short single-core jobs. As a result, the cluster can be assumed to run under full load during the simulated period. The exact parameters of all three scenarios are summarised in Table 3. For each load working point, it was assumed that a cluster's power consumption scales linearly between P_{idle} and P_{max} .

The optimisation results of the five setups are summarised in Table 4 for all workload scenarios. In addition, $E_{\text{total}}(u)$ alongside the individual contributions is displayed in Fig. 2 as a function of the utilisation u in case of BAF_{modern}'s backfilling scenario and illustrates the expected trend: $E_{\text{operation}}$ decreases along with the utilisation, as the cluster uses electricity with increasingly lower levels of associated carbon emissions. At the same time, E_{idle} and E_{embedded} along with the number of logical cores, i.e., the cluster scaling factor, increase to maintain a constant compute target.

A comparison across the five setups generally demonstrates that emission decrease through dynamic operation is more pronounced for workloads with high cluster utilisation. In addition, it is apparent that the emission reduction strongly depends on

Table 3 Workload scenarios used in the simulation

Load mode/%	0	10	50	100
Medium workload time fraction/%	25	30	35	10
Heavy workload time fraction/%	10	20	55	15
Backfilling workload time fraction/%	5	0	0	95

Setup dependent power consumption is assumed to scale linearly between P_{idle} and P_{max}

Table 4 Optimisation results to minimise carbon emissions for all setups and workload scenarios

Workload scenario	$u_{\text{opt.}}$	$X_{\text{emission}}/\text{kgCO}_2/\text{MWh}$	$E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$
BAF _{default}			
Medium	1.000	–	1.000
Heavy	0.975	699.90	0.998
Backfilling	0.745	515.01	0.955
BAF _{modern}			
Medium	0.983	714.13	0.999
Heavy	0.882	607.45	0.988
Backfilling	0.635	458.11	0.918
DEEP _{CM}			
Medium	1.000	–	1.000
Heavy	1.000	–	1.000
Backfilling	0.916	635.90	0.993
DEEP _{DAM}			
Medium	1.000	–	1.000
Heavy	1.000	–	1.000
Backfilling	1.000	–	1.000
GridKa _{ARM}			
Medium	1.000	–	1.000
Heavy	1.000	–	1.000
Backfilling	0.953	671.458	0.997

Listed is the optimal utilisation $u_{\text{opt.}}$, the associated threshold X_{emission} above which a computing cluster switches to idle state, and the relative emissions compared to constant operation $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$. In case that a constant operation yields the lowest total carbon emissions, no threshold on E_{power} is specified

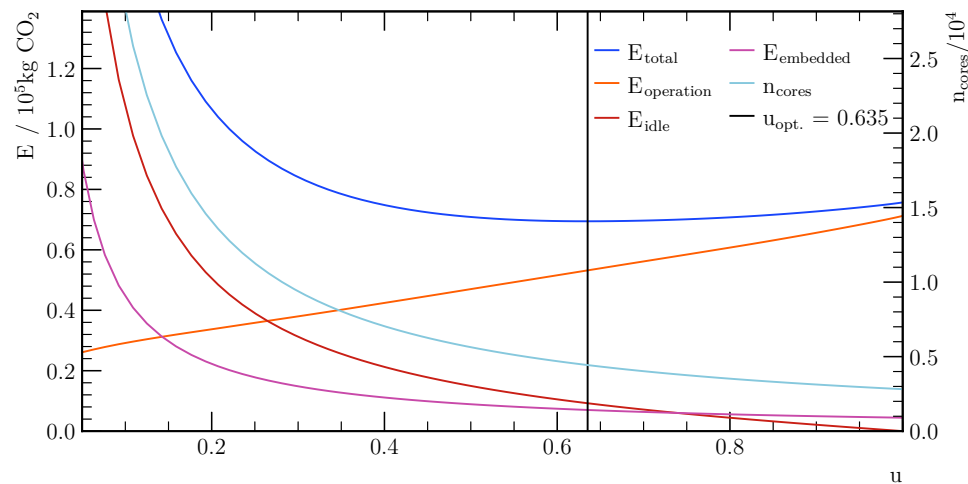


Fig. 2 Display of total emission alongside the individual contributions (left y-axis) and number of logical cores needed to maintain the default compute target (right y-axis) as a function of the utilisation u for the BAF_{modern}'s backfilling scenario. The optimal cluster utilisation is marked by the black vertical line. While $E_{\text{operation}}$ increases with rising utilisation, E_{idle} and E_{embedded} decrease. The latter results from reduced hardware requirements to meet the constant computing target

the ratio between P_{idle} and P_{max} . If power consumption is not reduced significantly when idling, the benefits of a dynamic computing cluster operation are negated. Consequently, the greatest potential for savings is found for the BAF_{modern} setup when

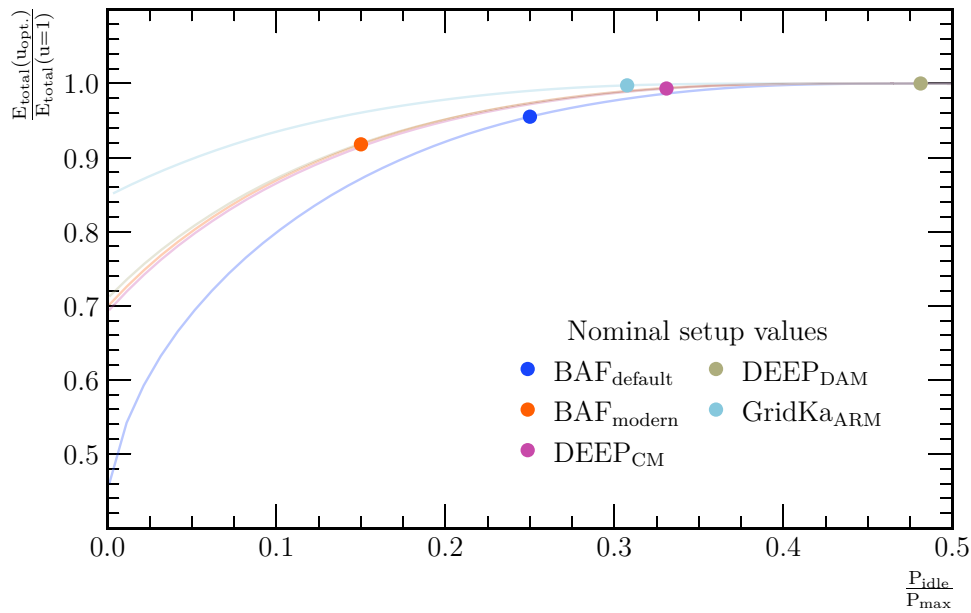


Fig. 3 Relative emissions compared to constant operation $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$ as a function of the power consumption ratio $P_{\text{idle}}/P_{\text{max}}$ for all setups in the backfilling scenario. Nominal power consumption ratios are marked

simulated for the backfilling workload scenario. At a utilisation of $u = 0.635$, carbon emissions can be reduced by around 8 % compared to $u = 1$. In contrast, the DEEP_{DAM} setup offers no potential for savings. Its relatively high power consumption in idle mode ($P_{\text{idle}}/P_{\text{max}} \approx 48\%$) results in a trivial minimum of $u = 1$.

The actual idle power consumption is not only hardware-dependent but can also be adjusted by altering hardware configuration.³ To illustrate potential variability and assess its effect, Fig. 3 shows the relative emission reduction as a function of the relative idle power consumption $P_{\text{idle}}/P_{\text{max}}$. For better visualisation, only the backfilling scenario is displayed.

The strong dependency of the emission reduction on the ratio is evident. Potential savings compared to a continuous operation decrease rapidly as the relative idle power consumption increases. Beyond a threshold of $P_{\text{idle}}/P_{\text{max}} \approx 0.4$, dynamic operation no longer offers any advantages regardless of the simulated setup. In contrast, assuming a cluster can be completely shut down instead of idling ($P_{\text{idle}}/P_{\text{max}} = 0$), savings of more than 50 % could be achieved depending on the setup.

To further estimate which effect the uncertainty with regard to embedded emissions has on the overall optimisation, the relative savings were derived as a function of embedded emissions. This test was only performed for the BAF_{modern}'s backfilling scenario. All other simulation parameters were kept unchanged. The results for an up- and down-variation of 50 % with respect to the nominal value of $1.8 \times 10^{-4} \text{kgCO}_2/\text{h}$ are listed in Table 5. The results show the expected trend: increased embedded emissions shift the minimum closer to $u = 1$ and vice versa. Even for the up-variation, the calculated

³ For example, BIOS energy settings.

Table 5 Optimisation results to minimise carbon emissions for the BAF_{modern} setup's backfilling scenarios

Workload scenario	$u_{\text{opt.}}$	$X_{\text{emission}}/\text{kgCO}_2/\text{MWh}$	$E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$
BAF _{modern}			
Backfilling	0.635	458.11	0.918
Backfilling ($E_{\text{embedded}} \times 0.5$)	0.594	436.78	0.896
Backfilling ($E_{\text{embedded}} \times 1.5$)	0.723	500.78	0.949

Listed is the optimal utilisation $u_{\text{opt.}}$, the associated threshold X_{emission} beyond which a computing cluster switches to idle state, and the relative emissions compared to constant operation $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$. In addition to the nominal value, the results are shown for a 50 % up- and down-variation of the embedded emissions

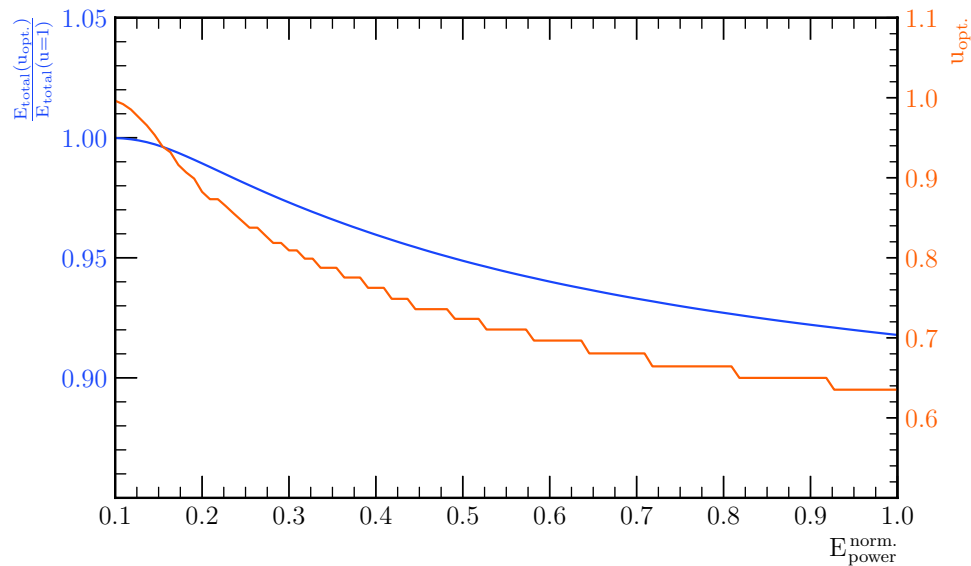


Fig. 4 Relative emissions compared to constant operation $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$ (left y-axis) and optimal utilisation $u_{\text{opt.}}$ (right y-axis) as a function of the normalised carbon emission associated with the electricity generated in Germany $E_{\text{total}}^{\text{norm}}$. Here, $E_{\text{total}}^{\text{norm}} = 1.0$ corresponds to the actual emission in 2024, while $E_{\text{total}}^{\text{norm}} = 0.1$ corresponds to a reduction by 90 %

optimum corresponds to a utilisation smaller than $u = 1$. The associated relative emission reduction lies at approximately 5 %. In the tested setup, dynamic operation of a cluster enables a carbon emission reduction across a fairly wide range of E_{embedded} . This is consistent with the results shown in Fig. 2 which already illustrated that the embedded emissions do not represent the optimisation's dominant factor.

Finally, a study was conducted to estimate how an increase of renewable energy sources, i.e. lower carbon emissions associated with electricity generation, would affect the saving potential of a dynamic computing cluster operation. As with the variation in embedded emissions, this test was carried out only for the BAF_{modern}'s backfilling scenario. To cover a wide range, E_{power} was reduced by up to 90 % in the simulation. The results regarding the relative emission reduction and the corresponding optimal utilisation $u_{\text{opt.}}$ are shown in Fig. 4. As expected, reducing E_{power} causes $u_{\text{opt.}}$ to increase and consequently the relative emissions $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$ approximate 1. However, even when reducing E_{power} down to 20 % of its original value ($E_{\text{total}}^{\text{norm}} = 0.2$), dynamic

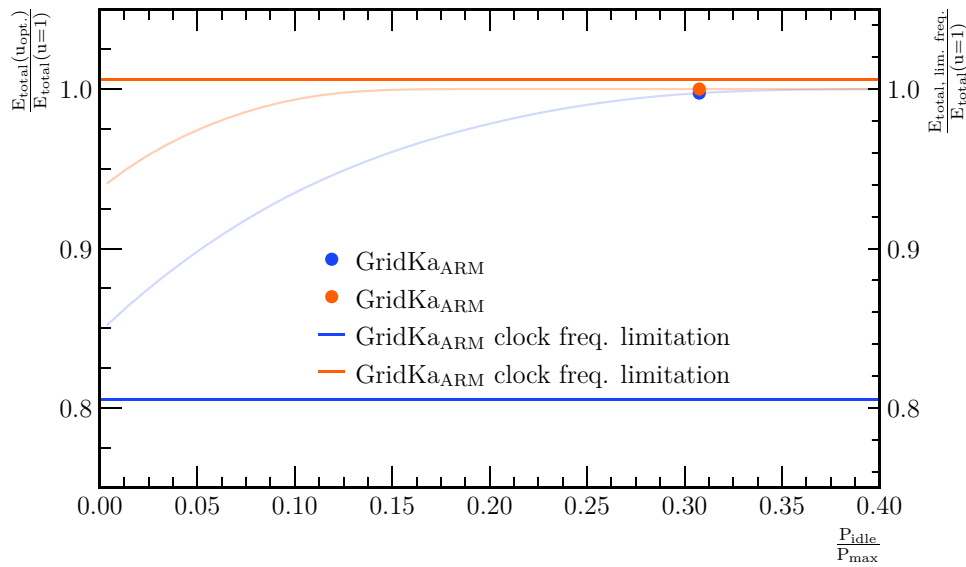


Fig. 5 Relative emissions compared to constant operation $E_{\text{total}}(u_{\text{opt.}})/E_{\text{total}}(u = 1)$ (left y-axis) as a function of the power consumption ratio $P_{\text{idle}}/P_{\text{max}}$ for the GridKa_{ARM} setup. The nominal power consumption ratio is marked in blue (orange) for the backfilling (medium) scenario. The blue (orange) horizontal line indicates the relative emissions $E_{\text{total, lim. freq.}}/E_{\text{total}}(u = 1)$ (right y-axis) for the backfilling (medium) scenario, assuming continuous utilisation of the cluster at a limited clock frequency. Due to the continuous utilisation varying the power consumption ratio $P_{\text{idle}}/P_{\text{max}}$ has no effect for those scenarios

operation still allows for emission reduction. Thus, the dynamic operation of a computing cluster continues to offer advantages even when assuming way higher shares of renewable energy sources.

Dynamic Operation and Clock Frequency Limitation: Comparison

The GridKa_{ARM} setup allows for an additional comparison with an alternative cluster operation mode, based on studies conducted in ref. [27] regarding the installed hardware's power consumption. In this study, the CPUs clock frequency was scanned to find the maximum power efficiency. The latter is defined as the ratio of CPU performance to the average power consumption. In this case, CPU performance was quantified using the HS23 score⁴ [29]. At the maximum derived in ref. [27], power consumption was reduced by around 40 %, while the HS23 score only dropped by 19 %.

If the cluster's CPUs are permanently limited to the associated frequency, carbon emissions can be reduced similar to the dynamic operation. To still meet the requirement of a constant compute target, the 19 % drop in the HS23 score was compensated by additional hardware and the accompanying increase in embedded emissions. Total emission for operation with a limited frequency are denoted as $E_{\text{total, lim. freq.}}$. The results of dynamic operation and operation with limited CPU frequency are compared in Fig. 5. Similar to Fig. 3, the nominal simulation is again displayed besides dynamic operation results for a varied power consumption ratio $P_{\text{idle}}/P_{\text{max}}$. A comparison of the results

⁴ HS23 is a standard benchmark performance unit to measure computing resources uniformly across the Worldwide LHC Computing Grid (WLCG) [28]. It measures performance using typical high-energy physics workloads from multiple experiments.

shows that limiting the clock frequency at the medium workload leads to increased emissions caused by the additional hardware. The dynamic operation reaches its optimum at the trivial case of $u = 1$. For the backfilling scenario, however, limiting the clock frequency results in an emissions reduction of almost 20 %, thereby significantly outperforming dynamic operation of the GridKa_{ARM} setup regardless of the power consumption ratio.

Thus, the feasibility of such a clock frequency limit is, not only highly dependent on the hardware but also affected by the cluster's typical workloads. However, given a clear power efficiency maximum is found, a clock frequency limit could be combined with dynamic operation of a computing cluster. A simulation of this combination is not part of this study but offers an interesting additional test case.

Emission Operation Threshold: Validation

In order to verify the emission minimisation results, in particular the derived threshold's stability, a validation was carried out based on the 2023 and 2025 data. As an example, the BAF_{modern} setup's backfilling scenario was used. The extrapolation defined in Sect. "Methodology" was applied. A comparison between the targeted utilisation u_{target} , associated with the threshold X_{emission} determined via the 2024 data, and the utilisation u_{extra} reached with the extrapolated thresholds $X_{\text{emission}}^{\text{extra}}$ is listed in Table 6. Resulting deviations are within 2 percent, which implies that the compute target is almost perfectly met in both validation periods.

The extrapolated thresholds $X_{\text{emission}}^{\text{extra}}$ are displayed in Fig. 6. In addition both are compared to the threshold $X_{\text{emission}}^{\text{target}}$, individually derived to match u_{target} in each validation period. This comparison yields relative deviations of 1.42 % and 0.68 % for the years 2023 and 2025, respectively. Therefore, taking into account the changed share of renewable energies, no significant deviations are found within both validation periods.

Optimisation with Respect to Operational Costs

Due to a lack of data, optimisation with respect to operating costs was solely carried out for the two different BAF computing cluster setups. Both setups were simulated under the assumption that 100 % of the electricity demand is purchased on the spot market. A value of $C_{\text{demand}} = 100 \text{ €/kW/year}$ was assumed for the demand charge. This was again only intended to serve as a rough estimate, as grid fees in Germany generally depend on the supplier, the region and the duration of use. The acquisition costs per logical core hour were assumed to lie at $C_{\text{acq., CPU hour}} = 5.35 \times 10^{-4} \text{ €/h}$ for both scenarios. This value was determined based on the costs of BAF hardware currently in use. The calculation only incorporates purchase prices for CPUs and RAM, while additional costs for infrastructure etc. are excluded. An internal comparison between original acquisition costs in early 2017 and the costs for more recent expansions in 2022 and 2023 revealed that the acquisition cost has remained approximately constant to date. It is difficult to predict how acquisition costs will change in the future. However, with global developments in the field of artificial intelligence and the associated demand for RAM modules [30] a cost increase seems increasingly likely.

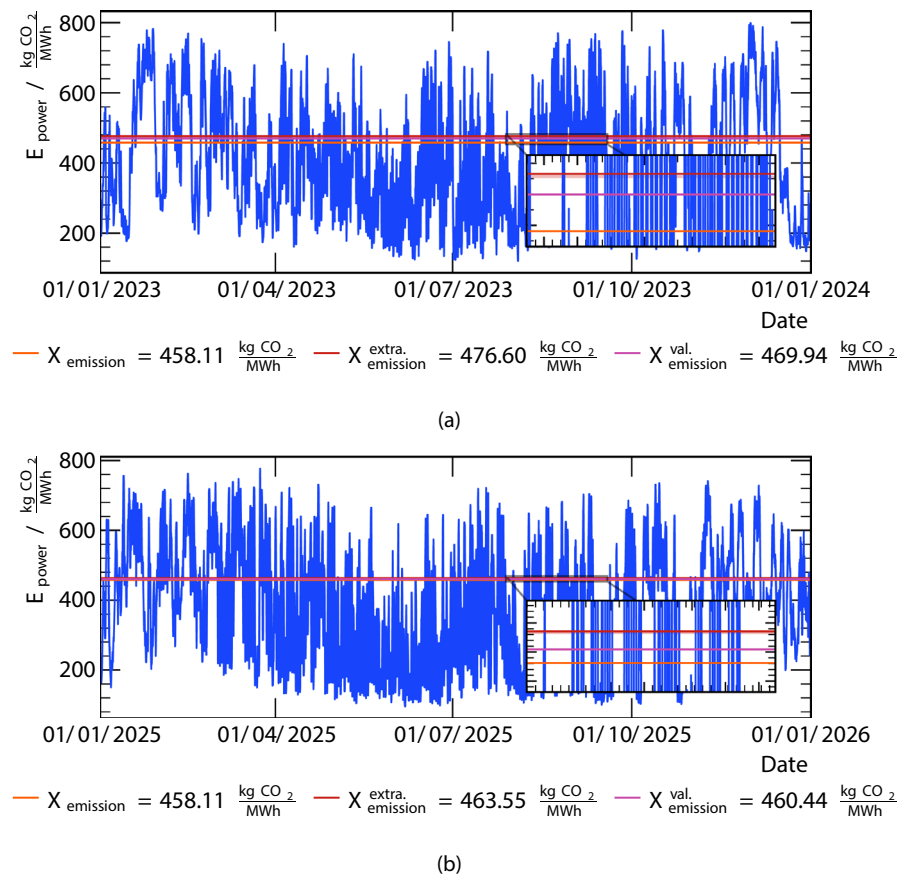


Fig. 6 Validation of the calculated threshold for the BAF_{modern}'s backfilling scenario in the years **a** 2023 and **b** 2025. Both validation periods display the original threshold value X_{emission} determined via the 2024 data, the extrapolated threshold $X_{\text{extra.emission}}$, taking into account the changed share of renewable energies as well as the threshold $X_{\text{target.emission}}$ derived individually in each validation period to match u_{target}

Table 6 Target utilisation u_{target} derived via the 2024 data and utilisation obtained in a validation period u_{extra} , taking into account the changed share of renewable energies for the BAF_{modern} setup's backfilling scenario

Validation period	u_{extra}	$X_{\text{extra.emission}}^{\text{extra}}/\text{kgCO}_2/\text{MWh}$	u_{target}	$X_{\text{emission}}^{\text{target}}/\text{kgCO}_2/\text{MWh}$
2023	0.648	476.60	0.635	469.94
2025	0.648	463.55	0.635	460.44

Associated emission thresholds are also listed

The assumed values used for both setups in terms of power consumption, i.e., P_{max} and P_{idle} , still correspond to those listed in Table 1. Similarly, all three workload scenarios from Table 3 were also reused.

Cost minimisation results are listed in Table 7 for both setups. Similar to the results in Sect. 'Optimisation with Respect to Carbon Emission', Fig. 7 additionally illustrates total and individual costs as a function of the utilisation u in case of BAF_{modern}'s backfilling scenario. As expected, a decreased utilisation causes not only an increase in C_{idle} ,

Table 7 Optimisation results to minimise operational costs for both BAF setups and workload scenarios

Workload scenario	$u_{\text{opt.}}$	$X_{\text{cost}}/\text{€/MWh}$	$C_{\text{total}}(u_{\text{opt.}})/C_{\text{total}}(u = 1)$
BAF _{default}			
Medium	0.999	470.78	0.999
Heavy	0.996	351.70	0.998
Backfilling	0.985	200.14	0.992
BAF _{modern}			
Medium	0.998	438.30	0.999
Heavy	0.996	330.05	0.998
Backfilling	0.983	189.32	0.991

Listed is the optimal utilisation $u_{\text{opt.}}$, the associated threshold X_{cost} beyond which a computing cluster switches to idle state, and the relative cost compared to constant operation $C_{\text{total}}(u_{\text{opt.}})/C_{\text{total}}(u = 1)$. In case that a constant operation yields the lowest total costs, no threshold on C_{power} is specified

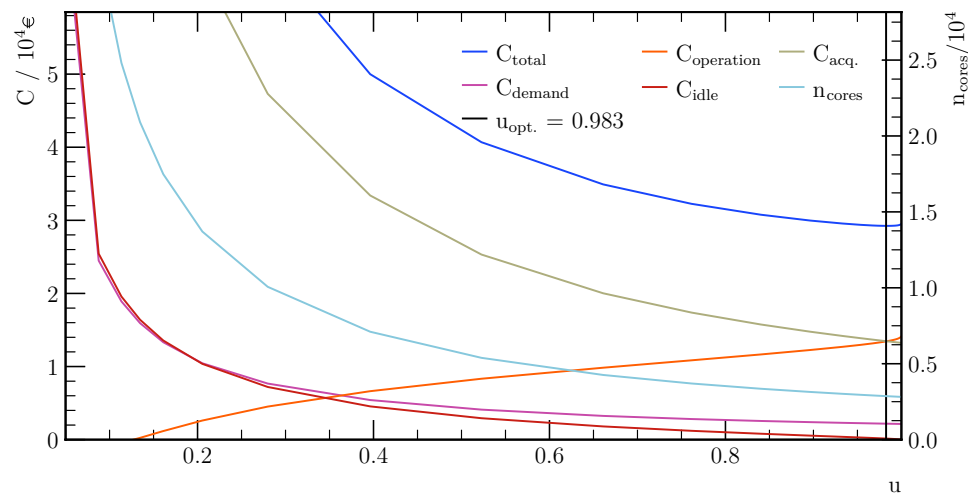


Fig. 7 Display of total costs alongside the individual contributions (left y-axis) and number of logical cores needed to maintain the default compute target (right y-axis) as a function of the utilisation u for the BAF_{modern} setup's backfilling scenario. The optimal cluster utilisation is marked by the black vertical line. While $C_{\text{operation}}$ increases with rising utilisation, C_{idle} , C_{demand} and $C_{\text{acq.}}$ decrease. The latter two results from reduced hardware requirements to meet the constant computing target

but also in C_{demand} and $C_{\text{acq.}}$. The latter both represent costs connected to the additional hardware, required to maintain a constant compute target.

A comparison of all listed scenarios demonstrates that the greatest cost-saving is achieved from simulating a dynamic operation under heavy cluster utilisation, as was also the case for the carbon emission studies. Nevertheless, the relative savings for all tested scenarios lie below 1 percent. This limited saving potential can be attributed to the dominating acquisition costs $C_{\text{acq.}}$, as shown in Fig. 7. They rise for utilisations of $u < 1$, as more hardware is required to keep the compute target constant.

However, the actual acquisition costs for different setups' hardware might vary significantly. To assess the potential of a dynamic operation as a function of the acquisition costs, further simulations were carried out for the BAF_{modern}'s backfilling scenario. Results for an acquisition cost reduction of 50 % are shown in Fig. 8. While $C_{\text{acq.}}$ is

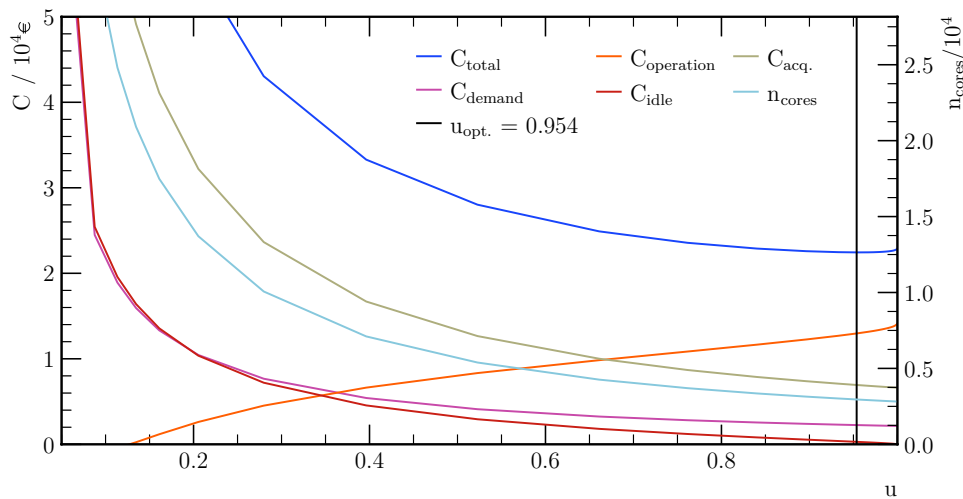


Fig. 8 Display of total costs alongside the individual contributions (left y-axis) and number of logical cores needed to maintain the default compute target (right y-axis) as a function of the utilisation u for the BAF_{modern} setup's backfilling scenario assuming $C_{acq.}$ is reduced by 50 %. The optimal cluster utilisation is marked by the black vertical line

no longer the dominant contribution for $u \gtrsim 0.7$, the relative savings compared to $C_{total}(u = 1)$ are still below 2 %. Consequently, a dynamic operation can, even under the assumption of reduced acquisition costs, only achieve quite limited savings.

Following the approach in Sect. “Optimisation with Respect to Carbon Emission” on emissions, the cost reduction potential was also evaluated across varying electricity prices. To this end, simulations up to a reduction of C_{power} to 10 % of its original value were carried out for the BAF_{modern}'s backfilling scenario. In principle, lower electricity prices naturally reduce the savings from a dynamic operation, i.e., $u_{opt.}$ increases. The simulation showed that small savings of less than 1 % are achievable as long as the electricity price C_{power} does not drop by more than 85 %. Thus, apart from the fact that operational cost savings showed to be quite limited across all tested scenarios, only a significantly reduced electricity price leads to the trivial result of $u_{opt.} = 1$.

Conclusion

Germany's growth in renewable energies is inevitably linked to relatively large fluctuations with respect to the total energy production. Large power consumers whose operation can be dynamically controlled are able to partially compensate these fluctuations. Through trading on the electricity spot market such a dynamic operation offers potential savings in carbon emission and operating costs. This study investigated the feasibility of dynamically operating a computing cluster used in science. These computing clusters represent an ideal test case, since short-term delays caused by any dynamic operation are often negligible as long as the overall compute target remains constant over prolonged periods.

On basis of the German energy generation data of 2024, dynamic operation for five different computing clusters installed at research institutes was simulated. The results were compared with respect to a potential carbon emission and operational cost

reduction. Savings with respect to the operating costs during the simulation period were found to be quite limited due to the dominant acquisition costs. For all tested setups and workloads relative savings stayed below 1 % with respect to a standard non-dynamic operation.

In contrast, the results regarding carbon emissions showed more pronounced differences between the various cluster configurations. The decisive factor for the effective implementation of dynamic operation was found to be the relative idle power consumption $P_{\text{idle}}/P_{\text{max}}$. Setups with a comparatively high value allowed for little to no emission reduction, whereas the BAF_{modern} setup with $P_{\text{idle}}/P_{\text{max}} = 0.15$ achieved savings of up to 8 %.

To validate the results, dynamic operation of the BAF_{modern} setup's backfilling scenario was additionally simulated for the years 2023 and 2025. Taking into account the changed share of renewable energies, a deviation from the targeted relative utilization of no more than 2 % was attainable. This demonstrated, at least within the limited validation period, that a fixed annual compute target can be met consistently.

Apart from the two specific setups, simulations were also carried out for variable modelling parameters. Values for relative idle power consumption and acquisition costs were varied independently in order to assess the effects of potential future developments or different hardware setups. To at least partially evaluate the effect of the embedded emissions' large uncertainties additional testes were carried out. It was demonstrated that the dynamic operation allowed for an emission reduction within the chosen uncertainty band.

The GridKa_{ARM} setup enabled a comparison between dynamic operation and continuous operation at a limited clock frequency. Results showed to be quite hardware- and workload-dependent but combining both methods might be a viable option to decrease the emission in the future.

While this study is intended to serve as a proof of concept, demonstrating the potential for savings, particularly in terms of carbon emissions, further steps are necessary before implementing the dynamic operation of a cluster in practice. Among other simplifications used in the simulation, an interface to stop and restart the cluster operation is mandatory. Computing jobs must, therefore, be able to be paused and resumed flexibly. Currently it is assumed, that switching between idle and normal operation happens instantaneously and can be performed as often as required. The number of these switching operation increases for lower utilisation values. In case of the BAF_{modern} setup's backfilling scenario the cluster had to pause and resume operation roughly 230 times in the simulated period with an average runtime interval length of 20.7 h over the whole year. For practical application, however, a ramp-up and ramp-down time must definitely be taken into account. This might affect and shift the optimisation results. In addition, a more sophisticated scaling factor has to be implemented and replace the simple scaling via the number of logical cores, currently utilized in this study. Otherwise, a realistic consumption, especially for inhomogeneously structured computing clusters consisting of various CPU and accelerator modules is hard to estimate.

Acknowledgements

The present analysis would not have been possible without the publicly available energy production and spot market price data by the Fraunhofer Institute for Solar Energy Systems (ISE). The authors would thus like to express their gratitude and support for open data. In addition, the authors are grateful for access to the resources of the BAF2 cluster at the

University of Bonn, the GridKa cluster at the Karlsruhe Institute of Technology, as well as the DEEP-System of the Jülich Supercomputing Centre to simulate the different setups. This work was supported by the Federal Ministry of Research, Technology and Space (BMFTR) within the project FIDIUM (Förderkennzeichen 05H21PDR1) and SUSFECIT (Förderkennzeichen 05D25PD4). The collaboration within the PUNCH4NFDI consortium in the NFDI (Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—project number 460248186) was instrumental in carrying out the project. It was also supported by the environment of the cluster of excellence EXC 3107 “Color meets Flavor” (project number 533766364) as well as the collaborative research centre “NuMerIQS” (project number 511713970). We are very grateful for the support by all participants and collaborators of the SUSFECIT consortium, most importantly Markus Schumacher and Michael Böhler.

Author Contributions

Conceptualization was done by Philip Bechtle who also supervised the complete study. Florian Kirfel wrote the first draft of the manuscript and conducted the complete simulation for all presented scenarios. Oliver Freyermuth and Michael Hübner conducted the BAF load tests to estimate the power consumption and derived the BAF acquisition costs. Similar load tests for the DEEP modular supercomputer prototype and the GridKa ARM worker nodes were performed by Stefan Krieg & Jochen Kreutz and Manuel Giffels & Matthias Schnepf, respectively. The latter two also provided the power efficiency measurements for different clock frequencies. Max Geffers worked on initial studies regarding cost and carbon emission estimation for a fixed computing cluster size. Sebastian Matberg performed early studies to test the checkpointing of compute jobs. All authors commented on previous versions of the manuscript and approved the final manuscript.

Funding

Open Access funding enabled and organized by Projekt DEAL.

Data Availability

The data used on electricity generation in Germany is publicly available at <https://energy-charts.info/> and was used with the permission of Leonhard Gandhi. The framework used for the simulation is available at https://gitlab.uni-bonn.de/grp_phy/drittmittelprojekte/erum-data/susfecit/eco-clusters-study.

Declarations

Competing interests

The authors declare no competing interests.

Received: 18 May 2026 Accepted: 25 August 2026

Published online: 15 September 2026

References

- Arrhenius S (1896) On the influence of carbonic acid in the air upon the temperature of the ground. London Edinburgh Dublin Philos Magazine J Sci 41:237–76. <https://doi.org/10.1080/14786449608620846>
- IPCC (2023) Climate Change 2023: synthesis report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Geneva, Switzerland: IPCC, 35–115. <https://doi.org/10.59327/IPCC/AR6-9789291691647>
- Stöver B, Reuschel S, Wolter MI, Daßler J, Bernardt F (2025) Die Kosten des Klimawandels für Deutschland 2025–2050. GWS Research Report Series 25-1. GWS—Institute of Economic Structures Research
- (2015) United Nations Framework Convention on Climate Change. Paris Agreement: Preamble
- Icha P, Lauf T (2026) Entwicklung der spezifischen Treibhausgas-Emissionen des deutschen Strommix in den Jahren 1990–2025. Umweltbundesamt,
- Federal Ministry for Economic Affairs and Energy. Renewable energies in figures—2025 edition. (2026) Available from: <https://energiewende.bundeswirtschaftsministerium.de/EWD/Redaktion/EN/Newsletter/2026/03/Meldung/direkt-finds.html> [Accessed on: 2026 Jul 30]
- Brown T, Schlachberger D, Kies A, Schramm S, Greiner M (2018) Synergies of sector coupling and transmission reinforcement in a cost-optimised, highly renewable European energy system. Energy 160:720–39. <https://doi.org/10.1016/j.energy.2018.06.222>
- Federal Ministry for Economic Affairs and Energy (2025) Status and development of the German data centre landscape. Executive Summary, Federal Ministry for Economic Affairs and Energy (BMWi)
- Cao Z, Zhou X, Hu H, Wang Z, Wen Y (2022) Toward a systematic survey for carbon neutral data centers. IEEE Commun Surv Tutor 24:895–936. <https://doi.org/10.1109/COMST.2022.3161275>
- Bundesnetzagentur. Record High for Solar Generation in Each Quarter. The electricity market in 2025. 2026 Jan. Available from: <https://www.smard.de/page/en/topic-article/217400/219038/record-high-for-solar-generation-in-each-quarter> [Accessed on: 2026 Feb 19]
- Freyermuth O, Wienemann P, Bechtle P, Desch K (2021) Operating an HPC/HTC cluster with fully containerized jobs using HTCondor, singularity, CephFS and CVMFS. Comput Softw Big Sci 5:9. <https://doi.org/10.1007/s41781-020-00050-y>
- Suarez E, Amaya J, Frank M, Freyermuth O, Gironi M, Kostrzewa B (2025) Energy efficiency trends in HPC: what high-energy and astrophysicists need to know. Frontiers Phys. <https://doi.org/10.3389/fphy.2025.1542474>

13. Lang VS, Bhalla NK, Gurdasani SS, Niknejadi P (2025) Know your footprint—Evaluation of the professional carbon footprint for individual researchers in high energy physics and related fields. *NPJ Clim Action* 4(1): 28. <https://doi.org/10.1038/s44168-025-00232-7>
14. Suarez E, Amaya J, Frank M, Freyermuth O, Gironi M, Kostrzewa B (2025) Energy-aware operation of HPC systems in Germany. *Frontiers High Perform Comput* 3:1520207. <https://doi.org/10.3389/fhpcp.2025>
15. Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A et al (2016) The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 3:1–9. <https://doi.org/10.1038/sdata.2016.18>
16. Bruers B, Cruces M, Demleitner M, Duckeck G, Düren M, Eich N et al (2024) Resource-aware research on universe and matter: call-to-action in digital transformation. *Eur Phys J Special Topics*. <https://doi.org/10.1140/epjs/s11734-024-01436-4>
17. (2023). Gesetz zur Steigerung der Energieeffizienz in Deutschland (Energieeffizienzgesetz). Bundesrepublik Deutschland
18. Gandhi L. Energy-Charts.info: public net electricity generation in Germany. 2026. Available from: <https://www.energy-charts.info/charts/power/chart.htm?l> [Accessed on: 2026 Feb 19]
19. Gandhi L. Energy-Charts.info: carbon dioxide emissions from electricity generation in Germany. 2026. Available from: https://www.energy-charts.info/charts/co2_emissions/chart.htm?l [Accessed on: 2026 Feb 19]
20. Gandhi L. Energy-charts.info: annual renewable share of public net electricity generation and load in Germany. 2026. Available from: https://www.energy-charts.info/charts/renewable_share/chart.htm?l [Accessed on: 2026 Feb 19]
21. Eicker N, Lippert T, Moschny T, Suarez E (2016) The DEEP project an alternative approach to heterogeneous cluster-computing in the many-core era. *Concurr Comput Pract Exp* 28:2394–411. <https://doi.org/10.1002/cpe.3562>
22. Jülich Supercomputing Centre (JSC) (2026) Available from: <https://www.fz-juelich.de/en/jsc> [Accessed on: 2026 Apr 16]
23. Grid-Computing-Zentrum Karlsruhe—GridKa. 2026. Available from: <https://www.scc.kit.edu/forschung/gridka.php> [Accessed on: 2026 Apr 16]
24. thinkstep AG and Dell Technologies (2019) Life cycle assessment of Dell PowerEdge R740. Technical Report. Prepared by thinkstep AG on behalf of Dell Technologies. Dell Technologies
25. Tannu S, Nair PJ (2023) The dirty secret of SSDs: embodied carbon. *SIGENERGY Energy Inform Rev* 3:4–9. <https://doi.org/10.1145/3630614.3630616>
26. Shmueli E, Feitelson DG (2005) Backfilling with lookahead to optimize the packing of parallel jobs. *J Parallel Distrib Comput* 65:1090–107. <https://doi.org/10.1016/j.jpdc.2005.05.003>
27. Krull AA, Schnepf MJ, Fischer M, Petzold A (2025) Experience with ARM WNs at the WLCG Tier1 GridKa. *EPJ Web Conf* 337:01147. <https://doi.org/10.1051/epjconf/202533701147>
28. Worldwide LHC Computing Grid. Worldwide LHC Computing Grid. 2026. Available from: <https://wlcg.web.cern.ch> [Accessed on: 2026 Jul 30]
29. Szczepanek N, Britton D, Girolamo AD, Ketele E, Glushkov I, Giordano D, et al. (2024) HEP Benchmark suite: enhancing efficiency and sustainability in worldwide LHC computing infrastructures. <https://doi.org/10.48550/arXiv.2408.12445>
30. Zhang T, Mailthody VS, Sun F, Ma L, Newburn CJ, Zhang T, et al. (2025) From minutes to seconds: redefining the five-minute rule for AI-era memory hierarchies. <https://doi.org/10.48550/arXiv.2511.03944>. [arXiv:2511.03944](https://arxiv.org/abs/2511.03944) [cs.AR]

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.