Unveiling Key Performance Indicators for the Energy Efficiency of Cloud Data Storage

Banijamali, Pouyeh; Fatima, Iffat; Lago, Patricia; Heitlager, Ilja

2024

document version
Peer reviewed version

Unveiling Key Performance Indicators for the Energy Efficiency of Cloud Data Storage

Pouyeh Banijamali
Vrije Universitiet Amsterdam
The Netherlands
p.banijamali@student.vu.nl

Iffat Fatima
Vrije Universitiet Amsterdam
The Netherlands
i.fatima@vu.nl

Patricia Lago
Vrije Universitiet Amsterdam
The Netherlands
p.lago@vu.nl

Ilja Heitlager
Schuberg Philis
The Netherlands
iheitlager@schubergphilis.com

***Abstract*—The continuous growth of data centers prompts research into their energy impact. While some cloud providers give an overall estimate of data centers energy efficiency, there is a lack of transparency about data for individual components and functions. We explore metrics to evaluate the energy efficiency of cloud data storage in particular and represent them as Key Performance Indicators (KPIs) for organizational monitoring. To this aim, we perform an exploratory study with Schuberg Philis (SBP), a leading IT provider in the Netherlands to explore sustainability monitoring for cloud data storage. We identify the data provided by a cloud provider and explore the corresponding metrics in the academic literature. Based on the SBP's sustainability roadmap, we propose three KPIs that are developed using an existing KPI framework. These KPIs are reviewed by the experts at SBP for their correctness and added value. The study provides, (i) an overview of the energy efficiency of cloud data storage, (ii) suggested KPIs to monitor and optimize the cloud data storage to meet the SBP's sustainability goals. The KPIs are energy consumption (kW/h), storage capacity utilization (%), and capacity per watt (GB/Watt), and (iii) actions per KPI to reach the sustainability targets. The results of the study provide insights into monitoring the energy efficiency of data storage for SBP for potential optimization while building a foundation for developing similar approaches for other practitioners and for targeting sustainability concerns.**

***Index Terms*—software, sustainability, key performance indicator, KPI, cloud, data storage, cloud storage, energy efficiency, power consumption**

I. Introduction

Over the last decade, the exponential growth in global server instances (647%) and storage capacity (2,500%) has raised concerns about the environmental impact of data centers [1]. The International Energy Agency reported that data center electricity consumption accounted for 1-1.3% of the total global electricity demand in 2022.[1]

Despite this surge, there remains a blind spot in precisely estimating the energy impact of data centers due to insufficient information on data center type, location, and energy efficiency[2]. Although, cloud providers report their carbon footprint and energy consumption, there is a lack of transparency in data collection and reporting practices. For instance, NetApp provides a sustainability score as an indicator for the environmental sustainability of the storage system without further documentation on how this score is calculated.[2] To address this, there is a growing need for unified approaches and measures to provide reliable data on the energy efficiency of data centers, focusing on specific functions and components.

It is vital to report data at an atomic level transparently, and how it is aggregated into high-level sustainability indicators. An indicator value without baseline measurement information is not reliable or comparable to measurements from other sources (cloud providers).

This study concentrates on one aspect of cloud data centers, *i.e.,* data storage, to explore metrics for quantifying sustainability indicators. Masanet *et al.* [2] highlight storage as the second-highest energy-consuming element in global data centers. Despite the development of numerous metrics for monitoring the energy consumption of data centers, there is a lack of monitoring addressing energy efficiency while focusing on specific resources such as data storage [3], [4], [5].

In this study, we explore the data storage infrastructure of Schuberg Philis (SBP), a leading IT provider specializing in innovative and sustainable IT solutions.[3] Utilizing NetApp[4], which is used as one of the data storage provider for SBP's cloud applications, we model Key Performance Indicators (KPIs) based on SBP's future roadmap, offering a continuous monitoring mechanism. We use the KPI Framework by Fatima *et al.* [6] that provides a systematic approach to developing sound KPIs for the performance of organizations for software-related goals. The resulting KPIs are designed to be reusable and can serve as a valuable resource for other organizations aiming to monitor and improve their data storage energy efficiency. These indicators offer flexibility, allowing organizations to tailor targets and actions based on their specific sustainability goals. Through this approach, our study provides an example use case that can be adapted for monitoring energy efficiency across diverse organizational contexts.

II. Background

Global data centers provide the main infrastructures and facilities for data storage. In the context of our study, it is crucial to understand the importance of energy efficiency of

[1]https://www.iea.org/energy-system/buildings/data-centres-and-data-transmission-networks

[2]https://docs.netapp.com/us-en/active-iq/BlueXP_sustainability_dashboard_overview.html,

[3]https://schubergphilis.com/en

[4]https://www.netapp.com/

cloud data centers while investigating the energy impact of data storage. Pérez-Lombard *et al.* [7] define energy consumption as the product of service demand and energy intensity, representing the absolute energy used. Energy saving entails a reduction in energy use. In contrast, energy efficiency refers to technological improvements aiming to decrease energy usage, calculated as the ratio of service output to energy input. Essentially, it is about getting more done with less energy.

Getting accurate data on the current and future energy impact of global data centers is challenging. Corbett *et al.* [8] reports an estimation of 46.33 kWh per TB of data in 2018 on a global scale based on a report by Lawrence Berkeley National Laboratory [9] done for US data centers. The study also considers the ongoing trends in using renewable energy sources. Andrae [10] estimates that by 2030, electricity usage by data centers could range from 446 TWh in the best-case scenario—reflecting an estimation of 20% annual improvement in energy efficiency trends—to 974 TWh in the expected case, which considers continuation of current trends in energy efficiency technologies and practices. This represents a significant increase from 299 TWh in 2020.

Looking ahead, by 2030 servers are expected to use 74% of the electricity in data centers, with infrastructure using 17%, and network and storage using 9%. Corbett's review [8] of the estimates from 2007 to 2021 showed that data storage infrastructures have a significant energy footprint, making it crucial to monitor their energy use.

One major issue contributing to energy inefficiency is *dark data*. Corbett *et al.* [8] introduce dark data as the information stored but not used. This unnecessary data puts an extra strain on storage infrastructures. To address this issue, the study suggests adopting tools such as value stream mapping and governance standards is crucial for enhancing sustainable data management practices. Recognizing these factors is vital for devising strategies to boost energy efficiency in data centers.

III. Related work

This section focuses on studies in the literature that provide metrics on the environmental sustainability of data centers. Redy *et al.* [3] present sustainability metrics like energy efficiency, greenhouse emissions, performance and productivity, thermal, humidity, and air management, storage performance, security performance, and financial metrics. Wang *et al.* [11] present energy efficiency metrics in terms of performance, space, and power consumption (SWaP). Shao *et al.* [12] categorize energy efficiency metrics into energy conservation, eco-design, and data center security.

For energy conservation, Power Usage Effectiveness (PUE) [4] is widely used. It shows the relation between the energy used by IT equipment and energy used by other data center facilities. Data Center Energy Productivity (DCeP) [5] characterizes the data center's useful work compared to the energy consumption. Data Center Infrastructure Efficiency (DCiE) [13] provides the ratio of IT equipment power compared to the total facility power.

IT Equipment Energy Efficiency (ITEE) metric [14], covers energy use across all IT components, such as compute, storage, and network equipment. This metric is the ratio of IT equipment's total capacity and total rated power. Schulz *et al.* [15] introduce a capacity metric that is used to measure the amount of data moved per second per energy used. Björn *et al.* [16] present a framework for energy efficiency metrics across various levels of the data center infrastructure, with a specific focus on storage, network, and server resources. Brocklehurst [17] presents KPIs for monitoring the energy efficiency of data storage. The study presents existing sustainability metrics interchangeably as KPIs. Direct mapping of KPIs to metrics is not always feasible as the context of goals is important to measure only necessary data.

The existing metrics for storage focus on capacity or energy consumption separately. This highlights the need to integrate both concepts to make meaningful analyses for energy efficiency. The variation in data storage facilities (*e.g.,* data storage location, energy sources, etc) necessitates structured approaches for definition and continuous monitoring of energy efficiency. Our study aims to fill this gap by introducing KPIs, within the context of an industrial case aimed at monitoring and potentially enhancing energy efficiency of data storage.

IV. Study design and Execution

A. Goal

The goal of our study is defined based on Wohlin *et al.* [18] and is depicted in Table I.

TABLE I
Goal of the study

Analyze	Data Storage Metrics and KPIs
For the purpose of	Monitoring
With Respect to	Energy Efficiency
From the point of view of	Data Storage Cloud Customer
In the context of	Cloud Data Storage

B. Research Questions

We delve into solving the problem of monitoring the energy efficiency of cloud data storage for potential optimization. We aim to address this problem via two research questions.

RQ1: What metrics are used to monitor the energy efficiency of cloud data storage?

RQ 2: What sound KPIs can be developed to monitor the energy efficiency of cloud data storage?

C. Overview of the Study Design

Figure 1 presents different phases of the study design.

Phase 0 is a preliminary phase that focuses on understanding the background and importance of the problem as well as extracting relevant metrics currently adopted for addressing sustainability concerns of data centers in literature. This phase is conducted by reviewing peer-reviewed papers as well as grey, and white literature.

Phase 1 focuses on answering RQ1 by collecting capacity utilization and power consumption datasets from NetApp. These datasets can offer insights into organization's data storage energy efficiency. From these datasets, we identify data points relevant to power consumption or the useful output of the data storage environment and practices for further analysis. The collection and analysis of data from NetApp is facilitated by the storage expert at the organization. The energy efficiency trends in data are visualized as graphs generated via Python. The data and scripts are provided in the replication package.[5]

Based on data analysis, we identify points in the data storage practices and environment where monitoring and improvement of data storage is needed for energy efficiency. These improvement points are discussed with the storage expert at the organization for possible actions. Additionally, using relevant metrics collected from the preliminary phase together with the result of data analysis, new metrics are defined. These metrics are used as input for the KPI framework [6] for Phase 2.

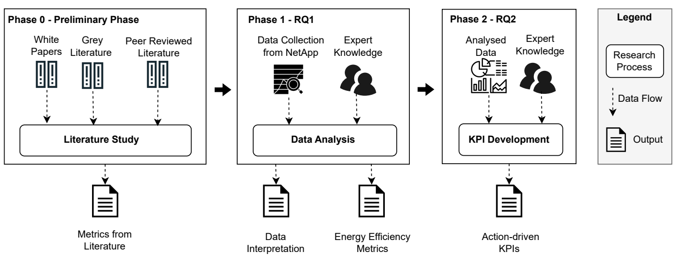

Fig. 1. Schematic overview of the study design

Phase 2 aims at answering RQ2. In this phase, we define KPIs by leveraging the metrics identified in Phase 1. Using the KPI Framework [6], we work with the organization's sustainability expert to set targets and actions for each KPI, facilitating the achievement of sustainability goals.

D. Study Execution

In this section, we describe the execution of our study.

Phase 0 - Preliminary Phase. During this phase, we consulted peer-reviewed papers, white papers, and grey literature. Initially, we established the context and significance of our study by exploring the literature on the energy impact of data centers. We also examined existing literature on metrics, measures, and Key Performance Indicators (KPIs) for data center energy efficiency. We observed a lack of KPIs for monitoring the energy efficiency of cloud data storage. Further, we identified a lack of a systematic approach to data aggregation for KPI development. We discovered many sustainability metrics (see Sec. III) for data centers and we made use of one metric (Capacity metric [15]) later on in Phase 1 of the study.

Phase 1 - Data Collection and Analysis for RQ1. In this phase, we identified the cloud data storage resources being currently used at the organization. We identified that the organization has 23 clusters each containing one or two nodes [6] managed via NetApp. The nodes are of four models (AFF-A200, AFF-A300, AFF-A700s, AFF-A800). NetApp provides projections of power consumption, direct carbon usage, and heat emission (see 2). We focused on power consumption as it can be used to monitor the energy efficiency of data storage. NetApp provides power consumption (kW) data as, (i) typical power consumption per cluster or, (ii) real-time power consumption per cluster.

NetApp requires an additional data collector for collecting real-time power consumption data. Following discussions with the organization's data storage expert, we determined that, due to current infrastructure conditions and project time constraints, real-time data usage was impractical and could be explored in future research. Therefore, we focused on the typical power consumption (kW) data. NetApp lacks detailed documentation on calculating typical power consumption, mentioning reliance on device specifications. To verify, we analyzed three reports at various times noting close but not identical values across all clusters.

We obtained two datasets for storage capacity from NetApp environment. One in real-time for October 2023 and another spanning monthly data from 2022-2023, including forecasts for 2024. In the former dataset, we had data points for Total Usable Capacity (Tebibyte (TiB)), Used Capacity (TiB), Used Capacity (%) as well as data on the impact of space-saving techniques including Space Saved by Snapshots (TiB), Space Saved by Clones (TiB), Space Saved by Volume Deduplication (TiB), Space Saved by Volume Compression (TiB), Space Saved by Compaction, Aggregate Deduplication and Compression (TiB) per node. The latter dataset only consisted of the Total Usable Capacity (TiB), Used Capacity (TiB), and Used Capacity Percentage per node. We conducted an exploratory data analysis to identify patterns (*e.g.,* power consumption) and elements (*e.g.,* nodes and clusters).

Based on the results of our data analysis, and analyzing the existing metrics in the literature regarding the energy efficiency of data storage, we identified the relevant metrics that could be used by the organization.

Phase 2 - KPI Development for RQ2. In this phase, we developed KPIs, targets, and actions using the KPI framework [6] and Phase 1 metrics. We defined goals and critical success factors by discussions with the sustainability expert based on organization's sustainability roadmap. We defined KPIs by incorporating metrics from our data analysis and one from the literature, ensuring a comprehensive approach addressing energy efficiency. The 3 KPIs cover energy consumption, capacity utilization, and capacity per watt ratio for storage infrastructure and practices. With the help of the organization's storage expert, we established targets and actions, initially based on the tactics provided by Vos *et al.* [19].

Following the discussions with the sustainability expert, we confirmed, refined, and introduced new actions. As a result of Phase 2, we present sound KPIs for the organization to monitor energy efficiency. The KPIs undergo validation using the checklist from the KPI Framework [6], revealing minor non-conformities to the checklist in the organization's current practices, described in Section VI.

[5] https://github.com/S2-group/GREENS-2024-KPIs-EnergyEfficiency-rep-pkg
[6] https://library.netapp.com/ecmdocs/ECMP1552961/html/GUID-FD018D8A-B07A-464E-9296-9B460F9BD332.html

V. Results

In this section, we explain the results of the conducted exploratory data analysis as well as the results of applying the KPI Framework to develop and define its building blocks. Combining the two, we answer our research questions.

A. *RQ1 - Data Collection and Analysis*

To answer RQ1, we present the results from Phase 1.

Power Consumption per Total Usable TB of Data For initial data analysis, we computed power consumption for nodes across various clusters, aiming for a detailed insight. Different models exhibited distinct power consumption patterns. Recognizing varying usable capacity among models, we opted to calculate power consumption (kW) per usable TB for each node, considering total usable capacity (TB). Nodes with model AFF-A220 displayed lower efficiency concerning power consumption relative to total capacity. As anticipated, consistent patterns emerged within similar models offering comparable TB of usable capacity. Figure 2 illustrates the results categorized by node models.

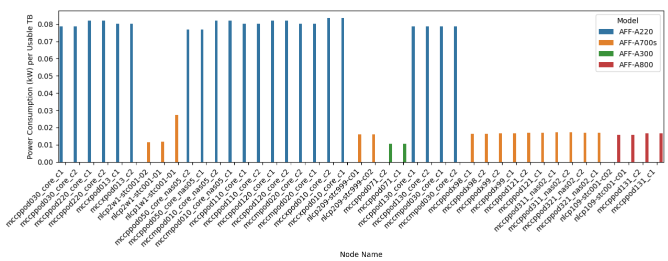

Fig. 2. Power Consumption (kW) Per TB of Usable Capacity per Node High-resolution figures can be found in Additional Material[5]

Storage Capacity Utilisation

We examined NetApp's storage space and capacity utilization data. Two datasets were analyzed, (i) real-time data from October 2023, and (ii) monthly trends for 2022-2023 and predictions for 2024. To ensure reliability, we calculated the mean or the median of the capacity utilisation per node for the period of 2022 and 2023. In order to choose between mean or median, we conducted the Shapiro-Wilcoxon test to check for normality as different nodes seemed to have different distributions of capacity utilisation over the mentioned time. Following that, we calculated a representative value for each node (mean in case of a normal distribution and median in case of non-normal distribution). The distribution of the representative values, presented in Fig 3, revealed model-specific trends. Notably, the AFF-A220 storage nodes exhibited inefficient capacity utilization compared to other models.

We further analyzed the data by creating bar plots for clusters to identify the inefficient clusters in terms of capacity utilisation during this period. We focused on the nodes of model AFF-A700s and AFF-A220 as these nodes covered the majority of our clusters. We made two observations regarding these clusters. Firstly, the majority of the clusters using AFF-A220 devices were inefficient showing very low values of

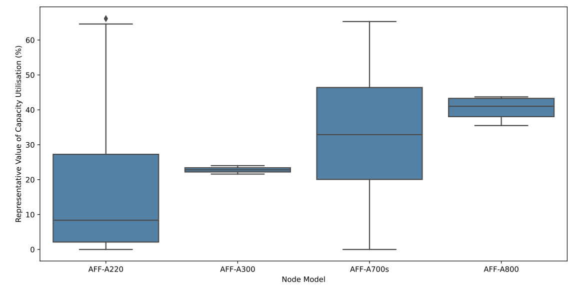

Fig. 3. Representative Used Capacity Percentage by Node Model (2022-2023)

capacity utilization. Secondly, for clusters with model AFF-A700s, at least three out of seven showed a notable disparity in capacity utilization between nodes in the same cluster. These observations were later on discussed with the storage expert in the organization and are further explained in Section VI.

Space Saving Techniques

As mentioned in [19], applying space-saving techniques on less frequently accessed data can lead to energy consumption savings. Therefore, we conducted a set of analyses on the data we obtained on the impact of space saving using different techniques. This data was collected for October 2023. After discussing with the storage expert, we realized that these techniques are mostly applied by NetApp and are rarely changed or managed by the organization and as a result can be relatively representative. Figure 4 shows the share of each space-saving technique provided by NetApp in space-saving across all nodes. We observed that snapshots[7] had a larger share in space saving compared to the other techniques and clones[8] were never used. NetApp considers snapshots as a space-saving technique. A Snapshot is described by NetApp as a copy read-only, point-in-time image of a volume that points to the same disk blocks as the active file system after creation; therefore, it does not use extra disk space. It is also mentioned that over time, the image consumes minimal storage space as it records only the changes created after every copy was made. Clones are also included in NetApp's space-saving techniques as they are created from a copy of the data referenced by snapshots.

We examined different techniques used across all the clusters depicted in Figure 5. While a few clusters leverage techniques such as compression or de-duplication for space-saving, the majority of the clusters only use snapshots for space-saving. This comparison was conducted by normalizing the amount of space saved by different techniques between 0 and 1 since they were not initially in the same range.

Capacity per Watt Ratio

Finally, we calculated the Capacity per Watt Ratio for all the nodes. This metric looks at the power consumption needed for

[7]https://docs.netapp.com/us-en/ontap/concepts/snapshot-copies-concept.html

[8]https://library.netapp.com/ecmdocs/ECMLP2854716/html/GUID-BBE793F6-F72F-4709-B71D-864133918DB1.html

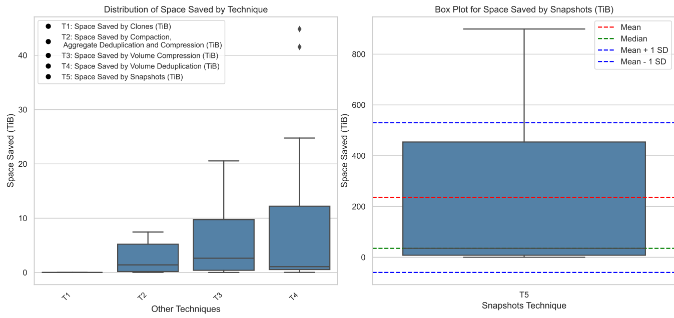

Fig. 4. Distribution of Space Saved by Technique

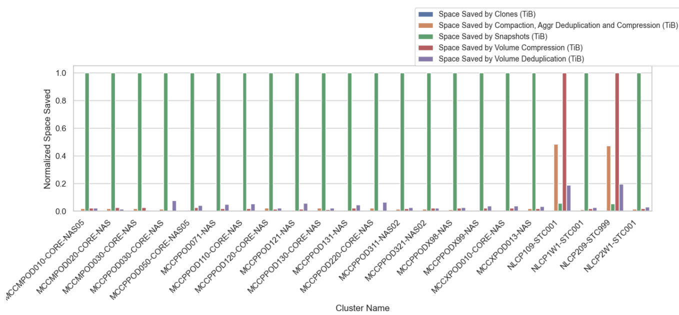

Fig. 5. Space Saved by Techniques per Cluster

storing a certain GB of data per second [15]. Therefore, maximizing this metric leads to more efficient power consumption. Additionally, very low values of this metric during a period can show inefficiency in terms of energy consumption as well. Figure 6 shows the calculated capacity per watt ratio for all the nodes from October 2023. The Figure shows a general inefficiency in clusters of AFF-A220 and one inefficient cluster of AFF-A800. This metric, if monitored continuously, can be very valuable for identifying the need for improvement in certain clusters in terms of energy efficiency. It can be seen that certain nodes have a value close to zero which shows inefficiency in terms of under-utilization of capacity.

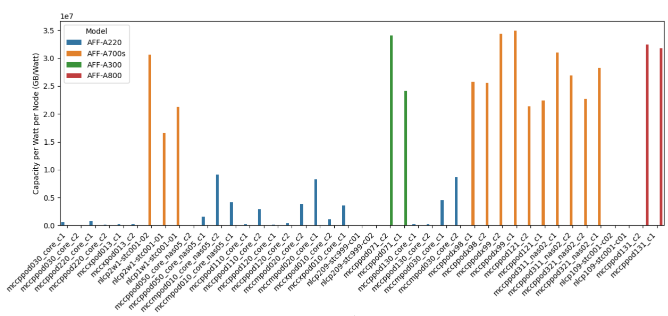

Fig. 6. Capacity per Watt Ratio per Node (GB/Watt)

B. RQ2 - Creating KPIs

After conducting the data analysis, we applied the KPI Framework with a top-down approach. We defined the goal and critical success factor (CSF) in consultation with the sustainability expert from the organization. Sustainability is integrated into the organization's strategy with a clear roadmap. Further using a bottom-up approach, we identified three metrics and their corresponding measures to address energy efficiency.

The first defined KPI is energy consumption (kW/h) (see Table II). This KPI addresses energy consumption as a critical factor in energy efficiency. Energy consumption can be measured on different levels (in this case, per node, per cluster, or for all the clusters in use). It can also be measured per usable or used TB of data for the IT equipment in question. Since energy consumption is calculated by measuring power consumption (kW) per unit of time, the corresponding defined measures for this metric are power consumption (kW) and unit of time (hour). Although we used typical power consumption values for our analysis, this metric can be calculated in real-time and then directly translated to energy consumption for continuous measurement.

It is important to note that, energy efficiency is not just about energy consumption as it involves a relation to the useful service output as well. This means that energy efficiency can increase by increasing the service output using the same amount of energy. In computer science, the service output is often expressed in terms of resource usage.

Therefore, the second KPI is defined as storage capacity utilization (%) (see Table III). This KPI addresses the storage system's useful output as another critical aspect of energy efficiency. This allows us to monitor and evaluate the efficiency of the storage components in terms of efficiency of their service output. This can also be measured on the level of storage nodes, clusters, or all clusters in this case. Additionally, the metrics must be continuously and frequently measured to identify inefficiencies and make improvements if necessary. We define this metric by calculating the percentage of the ratio of the following measures: used storage capacity (GB) and total usable storage capacity (GB). After examining service output and energy consumption independently, it's crucial to evaluate their combined effect on energy efficiency. Energy efficiency can be defined as: "the ratio between service output or results and the energy input required to provide it" [7].

$$Energy\ Efficiency = \frac{Service\ Output}{Energy\ Consumption}$$

We define the final KPI as capacity per watt by translating service output to the used capacity of the storage component (see Table IV). This metric also addresses energy efficiency as it considers the useful output of a storage component in ratio to the power consumed for achieving the output and can be defined as follows:

$$Capacity\ per\ Watt = \frac{Used\ Capacity\ (GB)}{Energy\ Consumed\ (Watt)}$$

We also provide a set of actions for each KPI based on the results of data analysis and suggestions from the storage and sustainability experts (see Table II, III, and IV).

TABLE II
Energy Consumption KPI

Energy Consumption KPI					
Goal	**CSF**	**KPI**		**Metric**	**Measure**
Monitoring SBP's environmental sustainability of data storage to achieve carbon neutrality by 2030	By the end of (2024) monitor the energy efficiency of the data storage used by organization for reporting and potential optimization	Energy Consumption (kW/h)		Energy Consumption (kWh) = Power Consumption (kW) * Time (hour)	Power Consumption (kW), Time (hour)
		Target	**Action**		
		To be minimized	Replace hardware for minimizing power consumption per GB of usable data		

TABLE III
Storage Capacity Utilisation KPI

Storage Capacity Utilisation KPI					
Goal	**CSF**	**KPI**		**Metric**	**Measure**
Monitoring SBP's environmental sustainability of data storage to achieve carbon neutrality by 2030	By the end of (2024) monitor the energy efficiency of the data storage used by the organization for reporting and potential optimization	Storage Capacity Utilisation (%)		Storage capacity utilization = (Used Capacity (GB) / Total Usable Capacity (GB)) * 100	Used Capacity (GB), Total Usable Capacity (GB)
		Target	**Action**		
		To be optimized	Deallocate unused resources Architecture redesign Optimize data management practices		

TABLE IV
Capacity per Watt KPI

Capacity per Watt KPI					
Goal	**CSF**	**KPI**		**Metric**	**Measure**
Monitoring SBP's environmental sustainability of data storage to achieve carbon neutrality by 2030	By the end of (2024) monitor the energy efficiency of the data storage used by the organization for reporting and potential optimization	Capacity per Watt (GB/Watt)		Capacity = Used Capacity (GB) / Power Consumed (Watt))	Power Consumption (Watt), Used Capacity (GB)
		Target	**Action**		
		To be maximized	Compress infrequently accessed data and combine inefficient nodes while considering the energy trade-off		

TABLE V
Self-assessment Checklist

Checklist Item	**Goal**	**Critical Success Factor**	**Key Performance Indicator (KPI)**	**Metric**	**Measure**
De-linked from HR-based financial incentives	Yes	Yes			
Target Driven			Partially*		
Action Driven			Yes		
Quantifiable				Yes	Yes
Frequently measured				Proposed**	
Continuously measured					Proposed**

(*) Single-value targets are not defined yet, (**) Data was exported upon request for this study and is not being frequently and continuously monitored in SBP

Additionally, used self-assessment checklist provided by the KPI framework [6] to evaluate our KPIs. Table V depicts the checklist for all the KPIs. As the KPIs do not have a single value target yet, hence this requirement is partially fulfilled. For this study, the metrics were calculated on request and currently are not monitored continuously, which is presented as a suggestion to the organization.

VI. Discussion

In this section, we delve into the interpretation of our observations and the expert opinions.

Energy-inefficient old hardware must be replaced. We discussed our observations about prominent patterns with the organization's storage expert. We observed very low capacity utilisation and capacity per watt ratio in most of the nodes of model AFF-A220 as well as a higher power consumption per TB of usable data in comparison to other models. We concluded that these nodes are relatively less energy efficient. Hence, we recommended the replacement of these nodes with newer and more efficient models such as the AFF-A700s which already exist in the environment. Discussion with the storage expert revealed that this inefficiency is caused by older hardware. Currently, older models are used for the organization's internal management purposes and not for the customer application environments. The storage expert mentioned that organization has a plan to discard old hardware and move the data to newer hardware for optimization and usage of fewer resources. This existing optimization plan made our recommendation relevant. However, it is important that the old hardware is removed in a sustainable manner.

Re-architecting based on application usage can improve capacity utilization. Among the clusters using the newer models (AFF-A700s and AFF-A800) we witnessed inefficient clusters with capacity utilization values close to zero. The expert explained that the architecture for these nodes is designed so that storage and compute nodes for one client are in the same rack. Some clusters have small data volumes as the application is focused on computational tasks. The storage expert mentioned that they are aware of this inefficiency issue. Most of the nodes have a capacity utilization between 30% to 40%, which was also depicted in our analysis. The expert further explained their plan to redesign the architecture. They aim to place all storage nodes in a single rack to increase overall capacity utilization. This will reduce resource usage and improve energy consumption. However, with continuous monitoring approaches, actions (such as, re-architecting) can be triggered and relevant processes accelerated.

Data management techniques must consider the sustainability impact during the decision-making process. We observed that in a noticeable number of clusters with AFF-A700s nodes, one node seemed to have a relatively higher capacity utilization compared to the other node in the same cluster. These nodes have active and passive roles. Data is only transferred to the passive node in the condition of failure. However, this is not the case for all the clusters. In this case, the expert mentioned that they need a more efficient way of data management where they have control over all the data and the nodes storing them.

The choice of space-saving techniques must consider the sustainability impact of chosen techniques and be prioritized appropriately. We observed that while snapshots had a significant role in space-saving, other techniques were never or rarely (cloning) used. In the discussion, it was explained that snapshots are used to avoid data loss in cases of failures. With regards to leveraging the other techniques, the expert mentioned that optimizing their data management approaches, as mentioned before, would help them benefit from other techniques more. He further explained that better control of the data management practices and a more precise overview of the location of the stored data would enable them to apply techniques such as de-duplication more efficiently across all their clusters. The sustainability impact of these space-saving techniques needs further analysis for its technical and economic impacts as well.

Optimization-based targets can be used till standards and regulations are mature enough to provide single value targets. After an informal discussion with the expert from the organization, we defined actions and partial targets for each KPI based on the discussion and the tactics from Vos *et al.* [19]. Setting concrete values of targets was challenging as the organization did not have data for setting targets on KPIs that involved power consumption. This is due to lack of available standards. Regarding capacity utilization KPI, the storage expert mentioned that best practices are applied in case of over-utilization (at around 80%). However, no value is set for under-utilization. Therefore, we defined partial targets to guide the organization towards taking necessary actions.

Action Plan. For the first KPI, energy consumption (kW/h) of different storage components must be continuously measured with the target of minimization. We identified that for power consumption (kW), AFF-A220 nodes are less efficient in comparison to other models and suggested their replacement. As a result, the action recommended for this KPI is to identify and replace or remove inefficient hardware devices in terms of energy consumption per usable GB of data. This action has already been planned for the nodes with the AFF-A220 model.

For the second KPI, the storage capacity utilization (%) of different components must be frequently measured to identify inefficient components or practices. The target should be optimizing this KPI as both over-utilization and under-utilization can lead to inefficiency of the data storage environment in terms of performance and energy consumption, respectively. Monitoring this KPI, we identified storage nodes with very low or 0% of storage capacity utilization. We suggested the action of resource de-allocation [19]. We also observed clusters with energy-efficient nodes (in terms of power consumption, *e.g.,* AFF-A700s nodes) being under-utilized. Later on, we discovered this was happening due to the architecture design used in the organization putting storage nodes and compute nodes together in similar racks. Therefore, the action suggested was a re-evaluation and re-design of storage infrastructure architecture for accomplishing higher levels of capacity utilization of energy-efficient nodes. This action has also been added to the plans of the data storage team in the organization. We observed that most of these inefficiencies were a result of data management practices that need improvement. The action suggested in this case was to optimize data management practices to achieve more efficient capacity utilization across all the nodes and clusters.

For our final KPI, the capacity per watt ratio should be frequently measured for different storage components. The target is to maximize the KPI value. It would lead to more units of data saved per unit of energy consumed. Monitoring this KPI, we witnessed very low values for the AF-A220 nodes. It was explained that this inefficiency was due to the nature of the application of these nodes and their old hardware. These nodes were only used for management tasks and the data stored on them was rarely used. The action recommended with regards to this KPI and from Vos *et al.* [19] was to compress infrequently accessed data (i.e. the data stored on clusters with AFF-A200 nodes). For maximizing the capacity per watt KPI, it is important to consider the trade-off of the required energy for compression. Applying this technique can eventually lead to increased energy efficiency in the storage environment.

VII. Challenges and Reflection

This section discusses the study's challenges and reflections.

Lack of unified guidelines or standards. Defining target values for KPIs was a challenging task due to the lack of clear standards in terms of power consumption. For metrics such as storage capacity utilization, the industry only provides best practices and no concrete values regarding under or over-utilization. This highlights the need for unified guidelines for consistency and transparency *e.g.,* through sustainability regulations.

Lack of transparency and unavailability of clear documentation is hindering research. We observed a lack of transparency about sustainability data and unclear documentation on the sustainability data from the cloud provider. This emphasizes the need for regulations on transparency as well as guidelines for providing reliable sustainability data by IT service providers. Collecting real-time power consumption data from NetApp proved challenging as the required data collectors were not suitable for all environments.

Hybrid-technique of developing KPIs can aid in balanced decision-making. Applying the KPI framework [6] with a hybrid top-down and bottom-up approach in this case study proved beneficial. Combining strategic planning with operational insights provides decision-makers with a balanced and informative perspective. Taking the available data from the organization into consideration is especially important in this case due to the challenges of gathering sustainability data in most organizations.

A cost-benefit analysis for KPI actions will add business value to the framework. While applying the KPI Framework was feasible and valuable, we observed the potential for improvement. Feedback from the organization indicated a practical interest in conducting a business-oriented cost-benefit assessment. Developing a detailed business case separately from defining the KPIs, focusing on proposed actions, can encourage industrial investment in sustainability initiatives.

VIII. Threats to Validity

In this section, we discuss the main threats to validity of the study and the steps taken to mitigate them.

Internal Validity. One internal validity concern in our study arises from the absence of real-time power consumption data, which may affect the precision of our findings regarding fluctuations over time. To mitigate this, we collected data at three different time points, finding consistent values, indicating the data represents typical power consumption. Additionally, we focused on power consumption per unit of storage (TB) to maintain independence from cloud storage usage and workload variations. We recognize the possibility of single researcher bias and mitigated it by conducting informal discussions with storage and sustainability experts within the organization, ensuring a diverse perspective. The data and analysis was reviewed by the second author. We also provide a replication package with data and code for transparency.

External Validity. We conduct this study in the context of one organization and data from NetApp. However, the metrics identified are independent from contextual information and serve the purpose of monitoring for energy efficiency. The KPIs are created based on the sustainability goal of the organization. However, they can be used by other organizations for their own sustainability goals by reusing the metrics. Organizations can tune the targets and actions based on their own goals and organizational strategy.

IX. Conclusion and Future Work

We conducted an exploratory study using data from the cloud data storage of SBP managed via NetApp. The study presents patterns in power consumption of different models of storage devices, insights into storage capacity utilization trends, and suggestions regarding space-saving techniques. Finally, the study defines three relevant KPIs, namely energy consumption (kW/h), storage capacity utilization (%), and capacity per watt (GB/Watt). The KPIs are introduced along with relevant actions, and targets based on the SBP's goals. These KPIs serve as a foundation for developing standard monitoring approaches for cloud data storage environments. The recommended actions that can provide insights for other industrial practitioners include hardware replacement, architectural redesign, and optimization of data management practices for enhancing overall energy efficiency.

Looking ahead, continuous monitoring trends over time could offer a more nuanced perspective on energy efficiency. This monitoring can aid organizations in making actionable decisions for sustainability. Additionally, expanding our scope to include other metrics, such as CO2 emissions or heat production, would provide a more comprehensive view of the environmental impact. The same methodology can be applied to develop KPIs for a cloud solution while including data storage along with other operational components. As a potential step for future work, a comparative analysis across multiple providers could reveal industry-wide patterns, showcase best practices, and contribute to the ongoing conversation about sustainable data storage.

Acknowledgments

We thank Schuberg Philis for their vital contribution of time, resources, and expertise, enabling this research.

References

[1] D. Mytton and M. Ashtine, "Sources of data center energy estimates: A comprehensive review," *Joule*, vol. 6, no. 9, 2022.

[2] E. Masanet, A. Shehabi, N. Lei, S. Smith, and J. Koomey, "Recalibrating global data center energy-use estimates," *Science*, 2020.

[3] V. D. Reddy, B. Setz, G. S. V. Rao, G. Gangadharan, and M. Aiello, "Metrics for Sustainable Data Centers," *IEEE Transactions on Sustainable Computing*, 2017.

[4] A. Rawson, J. Pfleuger, and T. Cader, "Green grid data center power efficiency metric: PUE and DCIE," Jan. 2008.

[5] J. Haas *et al.*, "A framework for data center energy productivity," tech. rep., Technical report, The Green Grid, 2008.

[6] I. Fatima, M. Funke, and P. Lago, "Providing Guidance to Software Practitioners: A Framework for Creating KPIs." Preprint., 2024.

[7] L. Pérez-Lombard, J. Ortiz, and D. Velázquez, "Revisiting energy efficiency fundamentals," *Energy Efficiency*, 2012.

[8] C. J. Corbett, "How Sustainable Is Big Data?," *Production and Operations Management*, vol. 27, Jan. 2018.

[9] A. Shehabi, S. Smith, D. Sartor, R. Brown, M. Herrlin, J. Koomey, E. Masanet, N. Horner, I. Azevedo, and W. Lintner, "United States Data Center Energy Usage Report," tech. rep., June 2016.

[10] A. Andrae, "New perspectives on internet electricity use in 2030," *Engineering and Applied Science Letters*, vol. 3, June 2020.

[11] L. Wang and S. U. Khan, "Review of performance metrics for green data centers: a taxonomy study," *The Journal of Supercomputing*, 2011.

[12] X. Shao, Z. Zhang, P. Song, Y. Feng, and X. Wang, "A review of energy efficiency evaluation metrics for data centers," *Energy and Buildings*, 2022.

[13] J. Haas, J. Froedge, J. Pflueger, and D. Azevedo, "Usage and public reporting guidelines for the green grid's infrastructure metrics (PUE/DCiE)," *White Paper*, vol. 22, 2009.

[14] M. Jamalzadeh and N. Behravan, "An exhaustive framework for better data centers, energy efficiency and greenness by using metrics," *Indian Journal of Comp. Sci. and Engineering (IJCSE)*, vol. 2, no. 6, 2012.

[15] G. Schulz, *The Green and Virtual Data Center*. Auerbach Publications, Jan. 2016.

[16] B. Schödwell, M. Wilkens, K. Erek, and R. Zarnekow, "Towards a holistic multi-level green performance indicator framework (GPIF) to improve the Energy Efficiency of data center operation - A resource usage-based approach," in *Electronics Goes Green 2012+*, 2012.

[17] F. Brocklehurst, "International Review of Energy Efficiency in Data Centres," 2021.

[18] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén, "Experimentation in Software Engineering," 2012.

[19] S. Vos, P. Lago, R. Verdecchia, and I. Heitlager, "Architectural Tactics to Optimize Software for Energy Efficiency in the Public Cloud," *2022 International Conference on ICT for Sustainability (ICT4S)*, 2022.