



Victoria Veselova

# Data Center Sustainability Reporting

Advancement of Assessment Methodology for Energy Consumption by Virtualized Resources in Data Centers

Metropolia University of Applied Sciences

Master's Degree

Degree Programme in Business Informatics

Master's Thesis

02.06.2023

## Abstract

Author: Victoria Veselova  
Title: Data Center Sustainability Reporting: Advancement of Assessment Methodology for Energy Consumption by Virtualized Resources in Data Centers  
Number of Pages: 94 pages  
Date: 2 June 2023

Degree: Master of Business Administration  
Degree Programme: Business Informatics  
Instructors: Antti Hovi, Senior Lecturer  
Zinaida Grabovskaia, PhD, Senior Lecturer

The objective of this study was to explore how the assessment methodology for energy consumption of virtualized resources in data centers, currently leveraged in the case company, can be advanced. The case company headquartered in Europe, provides IT services, consulting, and cloud services through its large-scale data centers. Data centers (DCs) of the case company need to provide sustainability reports to the customers and comply with Corporate Sustainability Reporting Directive (CSRD) entered into force in January 2023, which introduces the European Sustainability Reporting Standards (ESRS). While CSRD provides the guidelines how the companies should report on their environmental impact, the regulatory framework does not offer any standardized assessment methodologies for sustainability KPIs.

The current state analysis identified the lack of the standardized baseline for energy consumption measurements of virtualized resources in data centres of the case company, and the need for exploring additional metrics, which may impact on assessment for energy consumption of virtualized resources. The current state analysis revealed challenges on four stages of sustainability reporting process: Data Evaluation, Data Collection, Data Processing and Customer Reporting. The thesis focus was made on the stages related to the Data Evaluation and Data Processing, defined by the stakeholders as the most important areas for the improvement, where more research was needed.

The thesis used qualitative research methods including eleven in-depth both unstructured and semi-structured interviews with relevant stakeholders. The theoretical framework related to the topics of sustainability reporting and the Power Usage Effectiveness (PUE) metric in particular. The proposal was built in co-creation with key stakeholders, after a thorough analysis and evaluation by the experts, and it was immediately incorporated into the assessment methodology currently leveraged in the case company.

The outcome of this thesis is an enhanced assessment methodology for energy consumption of virtualized resources in data centers. The advanced methodology enables the case company to measure and monitor energy consumption, leading to more informed decision-making and the potential for optimizing resource allocation more accurately. The adoption of this methodology aligns with the case company's commitment to sustainability and reducing their environmental impact, showcasing their dedication to responsible and efficient data center operations.

Keywords: Data Center, Sustainability Reporting, Virtualized Resources, Energy Consumption, Methodology, Metric

## Contents

Glossary

List of Figures

1	Introduction	1
1.1	Business Context	2
1.2	Business Challenge, Objective and Outcome	3
1.3	Thesis Outline	4
1.4	Key Concepts	5
2	Method and Material	5
2.1	Research Approach	5
2.2	Research Design	7
2.3	Data Collection and Analysis	9
3	Current State Analysis of Sustainability Reporting Process	14
3.1	Overview of the Current State Analysis	14
3.2	Description of Current Sustainability Reporting Process	16
3.2.1	Corporate Sustainability Reporting Regulations and Standards	18
3.2.2	Data Evaluation	19
3.2.3	Data Sources	20
3.2.4	Data Collection	21
3.2.5	Data Processing	21
3.2.6	Customer Reporting	23
3.3	Key Findings from the Current State Analysis	24
3.3.1	SWOT analysis of the Current Sustainability Reporting Process	24
3.3.2	Key Challenges and Business Impact	27
3.4	Summary of the Current State Analysis Results	30
3.4.1	Selected Focus Areas	30
4	Existing Knowledge and Best Practice on Energy Consumption Assessment for Virtualized Resources in Data Centers	32
4.1	Environmental Sustainability Frameworks, Initiatives and Goals	32
4.2	Corporate Sustainability Reporting Directive	37
4.3	Data Center Energy Consumption	39
4.4	Energy Consumption Metrics for Virtualization Layer	44
4.5	Data Center Energy Efficiency Metrics	49

4.5.1	Defining Power Usage Effectiveness (PUE)	50
4.5.2	Power Usage Effectiveness (PUE) measurement levels	53
4.5.3	Critical Analysis of Power Usage Effectiveness (PUE)	57
4.5.4	Power Usage Effectiveness (PUE) complimentary metrics	59
4.6	Industry Standards and Best Practice	61
4.6.1	Google Cloud Carbon Footprint Reporting Methodology	62
4.6.2	Cloud Carbon Emissions Measurement and Analysis Tool	66
4.7	Conceptual Framework of Thesis	69
5	Building Proposal for Assessment Methodology of Energy Consumption by Virtualized Resources in Data Centers for the Case Company	72
5.1	Overview of the Proposal Building Stage	72
5.2	Findings from Data 2	73
5.3	Initial Proposal	75
5.3.1	Energy Consumption Metrics for Virtualization Layer	75
5.3.2	Data Center Energy Efficiency Metrics	76
5.3.3	Adoption of Industry Standards and Best Practices	78
5.4	Summary of the Initial Proposal	82
6	Validation of the Proposal	84
6.1	Overview of the Validation Stage	84
6.2	Final Proposal	85
6.3	Recommendations	87
7	Conclusion	90
7.1	Executive Summary	90
7.2	Thesis Evaluation	92
7.3	Closing Words	93
	References	94

## **Glossary**

ESG	Environmental, Social, and Governance
GRI	Global Reporting Initiative
SASB	Sustainability Accounting Standards Board
TCFD	Task Force on Climate-Related Disclosures
CDSB	Climate Disclosure Standards Board
CDP	Carbon Disclosure Project
IRRC	International Integrated Reporting Council
VRF	Value Reporting Foundation
EU	European Union
CSRD	Corporate Sustainability Reporting Directive
ESRS	European Sustainability Reporting Standards
SMEs	Small and medium-sized enterprises
KPIs	Key performance indicators
UNGC	United Nations Global Compact
SDGs	Sustainable Development Goals
GHG	Greenhouse Gas
CSA	Current State Analysis
BI	Business Intelligence

OSs	Operating Systems
EPD	Environmental Product Declaration
SFDR	Sustainable Finance Disclosure Regulation
CMDB	Configuration Management Database
SWOT	Strengths, Weaknesses, Opportunities and Threats
CPU	Central Processing Unit
RAM	Random Access Memory
EMS	Energy Management System
PUE	Power Usage Effectiveness
UPS	Uninterruptible Power Supplies
IT	Information Technology
PDUs	Power Distribution Units
IaaS	Infrastructure-as-a-Service
SaaS	Software-as-a-Service
PaaS	Platform-as-a-Service
I/O	Input/Output
VMs	Virtual Machines
VMM	Virtual Machine Monitor
PPE	Power to Performance Effectiveness

DCeP	Data Center Productivity
DCIE	Data Center Infrastructure Efficiency
CSPs	Cloud Service Providers
SBTs	Science Based Targets
NGOs	Non-governmental organizations
WRI	World Resources Institute
WBSCD	World Business Council for Sustainable Development
UNFCCC	United Nations Framework Convention on Climate Change
UNCCD	United Nations Convention to Combat Desertification
CBD	Convention of Biological Diversity
ISO	International Organization for Standardization
NFRD	Non-Financial Reporting Directive
VM	Virtual Machine
IPCC	Intergovernmental Panel on Climate Change
OS	Operating System
NAS	Network Attached Storage
SERT	Server Efficiency Rating Tool
PVAR	Power Variability
IPA	Infrastructure Power Adaptability

APC	Adaptability Power Curve
GUF	Grid Utilization Factor
WUE	Water Usage Effectiveness
SPEC	Standard Performance Evaluation Corporation
AWS	Amazon Web Services
GHGP	Greenhouse Gas Protocol
TTM	Trailing Twelve-Month
GCP	Google Cloud Platform
ERE	Energy Reuse Effectiveness
DC	Data Center

## List of Figures

- Figure 1 Research design of the thesis.
- Figure 2 Current sustainability reporting process in case company.
- Figure 3 Timeline for selected requirements and frameworks impacting case company.
- Figure 4 Overview of GHG Protocol Scopes and Emissions Across the Value Chain.
- Figure 5 Global enterprise technology emissions.
- Figure 6 The relation of CSRD to the European Green Deal, EU taxonomy, and SFD.
- Figure 7 Difference between IaaS, PaaS, and SaaS.
- Figure 8 Cloud computing market size in Europe from 2017 to 2030 (in billion EUR).
- Figure 9 The power consumption of a data center.
- Figure 10 Simple representation of a virtualized system.
- Figure 11 The formula for calculating Power Usage Effectiveness metric.
- Figure 12 PUE value calculation formula.
- Figure 13 pPUE computation formula.
- Figure 14 pPUE (zone) computation formula.
- Figure 15 A general scheme for the PUE computation.
- Figure 16 Overview of the subcomponents within a typical data center's facility and IT equipment.
- Figure 17 Graphical representation of the three PUE measurement levels for a typical data center.
- Figure 18 Number of contributors for the development of an application.
- Figure 19 Proposal Building steps.
- Figure 20 The presentation of key elements of the Proposal.
- Figure 21 Summary of the Proposal development process.

## 1 Introduction

Over the past 20 years, the disclosure landscape for Environmental, Social, and Governance (ESG) reporting has become increasingly complex. Multiple organizations, including the Global Reporting Initiative (GRI), Sustainability Accounting Standards Board (SASB), Task Force on Climate-Related Disclosures (TCFD), Climate Disclosure Standards Board (CDSB), Carbon Disclosure Project (CDP), International Integrated Reporting Council (IIRC), and Value Reporting Foundation (VRF), have created a variety of overlapping ESG reporting frameworks. Naturally, the need in common standards, the importance of consistency, transparency, and comparability has become evident. Entered into force on 5 January 2023, Corporate Sustainability Reporting Directive (CSRD) endorses the inclusion of non-financial metrics into corporate reports. (AuditBoard, 2022) CSRD replaces the European Union's Non-Financial Reporting Directive (NFRD). (Boston College Center for Corporate Citizenship, 2023) The CSRD introduces the European Sustainability Reporting Standards (ESRS), which will specify the information, format, and assurance of sustainability reporting. (PRI, 2022)

According to European Commission (2023), the CSRD regulations will establish a culture of transparency and guarantee that investors and other stakeholders will be able to obtain the necessary data to evaluate potential investment risks that emerge from climate change and other sustainability concerns. The new directive strengthens and extends the scope of the existing European Union (EU) reporting requirements, and it will be mandatory for approximately 50,000 large European companies and listed on the EU regulated markets small and medium-sized enterprises (SMEs). Additionally, under the CSRD, it is required that businesses conduct an audit of the sustainability data that they disclose, which bring sustainability reporting in line with financial reporting. (European Commission, 2023)

Companies that are already subject to the NFRD (based on their 2024 fiscal year data) will need to comply with the new requirements by January 2025. All other companies, excluding those already subject to the NFRD, will have until January 2026 to comply with the new requirements, based on their 2025 fiscal year data. Listed small and medium enterprises (SMEs) that request an extension will have until January 2027 to comply with the new requirements, based on their 2026 fiscal year data. (Onetrust, 2023)

The case company of this thesis is a subject to the CSRD requirements and needs to incorporate new reporting practices already in 2023. While CSRD provides the guidelines how the companies should report on their environmental impact, the regulatory framework does not offer standardized assessment methodologies for sustainability KPIs, according to the stakeholders from the case company. As a result, organizations face challenges in accurately evaluating and measuring the performance and reporting on the environmental impact.

This thesis research investigates the current practices of the company's sustainability reporting, established in the case company, analyses versatile challenges occurring on different stages of the reporting process and specifically assists in the improvement of practices on Data Evaluation and Data Processing stages. This study aims to advance the methodology for energy consumption assessment of virtualized resources in data centers. Data centres must accurately define all metrics that have direct or indirect impact on key performance indicators (KPIs) to produce clear and meaningful reports and accordingly consult customers, using virtualized resources, how to identify the opportunities for the improvement of their sustainability figures.

## 1.1 Business Context

The case company of this thesis is headquartered in Europe and operates on a global scale. It is a technology company, which serves thousands of enterprise and public sector customers. It offers IT services, IT consulting and cloud services, which are provided by large-scale data centers. The data centers owned by the case company run on 100% renewable energy, which is carbon-free. The company places high demands on sustainability and continuously works towards the reduction of climate impact. It aligns sustainability practices with international standards. Today the case company affiliates itself with several voluntary industry organizations and initiatives, such as Carbon Disclosure Project (CDP), United Nations Global Compact (UNGC), The European AI Alliance and others and contributes to United Nations Sustainable Development Goals (SDGs). Already in 2024, it will be mandatory for the case company to comply with Corporate Sustainability Reporting Directive (CSRD) regulations. The case company does not only intend to meet the requirements for corporate sustainable reporting standards, set by the regulators, but feels the strong need to go beyond the expectations of the stakeholders and assist in the improving of the environmental posture of its customers, by offering granular consulting services. The Sustainability Development Goals (SDGs) of the case company were selected through prioritization and are in line

with the guidance provided by the Global Reporting Initiative (GRI) and United Nations Global Compact (UNGC).

## 1.2 Business Challenge, Objective and Outcome

Currently, customers and other stakeholders increasingly expect the case company to report on their sustainability performance, requesting energy consumption and carbon emissions figures. Meanwhile the business is soon required to comply with the recently entered into force Corporate Sustainability Reporting Directive (CSRD). According to Ecochain (2022), the CSRD demands that companies report on their ESG metrics in a comprehensive manner, which is not yet familiar to the businesses. The quality of sustainability information is required to be of high standards. Data collection and reporting is expected to be robust, verifiable, and traceable, as external audit will be mandatory for the CSRD. Additionally, the CSRD requires Greenhouse Gas (GHG) protocol reporting for company emissions, including Scope 1, 2 and 3 emissions.

Since the clients of the case company will have to publicly disclose detailed and transparent information on how sustainability issues affect their business and what impacts they have on the society and the environment, the business intends not only to ensure the constant access to accurate environmental footprint figures, but to consult the clients on the potential for energy consumption and carbon emissions reductions. More accessible, in-depth, and verifiable non-financial data will allow customers to improve ESG related decision-making process and will help them to become more attractive to the investors and other stakeholders.

The company needs to stay in control with its own data, monitor improvements, and make footprint comparisons, as well as ensure that the best calculation methodologies for the key performance indicators (KPIs) are applied into practice. The case company was proactive and started more than a year ago the development of the new sustainability reporting and analytical solution. Regardless the solid baseline, the new service is still under the development, and it requires new competences and testing of the methodologies.

The main business challenge is related to the new corporate reporting requirements and associated calculating methodologies, which are novel in nature and not standardized by the regulatory bodies. The case company suggests that there is still room for the

refinement in the assessment methodology, leveraged for energy consumption calculation of virtualized resources in data centers, and it aims to ensure that all factors influencing the KPI are considered in calculations.

*The objective of this thesis is to improve the currently leveraged calculation methodology (focusing on the metrics which can impact on the assessment of the energy consumption by virtualized resources in data centers).*

The outcome of this thesis is the advanced energy consumption assessment methodology for virtualized resources in data center (focusing on the metrics which can impact on the assessment of the energy consumption by virtualized resources in data centers).

### 1.3 Thesis Outline

The scope of this thesis is to develop an advanced assessment methodology of the energy consumption by virtualized resources in data centers. This thesis will study theoretical and practical findings and best practices in use. The metrics impacting energy consumption by virtualized resources in data centers will be introduced in this thesis and the final conceptual framework will be built upon the findings. This thesis is divided into seven sections that follow a logical sequence to provide a comprehensive understanding of the topic.

Section 1, dedicated to the introduction, sets the context for the study, and emphasizes the importance and the urgency of the research, outlining the business challenge, objective, and outcome of this work. In Section 2, the method and materials are presented, detailing the research approach and design, data collection and analysis procedures. Section 3 provides a current state analysis of the existing practices in sustainability reporting. In Section 4, existing knowledge, and best practices on the assessment of the energy consumption by virtualized resourced in data centers are reviewed. Section 5 is introducing an initial proposal of the conceptual framework for the assessment of the energy consumption by virtualized resources in data centers. In Section 6, the proposed conceptual framework is validated by the case company. Finally, in Section 7 is summarizing the findings, offers thesis evaluation, as well as steps towards the implementation of the proposed conceptual framework and further recommendations.

## 1.4 Key Concepts

**Virtualized resources** in data centers refer to the abstraction of physical computing resources such as processors, memory, and storage, into multiple virtual environments that can be allocated to various applications and users. This technology allows for more efficient and flexible use of resources, as multiple virtual environments can run on a single physical server. Virtualization also enables dynamic resource allocation, allowing resources to be provisioned and deprovisioned as needed.

**Sustainability reporting** of data centers involves transparently measuring and communicating the environmental, social, and economic impacts of their operations. It includes aspects like energy efficiency, carbon footprint, resource management, environmental impact, social responsibility, and performance metrics. By providing this information to customers, data centers support informed decision-making and encourage sustainability practices. Through transparency and accountability, data centers can drive improvements in energy efficiency, reduce energy consumption, carbon footprint, and promote sustainable operations, contributing to a more environmentally responsible digital infrastructure and impact on sustainability posture of its customers.

## 2 Method and Material

This section introduces the methods utilized for research approach, research design, and data collection and analysis. The reader will gain an understanding of how the research was conducted.

### 2.1 Research Approach

According to Sreejesh et al. (2014: 4), there are two categories of business research: *basic research* and *applied research*. Basic research is conducted to expand knowledge and understanding of a particular field or phenomenon or concept, primarily for the sake of increasing knowledge rather than solving practical problems. On the other hand, applied research is conducted to discover the applications and uses of theories, knowledge, and principles in actual work or in solving problems. (Sreejesh et al., 2014: 4)

This study belongs to *applied research* family, as this type of research aims to solve practical problems and involves testing the effectiveness of existing theories and methodologies. The findings of applied research will be used for product development.

In this thesis, an *action research* approach is selected, as it is focused on solving immediate real-world problems and involves collaboration between researcher and practitioners. "Action research is intended to combine a strong and rigorous research activity with a respect for participants' knowledge and understanding. It therefore brings together theory and practical knowledge, to test each other with the purpose of developing practice". (Coleman, 2007: 484-485)

Data collection is an important stage of this approach and is typically done through a variety of methods, including surveys, focus groups, and interviews. However, it is essential to identify the type of data that is required for the research, before choosing the method of the collection. According to Toherdoost (2021:11-13), *qualitative data* as non-numerical data that is descriptive and cannot be shown as numbers. It answers "how and why" questions and covers data regarding feelings, perceptions, and emotions. It is collected through unstructured approaches such as interviews. *Quantitative data* is described as numerical data that is mathematically generated and computed. It is collected through structured approaches such as experiments and structured interviews. Data collection can be classified into two major categories: *primary data collection*, which involves obtaining first-hand information that is not yet published, and *secondary data collection*, which involves gathering data from published sources. The most popular types of primary data collection include interviews, focus group interviews, surveys, case-studies, questionnaires, and other techniques.

There are plenty of popular methods of secondary data collection, such as: general websites, published electronic sources, published printed sources, journals, e-journals, government records and many others. To provide a more comprehensive understanding of the research topic, qualitative data will be collected with a combination of research techniques commonly used in applied research. Both interviews, as the primary source of data and literature reviews, as the secondary source of data, will be collected for this research.

Two types of interviews are chosen for this thesis: unstructured and semistructured. According to Leavy (2014, 286), unstructured interviews have little pre-set structure and

are chosen when aspects of the discussion cannot be known to the researcher in advance and emerge in the course of spending time with the interviewee, while during semistructured interviews the researcher can have a greater control over the focus of the conversation on most important issues in the relation to the research project and can become visible as a knowledge producing participant in the process. The information obtained from the interviews will be recorded in the field notes. (Leavy, 2014: 286)

## 2.2 Research Design

This thesis follows a research design that includes five steps. The steps follow a specific order in the research process, where each subsequent step builds upon the previous one. Additionally, there are three distinct phases of data collection - Data 1, Data 2, and Data 3. Data collections act as inputs for the related steps, supporting the research design and producing the corresponding outputs.

Figure 1 below illustrates the research design of the study.

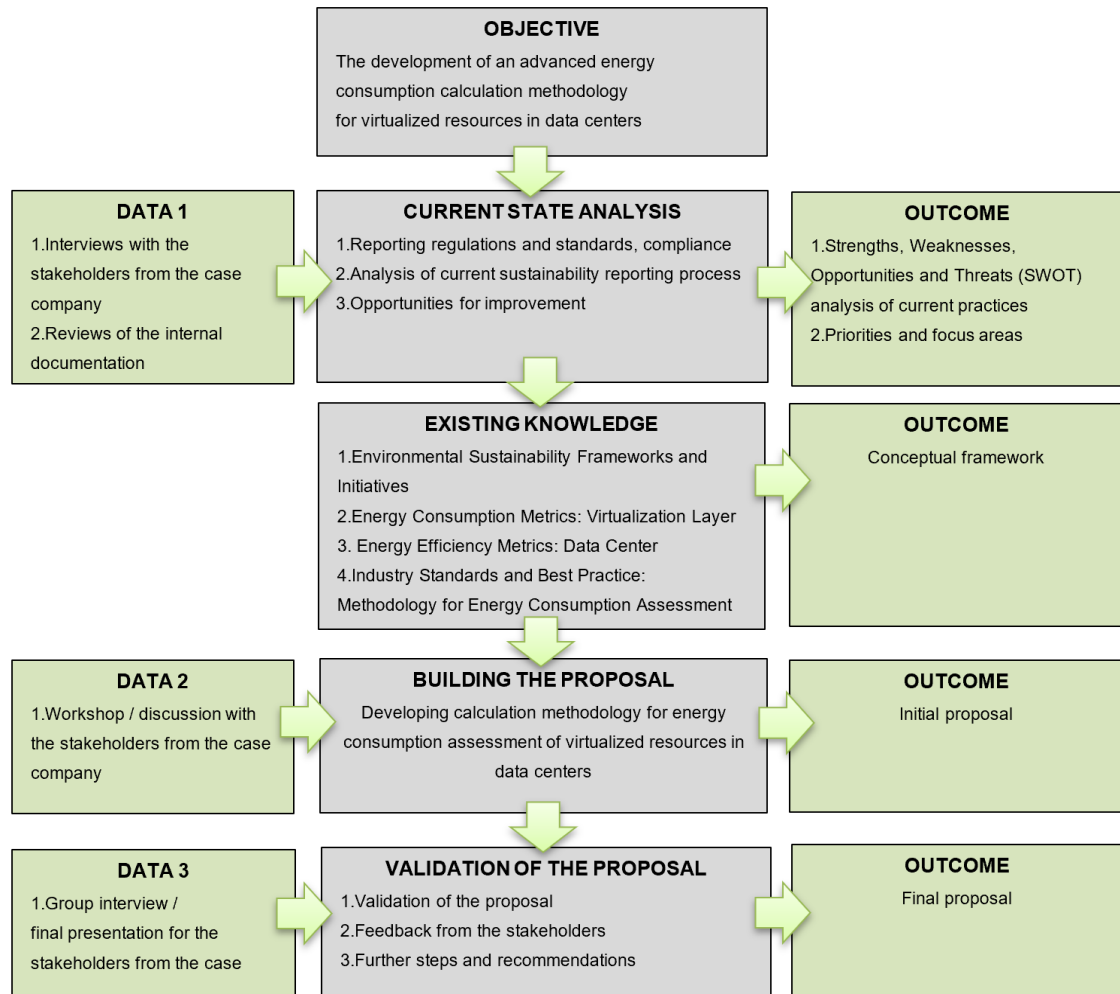


Figure 1. Research design of the thesis.

As seen from Figure 1, the objective of this thesis is set on the *the first step*, followed by the Current State Analysis (CSA) on *the second step*. The CSA is carried out based on Data 1 collection, which consists of interviews with the stakeholders, documented in the field notes, and internal documentation reviews. The aim of this step is to identify the requirements and regulations related to the reporting processes, identify the strengths and weaknesses in current practices of the reporting and choose the focus of the studies.

*The third step* of the research design focuses on exploring existing knowledge. The investigation includes best practices currently available on the assessment of energy consumption by virtualized resources in data centers and topics for literature reviews, most relevant for the study and useful for including on the stage of building the proposal. The outcome of the third step is a conceptual framework for the advanced calculation methodology.

The goal of the *fourth step* is to develop a proposal for an advanced calculation methodology. The combination of the data gathered in the second phase with the findings from step 3, provides the necessary input for the creation of the proposal. The step will involve gathering additional information in Data 2 collection, which includes conducting group interviews and documenting the insights in the field notes. The proposal serves as the initial draft for the suggested calculation methodology and is based on the outcomes of the current state analysis, relevant best practices, existing knowledge, and the input received from the stakeholders from the case company.

The validation of the proposal occurs on *the fifth step* of the research design, when the stakeholders of the case company verify the outcome of the thesis. Data 3 collection includes group interview with key stakeholders, documented in field notes. The proposed calculation methodology is reviewed by the key stakeholders, who approve it or otherwise suggest final modifications and updates.

### 2.3 Data Collection and Analysis

This research incorporates a diverse range of data sources, which were gathered through multiple rounds of data collection.

Table 1 provides a comprehensive overview of the Data collections 1-3 utilized in this study.

Table 1. Details of Data collections 1-3 used in this study.

#	Participants / role	Data type	Topic, description	Date, length	Documented as
<b>Data 1: Current State Analysis</b>					
1	Respondent 1: Enterprise Architect	Interview via Microsoft Teams application	The case company open innovation (OI) initiative and interview about current process based on the respondents' experiences	27 Jan 2023, 1 hour	Field notes
2	Respondent 2: Sustainability Manager Environment, Respondent 3: Head of Sustainability	Group Interview via Microsoft Teams application	Interview about current process based on the respondents' experiences	3 Feb 2023, 60min	Field notes

3	Respondent 3: Head of Sustainability Respondent 4: Service Owner, Sustainability Hub	Group Interview via Microsoft Teams application	Interview about current process based on the respondents' experiences	10 Feb 2023, 30min	Field notes and recording
4	Respondent 3: Head of Sustainability Respondent 4: Service Owner, Sustainability Hub	Group Interview via Microsoft Teams application	Interview about current process based on the respondents' experiences	22 Feb 2023, 30min	Field notes and recording
5	Respondent 3: Head of Sustainability	Interview via Microsoft Teams application	Interview about current process related to the respondent's experiences	23 March 2023, 1 hour	Field notes and recording
6	Respondent 3: Head of Sustainability	Interview via Microsoft Teams application	Interview about current process related to the respondent's experiences	27 March 2023, 3 hours	Field notes and recording
7	Respondent 2: Sustainability Manager Environment	Interview via Microsoft Teams application	Interview about current process related to the respondent's experiences	17 April 2023, 30min	Field notes
8	Respondent 5: Lead Project Manager	Interview via Microsoft Teams application	Interview about current process related to the respondent's experiences	17 April 2023, 40min	Field notes and recording
9	Respondent 6: Analytics and Reporting Lead	Interview via Microsoft Teams application	Interview about current process related to the respondent's experiences	27 April 2023, 30 min	Field notes and recording
<b>Data 2: Building the Proposal</b>					
10	Participants 3: Head of Sustainability Participant 4: Service Owner, Sustainability Hub	Workshop/ discussion via Microsoft Teams application	Proposal building	19 May 2023, 30min	Field notes and recording
<b>Data 3: Validation of the Proposal</b>					
11	Respondent 4: Service Owner,	Final Interview	Validation, evaluation of the Proposal	26 May	Field notes and recording

	Sustainability Hub	via Microsoft Teams application		2023, 30min	
--	--------------------	---------------------------------	--	-------------	--

As seen from Table 1, data for this Thesis was collected in three rounds.

The first round, collecting Data 1, was conducted for the current state analysis, and included interviews with various stakeholders from different business units and in different positions: Enterprise Architect; Sustainability Manager Environment; Head of Sustainability; Service Owner, Sustainability Hub; Lead Project Manager; Analytics and Reporting Lead. Each interviewed respondent was a participant of a common project for the development of sustainability reporting and consulting solution. For better understanding of the current state of the sustainability reporting processes, the internal documents were reviewed.

In the next round, Data 2 was collected to gather suggestions from the case company sustainability unit for developing the proposal. This set of data was gathered to receive information on any new or overlooked requirements, and to get feedback on the original proposal. The group interview, recorded in the field notes, was the main method of collecting this data. The two major stakeholders who were interviewed: Head of Sustainability and Service Owner, Sustainability Hub.

In the third round, Data 3 was collected in field notes when conducting group interviews for validation of the initial proposal. Data 3 included feedback for the proposal from the case company sustainability unit. This round was used for collecting the final validation and approval of the proposed advanced calculation methodology for virtualized resourced in data centers. During the validation phase, there were discussions held with the primary stakeholders about the implementation and potential consequences for the organization. This round also led to the identification of final improvement recommendations from the key stakeholders who were interviewed for the validation process: Head of Sustainability and Service Owner, Sustainability Hub.

As it is seen from the description of the Data collection rounds, in this study, interviews were the primary method of data collection. The interviews were carried out in an unstructured and semi-structured formats, the meetings with the respondents were held via Microsoft Teams application with part of the questions created in advance and with an open-end approach, in which open-ended questions were asked, allowing the

interviewees to respond in their own words and provide detailed, personalized answers. Open-end interviews helped to gather rich, detailed data on stakeholder's experiences, opinions, and attitudes towards existing working practices, and opened complex and sensitive issues. The follow-up questions were also used to further clarify the interviewee's responses.

The interviews were recorded, and the field notes taken, but this information will not be shared publicly, as this study is subject to confidentiality agreement with the case company.

Data 1 collection also included the review internal documentation, such as project's background information, numerical calculations, sustainability reports, sustainability regulations and requirements and prototypes of business intelligence (BI) dashboards.

The data was analyzed with thematic analysis, which is used for various types of qualitative data and helps to identify and report on patterns within data.

The internal documents included in the study are presented in Table 2.

Table 2. Internal documents used in the current state analysis, Data 1.

#	Name of the document	Number of pages/other content	Description
A	Sustainability Calculation.xlsx	10 pages	Numerical calculations
B	X Sustainability Customer Team.pptx	28 pages	Project background information
C	X Sustainability Customer.xlsx	1 page	Numerical calculations
D	Report Prototype 1.1	1 Business intelligence (BI) dashboard	Energy consumption and CO <sub>2</sub> e measurements
E	Report Prototype 1.2	1 Business intelligence (BI) dashboard	Hardware/IT equipment measurements
F	Report Prototype 2	1 Business intelligence (BI) dashboard	Energy consumption and CO <sub>2</sub> e measurements
G	Calculations 2022.xlsx	1 page	Numerical calculations

H	Virtual analyze.xlsx	1 page	Numerical calculations
I	Consumption_Year2022.docx	2 pages	Energy consumption report for the customer
J	Susytainability_Report_2022.pdf	92 pages	Case company own annual report 2022
K	Case Company Handprint Guidebook.pptx	47 pages	Carbon handprint assessment guide
L	Energy consumption 2023.docx	9 pages	Monthly report for the customer
M	Energy consumption 2022-2023.docx	15 pages	Semi-annual report for the customer
N	Sustainability Hub app model. ptpx	1 page	Product hierarchy and calculation methods description
O	Sustainability Hub app model. ptpx	1 page	Report structure and KPI categorization description
P	Customer's Landscape 2023	1 Business intelligence (BI) dashboard	Energy consumption and carbon emissions

As seen from Table 2, this study also analyzed a number of internal documents. The main documents included numerical calculations and reports needed for the assessment of the existing calculation methodologies. The documents were analyzed for Data collection 1 round, the current state analysis, to get an understanding of the current sustainability regulations applied to the case company, existing sustainability reporting practices, the strengths and weaknesses of the leveraged approaches and opportunities for the improvement of the service or methodologies for data identification and collection for further assessment of the customers' environmental posture. The biggest part of data was analyzed for the current state analysis, to establish the current state of corporate sustainability reporting practices in the case company. The findings from the current state analysis are discussed in Section 3 below.

### 3 Current State Analysis of Sustainability Reporting Process

This section discusses and analyses the current state of data centers sustainability reporting process to the customers, established in the case company. The analysis is made to understand the strengths and weaknesses of the current practices to find the opportunities for improvement and choose the focus for this thesis. First, the study investigates the reporting regulations and standards the case company must comply with. Then it evaluates all five stages of the reporting process: *Data Evaluation*, *Data Sources*, *Data Collection*, *Data Processing* and *Customer Reporting*.

#### 3.1 Overview of the Current State Analysis

The objective of the Current State Analysis was to define the key elements that belong to the sustainability reporting process, identify strengths and weaknesses of the existing working practices, and find out the opportunities for the improvement, which lead to the choice of focus of this study. Unstructured and semistructured in-depth interviews were the primary data sources and all of them were conducted virtually via Microsoft Teams application, due to the different geographical location of the respondents. Additionally, internal documentation was reviewed. It ensured enough data for a comprehensive analysis of the current state, which was conducted in seven steps.

There are several different Business Divisions in the case company. On the *first* step the initial interview was conducted with the Enterprise Architect from the Business Division A. The invitation with the agenda was sent a week prior to the session. This approach ensured an effective discussion over emerging new sustainability reporting practices and related challenges in the Business Division A. The internal documentation was reviewed during and after the interview. After the first meeting, a short discussion with a Lead Business Consultant, who used to be involved in the same project related to corporate sustainability reporting, influenced the change of cooperation from Business Division A to Business Division B. Due to the nature of the discussion, it was not recorded in this study, as an in-depth interview.

On the *second* step, the set of 3 group interviews with the representatives from Business Division B, such as Sustainability Manager Environment, Head of Sustainability and Service Owner, Sustainability Hub, was conducted in a row to identify current state of corporate sustainability reporting process, opportunities for improvement, challenges,

stakeholder engagement and further cooperation with the chosen focus for the study. The field notes from all 3 interviews were taken, two last interviews were also recorded, as the meetings included the overview of the internal documentation, which was later additionally examined and profoundly analysed.

On the *third* step, a set of two group interview was conducted with the Head of Sustainability. The objective of these interviews was to gather more information about the current state of corporate sustainability reporting process. Weaknesses, strengths, and opportunities for the improvement were also identified. The emphases in the discussions were also made on the choice of scope of this thesis. The current methodology for the assessment of energy consumption of virtualized resources in data centers was carefully studied. Both interviews were recorded for later profound analysis and field notes were taken. The second interview also included the internal documentation review.

On the *fourth* step, an interview with Sustainability Manager Environment was carried out to find out more about reporting guidelines, regulatory compliance and current practices on energy consumption and carbon footprint reporting. The internal documentation was reviewed during and after the meeting. The interview was recorded, and the field notes were taken.

On the *fifth* step, an interview with the Lead Project Manager was conducted to gain more insights to the current process of corporate sustainability reporting and another perspective on the strengths, weaknesses, challenges, and opportunities for development. The internal documentation was reviewed during and after the interview, which was recorded. Field notes for further analysis were taken as well.

On the *sixth* step, the last interview was organized with Analytics and Reporting Lead, who shared his vision on the current corporate sustainability reporting process and gave more insights on business intelligence (BI) tool, leveraged for the automation of the reporting. The interview was recorded, the field notes were taken and internal documentation, presented in the interview was analysed afterwards.

The final, *seventh* step included the analysis of strengths and weaknesses of the current corporate sustainability reporting process and the designation of the study focus. All the findings were summarised.

### 3.2 Description of Current Sustainability Reporting Process

The interviews with the employees in the case company gave a comprehensive picture of the current sustainability reporting process and regulatory requirements. The visualization of high-level sustainability reporting process can be seen in Figure 2.

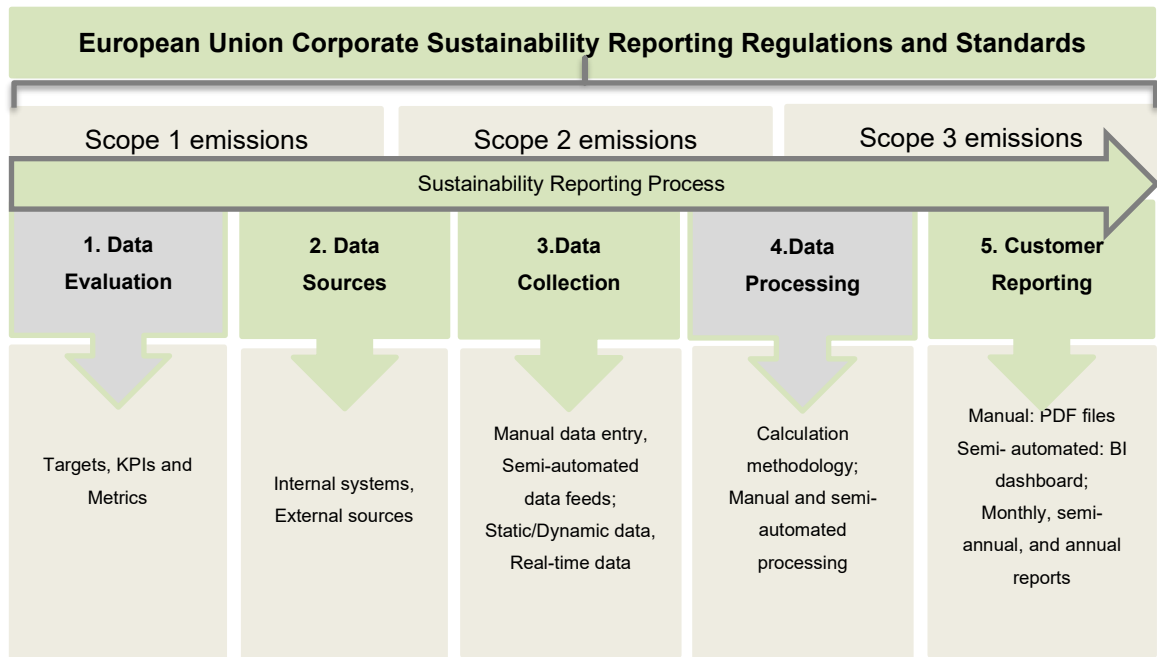


Figure 2. Current sustainability reporting process in the case company.

As seen from Figure 2, the current sustainability reporting process is tied with European Union (EU) corporate sustainability reporting regulations and standards, in which Scope 1, Scope 2, and Scope 3 emissions are widely used concepts. Different regulatory frameworks, which recognize the significance of measuring and disclosing emissions within each scope were investigated. These frameworks give guidelines and reporting requirements to ensure organizations provide transparent and comprehensive information on their direct and indirect greenhouse gas emissions, helping stakeholders assess their environmental performance and progress in reducing emissions. Additionally, the interviews with the employees revealed which Scopes the case company delivers the reports to their customers.

The sustainability reporting process itself is divided into five stages: Data Evaluation, Data Sources, Data Collection, Data Processing and Customer Reporting. This is illustrated in the following quote by the Respondent 4:

Together with the customers we check what kind of goals they want to set up, what kind of KPIs they want to use and what are their priorities. From here we dissect the KPI to find out what different variables are needed to fulfil the KPI. From there we find out where is the source of data. From here starts the system integration work, continuous collection of data. (*Respondent 4*)

On the *Data Evaluation* stage, the company assesses the relevant sustainability indicators and metrics that need to be reported. It involves identifying key performance areas, such as energy consumption and greenhouse gas emissions. The company determines which data points are important for measuring and monitoring their sustainability efforts.

The *Data Sources* stage involves identifying and gathering data from various sources within the company. These sources can include internal systems, such as accounting, operations in Data Centers, and HR databases, as well as external sources like suppliers, customers, and industry databases. The goal is to compile a comprehensive dataset that covers all the relevant sustainability metrics.

In *Data Collection* stage, the company collects the identified data from the different sources. This can involve using various methods, such as, automated, semi-automated data feeds, or manual data entry. Data collection also includes setting up monitoring systems to track real-time data, such as energy usage. Some data is still static, but it could be dynamic.

The *Data Processing* stage starts once the data is collected, it needs to be processed and organized for analysis. This stage involves cleaning the data, ensuring accuracy and consistency, and performing necessary calculations and conversions. Data may be normalized, standardized, or transformed to facilitate meaningful analysis and comparisons across different periods or units. Currently part of the data is processed manually, and part of the analysing process is semi-automated. The calculation methodology is under the development.

The final stage in the process is *Customer Reporting*, which focuses on communicating the collected and processed data to the customers. Sustainability reports are prepared to summarize the customer's performance in a transparent and comprehensible manner. These reports may include narratives, charts, graphs, and other visual aids to enhance understanding and engagement with the sustainability efforts. Part of the reports is currently done manually, and some are semi-automated. Two prototypes of the reports

are created in business intelligence (BI) reporting and analytical tool, the visualization of data helps to communicate complex information in a more accessible and engaging manner. For today, the reports for the rest of the customers are created manually in Microsoft Word and Excel files. The corporate sustainability reporting regulations and standards, as well as sustainability reporting stages are described and analysed in detail below.

### 3.2.1 Corporate Sustainability Reporting Regulations and Standards

The case company makes sustainability efforts not only to comply with laws and regulations, but it is determined to be an ethical leader in the technology industry and takes a proactive stance, monitoring changes in the requirements and frameworks that may in the future come into effect and impact the company, planning the implementation of the necessary changes in the reporting practices well in advance, which also ensures the premium quality of the service provided to the customers.

Figure 3 below illustrates timeline for selected requirements and frameworks impacting case company.

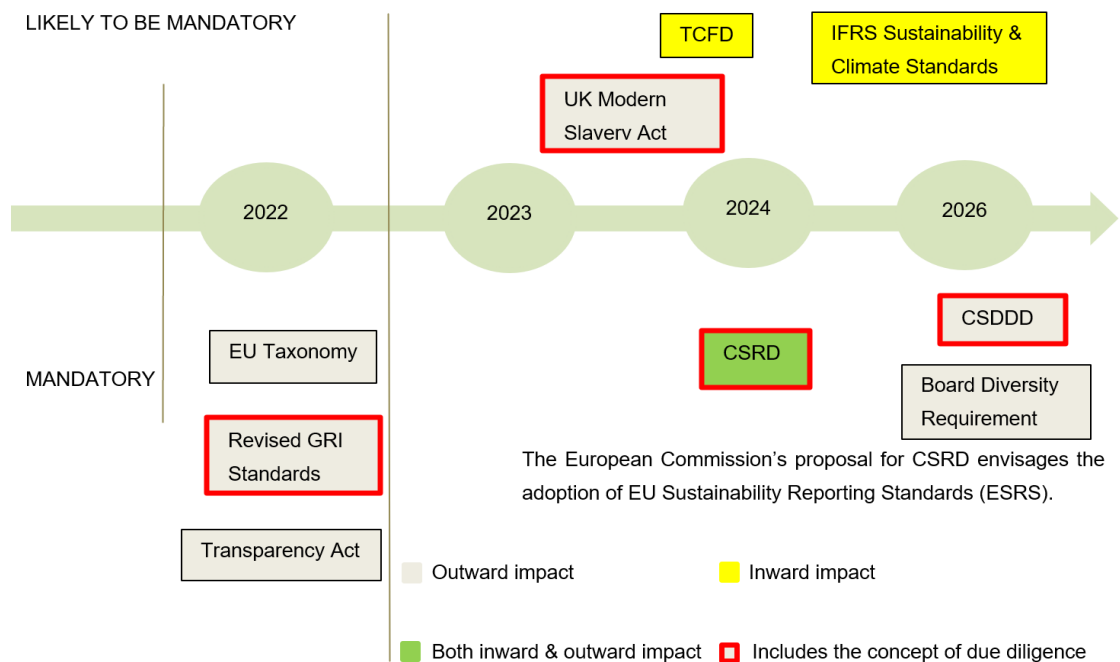


Figure 3. Timeline for selected requirements and frameworks impacting case company. (The Figure was derived from internal documentation and adapted)

As it is seen from Figure 3, Corporate Sustainability Reporting Directive (CSRD), which entered into force in January 2023, has both inward and outward impact on the case company, which means that the directive affects both the internal operations and reporting practices of the case company, as well as the external disclosure and transparency requirements imposed by the directive. It will require the case company to enhance its data collection, measurement, and reporting systems to comply with the new sustainability reporting standards, which will lead to extensive changes in terms of reporting and business tactics. The case company started to prepare and develop new reporting practices back in 2022, when first customers began to ask for own energy consumption and carbon footprint emissions figures, as it was stated in one of the internal documents and mentioned during the group interview by the Respondent 2.

The case company is providing to its customers the reports on both Scope 2 and Scope 3 emissions and is also measuring the emissions produced by the end-user devices. This is illustrated in the following quote by Respondent 4:

We also collect data from other suppliers for the specific customers. We show together these figures on one dashboard. Based on energy consumption, based on different variables, we can also show emissions from other suppliers.  
*(Respondent 4)*

The case company is disclosing information related to indirect emissions from purchased electricity (Scope 2) and emissions from value chain activities, such as transportation, waste management, or business travel (Scope 3) and assessing and tracking the emissions resulting from the use or operation of their products or services by end-users. The case company is actively monitoring and reporting emissions across different scopes, including both upstream and downstream activities.

While describing reporting procedure for one of the customers during the group interview, Respondent 4 stated “For this customer we are measuring IT equipment even in larger way: laptops, workstations, monitors, mobile phones, printers, tablets.”

### 3.2.2 Data Evaluation

During the data evaluation stage, the case company engages in a comprehensive dialogue with its customers to identify their specific sustainability expectations and requirements. This collaborative process helps the company understand which Key

Performance Indicators (KPIs) and metrics are most relevant to its stakeholders and aligns the reporting process with customer expectations. These discussions enable the organization to incorporate customer perspectives into its sustainability objectives and ensure that the selected indicators accurately reflect the company's performance in areas of interest to its customers. However, this engagement process can present challenges, such as managing diverse customer expectations and ensuring that the selected indicators are meaningful and measurable within the company's operations. Striking the right balance between customer priorities and the feasibility of data collection is crucial to generating reliable and insightful sustainability reports. Additionally, choosing the appropriate metrics could be a complex and challenging task. It requires careful consideration of various factors. The challenge lies in selecting metrics that accurately reflecting the customer's sustainability performance while also aligning with industry standards and reporting frameworks, which evolving landscape and emerging trends add to the complexity, requiring companies to stay updated and adapt their metrics accordingly to effectively communicate their sustainability efforts.

### 3.2.3 Data Sources

The main source of data, leveraged in calculations for reporting KPIs, is Configuration Management Database (CMDB).

Invoicing system is the other source of the data, needed for volume-based calculations, when the purchased volume of Central Processing Unit (CPU) or Random Access Memory (RAM) is converted to the energy consumption figures.

Energy Management System (EMS) is the source of Power Usage Effectiveness (PUE) values, which are specific to every data centre.

Environmental Product Declaration (EPD) covers electricity generated by Vattenfall's wind farms, located in Sweden.

### 3.2.4 Data Collection

For the part of the customers the data is collected only once a year retrospectively, for the others semi-annually or monthly. Some data is collected manually or in a semi-automated way, which is still time-consuming. It is positive that the case company can have a dynamic pull out of the total number of servers in data centers every month, based on location, also total amount of energy consumption in each data center respectively is available monthly.

The CMDB is populated automatically, but data needed for the reports is collected from CMDB manually, according to the timeframe requested by each customer, yearly, once in six months or monthly.

Invoicing data is reported automatically to the invoicing system but must be reviewed manually to align the purchased volumes of CPU or RAM with the energy consumption figures. Customers' invoicing information is available monthly.

Based on Environmental Product Declaration (EPD) certificates for data centers, emission factors are collected in cases, when renewable energy source is selected (market-based energy). EPD certifications provide information about the environmental performance of data centers, particularly in terms of their energy carbon intensity.

### 3.2.5 Data Processing

Once the data is collected, the subsequent stage involves processing and analysing it to derive meaningful insights. This crucial step entails aggregating, validating, and analysing the data to calculate relevant metrics and indicators. Challenges encountered during this stage may include data normalization and standardization, managing data inconsistencies, efficiently handling large data sets, and addressing any potential biases or limitations in the data. Overcoming these challenges ensures that the processed data provides reliable and accurate information for decision-making and reporting purposes.

Currently, data extraction is performed manually on an ad-hoc basis, which poses challenges for the organization. Manual data processing can hinder the ability to respond promptly to changes in the business environment, limiting informed decision-making and

adaptability to new challenges and opportunities. Moreover, as data volumes increase, manual processing becomes increasingly demanding and resource-intensive, making it difficult to scale operations to meet growing demands. Inconsistencies may arise in data collection, analysis, and reporting, leading to potential inaccuracies. The time-consuming and labour-intensive nature of manual data processing also slows down operations and reduces overall efficiency, while introducing a higher risk of human error.

Product specific energy factors are currently leveraged in calculating methodology. This helps to visualise consumption to product level. Customers purchase various products, like physical and virtual capacity.

The case company's methodology for assessing energy consumption of virtualized resources in data centers was evolving with time and preliminary was based on purchased volume of Central Processing Unit (CPU) or Random Access Memory (RAM). Product owners were consulting whether CPU or RAM should be used as a baseline for the calculations, depending on the case. However, the practice showed that the CPU volume does not allow to see the trend, while RAM's pattern of consumption can be detected and measured hourly. Currently billing data is analysed manually to determine the amount of RAM, or the volume of storage purchased. The usage and cost data are converted into energy consumption figures. These energy consumption figures are then used to estimate the associated carbon emissions. This process involves quantifying the energy consumed by the virtualized resources based on their allocated CPU, RAM, and storage capacities, as discussed earlier. The case company provides reports to its customers on both location-based and market-based emissions resulting from the usage of virtualized resources.

CPU usage is typically measured in terms of computing capacity or processing power. The unit commonly used to quantify CPU usage is the number of cores or threads. A core represents an individual processing unit within a CPU, while a thread refers to a sequence of instructions that can be executed independently by the CPU. CPU usage is calculated based on the allocated cores or threads to virtual machines or other virtual resources. RAM, or memory, is measured in units such as bytes (B), kilobytes (KB), megabytes (MB), gigabytes (GB), or terabytes (TB). These units represent the amount of data that can be stored and accessed by the virtualized resources. RAM usage is calculated based on the allocated memory capacity to virtual machines or other virtual resources. Storage capacity in data centers is typically measured in units such as bytes

(B), kilobytes (KB), megabytes (MB), gigabytes (GB), terabytes (TB), or petabytes (PB). These units represent the amount of data that can be stored on the storage devices, such as hard drives or solid-state drives. Storage usage is calculated based on the allocated storage capacity to virtual machines or other virtual resources.

Location-based and market-based methods are Scope 2 accounting methods, which involve the allocation of emissions from electricity generation to end-users. The location-based method calculates emissions based on the average emissions intensity of the grids where energy consumption takes place. This method relies on grid-average emission factor data. On the other hand, the market-based method considers emissions from the specific electricity sources chosen by companies (or used due to the lack of choice). (Greenhouse Gas Protocol, 2015)

Location-based reporting provides a more accurate representation of the environmental impact associated with energy consumption.

In the case of market-based emissions, emission factors are used in the calculations to account for the environmental impact of the chosen energy sources, particularly when renewable energy is selected.

### 3.2.6 Customer Reporting

Two pilot customers are receiving the reports via business intelligence (BI) analytical and reporting tool. They have access to the dashboards, where they can see the information in a trend view year over and month over month. Additionally, energy consumption figures could be filtered on the service base. There is an opportunity to improve the granularity of the reports and offer the filtering not only on the types of services, but also used Operating Systems (OSs) or different cloud services. Both Scope 2 and Scope 3 emissions are covered in the reports. The other part of the reports is done manually on an ad-hoc basis, which requires a lot of efforts and time and may cause human errors. These reports are sent to the clients in Microsoft PDF files via emails. Such reports include historical data and information. Some of them are generated annually, some semi-annually, and the others monthly. These reports do not currently have any strict structure, they are not standardized.

### 3.3 Key Findings from the Current State Analysis

As seen from the description above, the current state of sustainability reporting process, established within the case company is now undergoing big changes. Manual, semi-manual and ad hoc actions will soon be automated, reporting frameworks will be standardized, and assessment methodologies potentially enhanced. More detailed conclusions are summarized below in the SWOT analysis.

#### 3.3.1 SWOT analysis of the Current Sustainability Reporting Process

The analysis of Data Collection 1 revealed the strengths and weaknesses of current sustainability reporting process. Opportunities and threats were also highlighted. The results of Strengths, Weaknesses, Opportunities, and Threats (SWOT) analysis are summarized in Table 3 below.

Table 3. SWOT analysis of the current sustainability reporting process.

Strengths	Weaknesses
<ul style="list-style-type: none"> <li>• Adherence to regulations and a proactive stance towards new reporting requirements</li> <li>• Good knowledge of data sources</li> <li>• Reporting both GHG emissions and energy consumption</li> <li>• Reporting customers' both Scope 2 and Scope 3 emissions</li> <li>• Profound understanding of customer needs</li> <li>• All necessary resources for offering a complete reporting and consulting solution,</li> <li>• Emissions from other suppliers could be also tracked, calculated, and reported.</li> <li>• Strong competencies in the field of Sustainability</li> <li>• Good understanding of data sources</li> <li>• Support for decision making from product owners</li> <li>• Ability to align customers' consumption with the invoice volume</li> </ul>	<ul style="list-style-type: none"> <li>• Not all data is available</li> <li>• Not all data available in real-time and not all is dynamic</li> <li>• Data for some reports is available with one-year latency</li> <li>• Some data is processed manually, some semi-manually</li> <li>• Some reports are done manually, some semi-manually</li> <li>• Internal practice of reporting is under the development</li> <li>• Calculation methodologies are under the development</li> <li>• Some KPIs might be missing metrics</li> <li>• Information flow between units and sub-units might be challenging at times</li> <li>• No standards for the reporting</li> </ul>
Opportunities	Threats

<ul style="list-style-type: none"> <li>• Static data can be switched to dynamic</li> <li>• Data can be available in real-time</li> <li>• Reporting process can be fully automated</li> <li>• BI dashboard with the year over year, month over month trend views; data might be also available retrospectively</li> <li>• Satisfy customers requests for transparency</li> <li>• Complete consulting service for customers on opportunities for energy consumption and emission reduction</li> <li>• Customers can make smarter decisions with the help of BI reporting and analytics tool, and consulting services</li> <li>• BI dashboards can give easy access to KPIs</li> <li>• Granularity of the reports</li> <li>• Reports drill down to product/service level</li> <li>• Data can be used for other purposes in the reporting, such as analytics and forecasting</li> </ul>	<ul style="list-style-type: none"> <li>• Accuracy of the reports</li> <li>• Time and resources of the employees</li> </ul>
---	--

As seen from Table 3, the results include 11 strengths, 10 weaknesses, 11 opportunities, and 2 threats.

The current sustainability reporting process demonstrates strong regulatory compliance and a proactive approach to new reporting requirements. This enhances the company's reputation and credibility. The process includes measurement and reporting of both GHG emissions and energy consumption, providing a comprehensive view of environmental impact. It helps identify energy-saving opportunities and optimize resource allocation, leading to cost savings and operational efficiency. This is illustrated in the following quote by the Respondent 4:

Hyperscalers, such as AWS, Microsoft, Google are reporting only greenhouse gas emissions. We try not only to do that but see the background – energy consumption. Since you can avoid greenhouse gas emissions choosing the right kind of electricity. We can move from one platform to another, from one cloud solution to another to make sure they get more efficient. (*Respondent 4*)

Additionally, during one of the interviews, dedicated to Current State Analysis, the interviewee mentioned that “Customers want to get consulted how they can improve their energy efficiency”. (*Respondent 3*)

The case company also tracks emissions of other suppliers for the customers, promoting collaboration throughout the value chain. The case company possesses expertise in sustainability, ensuring comprehensive reporting and consulting solutions. It understands data sources, ensuring reliable and accurate reporting. Decision-making is supported by product owners, integrating sustainability considerations. Customers' consumption is aligned with invoice volume, promoting transparency and empowering

sustainable choices. Overall, these strengths enable effective environmental impact management and stakeholder engagement.

Even though there was a number of weaknesses identified during the Current State Analysis, they are actively addressed by the case company. Efforts are underway to automate the reporting process and integrate data sources with a Business Intelligence (BI) reporting and analytical tool. The automation will improve the availability and real-time nature of data, eliminating manual and semi-manual processes. The ongoing development of internal practices and calculation methodologies will ensure greater accuracy and reliability in reporting. Additionally, steps are being taken to incorporate missing metrics into Key Performance Indicators (KPIs) and streamline the information flow between units and sub-units. The company recognizes the importance of standardized reporting formats and is working towards their implementation. These initiatives highlight the company's commitment to leveraging technology and streamlining processes, resulting in a more efficient and comprehensive sustainability reporting framework.

While the identified weaknesses in reporting process pose certain threats, it is important to note that these threats can be mitigated and eventually eliminated through the implementation of improvements. The accuracy of the reports can be enhanced by addressing data availability issues, transitioning to real-time data, and automating the reporting process. By integrating data sources with a Business Intelligence (BI) reporting and analytical tool, the company can ensure more accurate and reliable reporting. Similarly, the challenges related to time and resources can be overcome by streamlining the reporting process through automation and utilizing the capabilities of the BI tool. Once these weaknesses are effectively addressed and eliminated, the threats to the accuracy of reports and the strain on employees' time and resources will no longer exist, creating a more efficient and sustainable reporting framework.

The current sustainability reporting process presents opportunities for significant improvements that can transform weaknesses into strengths. By transitioning from static to dynamic data, the reporting process can become more responsive and provide real-time insights. The full automation of the reporting process will enhance efficiency and accuracy, allowing for timely and streamlined reporting. With the integration of a Business Intelligence (BI) dashboards, the company can offer year-over-year and month-over-month trend views, with potential to provide additionally the data

retrospectively to some extent. This level of transparency and comprehensive reporting will satisfy customers' requests for transparency and enable them to make informed decisions. Additionally, the company can leverage in the future the BI analytics tool to provide customers with a complete consulting service on opportunities for energy consumption and emission reduction. The BI dashboards will allow easy access to KPIs, empowering customers with the necessary information to make smarter decisions. Moreover, the granularity of the reports and the ability to drill down to the product or service level will provide deeper insights and support targeted sustainability initiatives.

### 3.3.2 Key Challenges and Business Impact

The key challenges that occur on different stages of sustainability reporting process and make an impact on business were discussed during the set of group interviews. Key challenges are presented in Table 4.

Table 4. Key Challenges of sustainability reporting process and Business Impact.

Stage	Challenge	Business Impact
<b>Data Evaluation</b>	Identifying metrics for Key Performance Indicators (KPIs) related to energy consumption of virtualized resources in data centers	It can influence the ability to effectively and accurately measure and optimize energy consumption
<b>Data Collection</b>	In some cases, data is collected manually	It can significantly impact businesses by consuming valuable time and resources, introducing human error, and hindering the ability to gather data in a timely and efficient manner for sustainability reporting and decision-making processes
<b>Data Collection</b>	In some cases, relying on retrospective data rather than real-time data and static figures instead of dynamic information	It can limit businesses' ability to gain up-to-date insights, make timely decisions, and adapt strategies effectively in response to changing sustainability trends and goals
<b>Data Processing</b>	Developing an accurate energy consumption assessment methodology for virtualized resources in data centers	It can hinder the ability to effectively measure, optimize, and communicate their energy efficiency efforts, potentially

		leading to higher operational costs and missed sustainability targets
<b>Customer Reporting</b>	Lack of standardized reporting formats	It can make it difficult to ensure consistency, comparability, and effective communication of information, potentially leading to confusion, inefficiency, and reduced stakeholder understanding and engagement

As it can be seen from Table 4, the Challenge column contains five distinct challenges, affecting various aspects of the reporting process and overall sustainability reporting performance. Stages with which challenges are associated, are marked in the Stage column. Each challenge has a direct impact on business. All the impacts are listed in the Business Impact column.

Within the Data Evaluation Stage, there is a challenge related to the identification of metrics for Key Performance Indicators (KPIs) that are specifically linked to the energy consumption of virtualized resources in data centers. This challenge holds substantial business implications as it directly impacts the organization's capacity to measure and optimize energy consumption effectively and accurately. In the absence of suitable metrics, businesses may encounter difficulties in evaluating their energy efficiency initiatives and making well-informed decisions aimed at reducing consumption and related expenses. To address this challenge, further in-depth theoretical studies and comprehensive research on existing best practices are suggested.

In the Data Collection Stage, a challenge arises from the manual collection of data in certain cases. This presents several difficulties and has notable business impacts. Manual data collection consumes valuable time and resources, increases the risk of human errors, and hampers the ability to gather data in a timely and efficient manner. These challenges can have adverse effects on sustainability reporting processes and hinder decision-making that relies on accurate and up-to-date information. To address this challenge, implementing automation solutions and technologies can greatly enhance the data collection process, streamlining efficiency, reducing errors, and enabling real-time data availability. By automating data collection, businesses can overcome the limitations associated with manual processes and improve the overall effectiveness of their sustainability reporting efforts.

Another challenge within the Data Collection Stage is the reliance on retrospective data instead of real-time data and the utilization of static figures instead of dynamic information. This challenge can limit businesses' ability to gain up-to-date insights, make timely decisions, and adapt strategies effectively in response to changing sustainability trends and goals. Real-time and dynamic data are essential for proactive sustainability management and maintaining a competitive edge in an evolving business landscape. By embracing technologies and methodologies that provide up-to-date and dynamic data, business can enhance the ability to respond swiftly and strategically to emerging sustainability challenges and opportunities.

The Data Processing Stage, the challenge is the development of an accurate energy consumption assessment methodology for virtualized resources in data centers. This challenge can hinder businesses' ability to effectively measure, optimize, and communicate their energy efficiency efforts. Without a robust methodology, organizations may face difficulties in identifying energy-saving opportunities, optimizing resource allocation, and achieving sustainability targets, potentially resulting in higher operational costs, and missed environmental objectives. To address this challenge, further research and the exploration of innovative technologies and techniques are essential. By developing a precise energy consumption assessment methodology, businesses can enhance their ability to track and manage energy usage, identify areas for improvement, and drive meaningful progress towards their sustainability goals.

The challenge, identified in the Customer Reporting Stage is the lack of standardized reporting formats. This challenge can make it difficult to ensure consistency, comparability, and effective communication of information across different sustainability reports. Without standardized formats, stakeholders may encounter confusion, encounter inefficiencies in data analysis, and experience reduced understanding and engagement with regards to sustainability performance. To address this challenge, it is necessary to implement standardized reporting formats that provide a clear structure and framework for presenting sustainability information. By establishing consistent reporting formats, businesses can enhance transparency, facilitate meaningful comparisons, and improve stakeholder comprehension of sustainability efforts. This can foster greater accountability, informed decision-making, and active participation in sustainability initiatives.

### 3.4 Summary of the Current State Analysis Results

The sustainability reporting process analysis revealed several strengths, weaknesses, opportunities, and threats. The strengths include regulatory compliance, comprehensive reporting of GHG emissions and energy consumption, understanding of customer needs, availability of resources, sustainability competencies, data source understanding, and decision-making support. However, certain weaknesses need attention, such as data availability, real-time accessibility, manual processes, and the absence of standardized reporting practices. Nevertheless, opportunities exist to enhance data dynamics, automate reporting, provide consulting services, and improve access to KPIs through BI dashboards. It is important to address challenges related to accuracy, employee time, and resource management, including metrics, data collection, energy assessment, and reporting formats. These challenges are actively being addressed to ensure a robust and efficient reporting framework for corporate sustainability.

#### 3.4.1 Selected Focus Areas

When choosing the focus areas for this study, all the weaknesses and threats presented in Figure 6, as well as challenges listed in Table 3 were evaluated and prioritised together with stakeholders. The case company is aware of all existing weaknesses and recognizes them as opportunities for improvement. The stakeholders are actively engaged in establishing new practices to enhance the quality of the reports and to build more efficient and integrated reporting process. The case company is already proactively addressing the threat related to *Time and Resources* by undertaking the automation of its reporting processes. However, the stakeholders are interested to cooperate in tackling weaknesses associated with the *Accuracy of the Reports*.

The hypothesis suggests that the existing energy consumption assessment methodology for virtualized resources in data centers could be advanced. This hypothesis was created due the novel nature of the existing methodology. Hence, the focus areas of this study will be made on two stages of sustainability reporting process: *Data Evaluation* and *Data Processing*, on which the metrics for the KPIs are selected and the calculation methodology is developed. The emphasis will be done on refining the metrics and assessment methodology for potentially more accurate calculation of energy consumption by virtualized resources in data centers and ensuring that the calculations align with industry standards and best practices. By concentrating efforts on the chosen

focus areas, the accuracy and reliability of sustainability reports can be improved, which justifies the choice of the thesis scope.

These selected focus areas will be explored in-depth during the study, examining existing literature and best practices in Section 4.

## 4 Existing Knowledge and Best Practice on Energy Consumption Assessment for Virtualized Resources in Data Centers

This section discusses environmental sustainability frameworks and initiatives, as well as topics related to energy consumption in data centers, covering metrics for measuring data center energy efficiency, virtualization-specific metrics, and the industry best practice for calculating energy consumption of virtualized resources. Existing knowledge and best practice in these domains are the foundation for developing *Conceptual Framework* that provides insights for advancing energy consumption assessment methodology for virtualized resources in data centers.

### 4.1 Environmental Sustainability Frameworks, Initiatives and Goals

In recent years, the growing recognition of climate change and its potential consequences, as well as global imperative to address environmental sustainability has become increasingly evident. Organizations across various sectors are embracing their responsibility to mitigate greenhouse gas emissions and work towards achieving net zero targets. This section discusses the four pillars of the environmental sustainability frameworks, initiatives, and goals: Greenhouse Gas (GHG) Protocol, Science Based Targets (SBTs), ISO 14064, and Net Zero.

Environmental sustainability, as a global concern, created the need in the development of standardized frameworks to measure and manage greenhouse gas (GHG) emissions. The GHG Protocol was established through a collaborative 20-year partnership between the World Resources Institute (WRI) and the World Business Council for Sustainable Development (WBCSD) and provides comprehensive and globally recognized standardized frameworks for measuring and managing GHG emissions across various sectors, both in public and private sector operations, value chains and mitigation actions. The GHG Protocol collaborates with a range of stakeholders, including governmental bodies, industry associations, non-governmental organizations (NGOs), businesses and other entities. The GHG Protocol's guidelines allow the organizations to accurately quantify and report the emissions, make informed decisions and target to reduction strategies. The GHG Protocol has gained a worldwide adoption among companies and organizations. Its Corporate Accounting and Reporting Standard serves as the foundation for numerous corporate GHG reporting programs globally, with a remarkable 92% of Fortune 500 companies, responding to the CDP in 2016, utilizing GHG Protocol

directly or indirectly. Over the years, the GHG Protocol has undergone updates and enhancements, refined its guidance and introduced calculation tools to facilitate emissions quantification and the measurement of climate change mitigation benefits. (Greenhouse Gas Protocol, 2023)

The Greenhouse Gas (GHG) Protocol Initiative categorises business GHGs as Scope 1, Scope 2, and Scope 3 emissions. (Greenhouse Gas Protocol, 2023) Reporting Scope 1 and Scope 2 emissions are already mandatory for many organizations not only in EU, but globally. (Green Business Bureau, 2022) The Scope 1 stands for direct GHGs emissions that are released from the sources owned or controlled by business. The Scope 2 emissions are indirect GHGs and account for the electricity purchased and consumed by the company. The Scope 3, for today, is still an optional reporting category, which emissions are the consequences of the company's activities, originated from the sources which are not under the ownership or the control of the organization. (Greenhouse Gas Protocol, 2023)

Greenhouse Gas (GHG) protocol offers “calculation tools and guidelines to quantify and manage emissions in commercial data centers and cloud platforms”. These efforts aim to clarify the assessment procedures for energy consumption and facilitate a more holistic understanding of energy usage. (Cappiello et al., 2013)

Figure 4 below presents the diagram of scopes and emissions across the value chain.

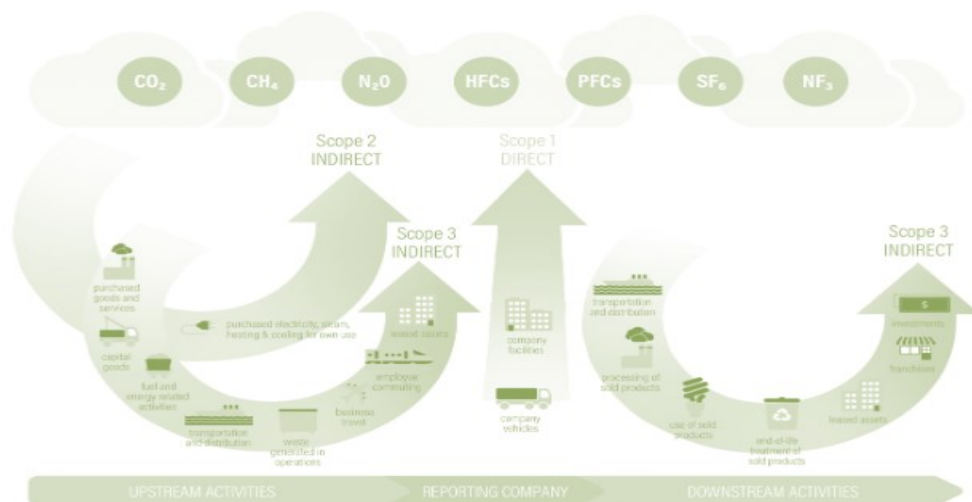


Figure 4. Overview of GHG Protocol Scopes and Emissions Across the Value Chain. (World Resources Institute, 2023)

Figure 4 is a visual representation of an organization's Greenhouse Gas emissions, which are divided into Scope 1, 2 and 3. Even though there are no legal requirements to report on Scope 3 emissions yet, most total corporate emissions come from the Scope 3 sources. (Microsoft, 2023)

According to Carbon Disclosure Project (CDP), Scope 3 emissions account for 75% of company's emissions on average. (World Resources Institute, 2022) It means that many companies are missing out on significant opportunities for improvement. There are many direct and hidden factors that contribute on the emissions, caused by cloud services, and they fall into the Scope 3. Businesses that have complex supply lines and global distribution networks face difficulties in identifying and mitigating emissions throughout their supply and value chains. (Drax, 2022) The case company is assisting customers via consulting and analytical services in identifying the emissions generated both under the Scope 2 and Scope 3.

According to McKinsey (2023), end-user devices like laptops, tablets, smartphones, and printers contribute 1.5 to 2.0 times more carbon emissions globally compared to data centers. This is primarily because companies possess a larger number of end-user devices than servers in their on-premises data centers. Moreover, end-user devices are usually replaced more frequently, with smartphones having an average refresh cycle of two years, laptops four years, and printers five years. On the other hand, servers are typically replaced every five years, although 19 percent of organizations extend this replacement cycle. (McKinsey, 2023)

Figure 5 below illustrates that the biggest carbon culprit is end-user devices.

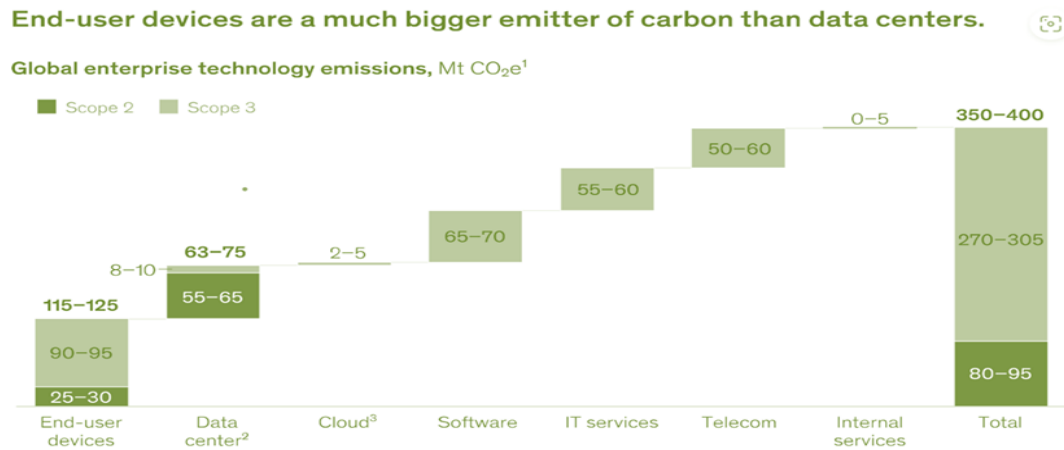


Figure 5. Global enterprise technology emissions. (McKinsey Digital, 2023)

It can be seen from Figure 5, that end-user devices generate significantly more carbon emissions compared to data centers.

The objective of the Science Based Targets (SBTs) Network is to drive systemic changes in economic systems, while safeguarding the Earth's shared resources: the air, water, land, biodiversity, and oceans and help organizations to set ambitions and science-based goals for reducing greenhouse gas emissions. It brings together a diverse group of experts from over 60 non-governmental organizations (NGOs), business associations and consultancies, whose collaborative efforts aim to collectively determine the necessary actions required to operate within the Earth's ecological limits and effectively fulfil the needs of the society. The SBTs initiative has established itself as a critical platform for setting science-based targets across various domains. The vision of SBTs Network is centred around creating a global economy where companies and cities operate within environmentally sustainable boundaries, while ensuring social equality. This is achieved through the setting of science-based targets, which serve as transformative benchmarks to mitigate environmental impacts. The SBTs Network focuses on developing methodologies that enable cities and companies to set integrated targets that encompass all Earth systems. While significant progress has already been made in establishing science-based targets for climate change, the SBTs aims to extend this approach to cover targets related to water, land, biodiversity, and the oceans. The other objective of SBTs Network is to demonstrate substantial progress in alignment with key global policy milestones such as the Sustainable Development Goals (SDGs) and the targets set under international conventions like the United Nations Framework Convention on Climate Change (UNFCCC), United Nations Convention to Combat

Desertification (UNCCD), and Convention of Biological Diversity (CBD). This progress is tracked through the development of an indicator framework that monitors the advancements made by SBTs setters. The significance of integrating the adoption of science-based targets within capital markets is also recognized. To achieve this, the collaborates with policymakers, financial institutions, and service providers such as benchmarkers and credit agencies is made. (Science Based Targets Network, 2023)

ISO 14064 is an internationally recognized standard for greenhouse gas (GHG) emissions inventories and verification. It was developed by the International Organization for Standardization (ISO) over a four-year period and involved the participation of over 175 experts from 45 countries. The standard consists of three parts, with Part 1 focusing on the quantification and reporting of GHG emissions at the organizational level. It provides minimum requirements for GHG inventories, ensuring credibility and consistency in independent auditing. ISO 14064 offers organizations improved consistency, flexibility, and reduced effort in conducting voluntary GHG inventories. Part 3 of the standard introduces a standardized process for verifying GHG inventories and reduction processes. The standard serves as a valuable reference for policymakers, organizations, and program developers in addressing the global challenge of climate change. ISO 14064 provides a technical specification and is climate policy neutral, serving as a guide for the private and public sectors in developing GHG inventories and forming the basis for climate change initiatives. By following ISO 14064, organizations can enhance the credibility of their GHG reduction efforts, facilitate international cooperation, and promote consistency in products and services across national boundaries. (EPA, 2023)

According to the international scientific community, to mitigate severe climate impacts, it is crucial to reduce global carbon dioxide (CO<sub>2</sub>) emissions caused by human activities by approximately 45 percent from 2010 levels by 2030, ultimately reaching net zero emissions by around 2050. The degree of global warming is directly linked to the cumulative amount of CO<sub>2</sub> emissions, which means that if global emissions exceed zero, the planet will continue to experience increased warming. Consequently, the damages resulting from climate change will continue to worsen if emissions persist. The Net Zero concept refers to achieving a state where the amount of greenhouse gases released into the atmosphere is balanced by the amount removed from the atmosphere. This equilibrium is significant because it signifies a halt to global warming, particularly for carbon dioxide (CO<sub>2</sub>). The importance of net zero is emphasized in the Paris Agreement,

which calls for countries to achieve a balance between human-caused emissions and greenhouse gas removals in the latter half of this century. The concept of "going net zero" involves reducing greenhouse gas emissions and ensuring that any remaining emissions are offset by removal efforts. The term "net" is crucial because it acknowledges the difficulty of eliminating all emissions completely within the required timeframe. In addition to substantial emission reductions, it is necessary to scale up removal efforts. The removal of greenhouse gases must be permanent, so that Net Zero is effective, meaning that emissions do not enter the atmosphere again over time due to factors such as deforestation or inadequate carbon storage. Net zero serves as the overarching goal for mitigating global warming in the latter half of the century and is internationally recognized as such. Various terms like "Carbon Neutral," "Net Zero," and "Climate Neutral" are used to account for emissions sources and sinks in different contexts, providing clarity on what is included or excluded in calculations or targets. The Intergovernmental Panel on Climate Change (IPCC) has concluded that achieving net zero CO<sub>2</sub> emissions by 2050 is necessary to remain consistent with limiting global warming to 1.5 degrees Celsius. (Net Zero Climate, 2023)

#### 4.2 Corporate Sustainability Reporting Directive

The European Commission has introduced the Corporate Sustainability Reporting Directive (CSRD) as a regulatory framework to improve the transparency and disclosure of sustainability information by companies in the European Union (EU). Its objectives include standardizing sustainability reporting requirements and increasing the reliability and comparability of sustainability data provided by these companies. The directive entered into force on 5 January 2023. (European Commission, 2023)

The CSRD expands reporting obligations beyond the NFRD, encompassing environmental protection, board diversity, social responsibility, employee relations, anti-corruption, and more. It introduces the concept of double materiality, considering sustainability aspects from both outside-in and inside-out perspectives. The CSRD also includes reporting on intangible assets, aligning with regulations such as the Sustainable Finance Disclosure Regulation (SFDR) and the EU Taxonomy Regulation. The CSRD aims to standardize ESG reporting, integrate it with financial reporting, and drive strategic corporate management across all departments. Establishing a technological foundation for managing finance and non-finance data volumes, monitoring KPIs, and steering

strategic initiatives becomes crucial for effective ESG reporting and performance management. (Walters Kluwer, 2023)

The CSRD is a crucial part of the EU's Green Deal, which aims to transition the economy to a sustainable, responsible, and digital model. It supports the Action Plan for Financing Sustainable Growth, which focuses on redirecting capital flows to sustainable investments, integrating sustainability into risk management, and promoting transparency and long-termism in economic activities. The CSRD enhances transparency by requiring companies to disclose finance-grade ESG information in their annual reports, making the data more credible and comparable. This enables stakeholders to make informed decisions and directs more capital towards sustainable businesses. Additionally, it promotes corporate accountability by encouraging companies to integrate ESG considerations into their practices, making data on their environmental and societal impact publicly available. (Onetrust, 2023)

Figure 6 below shows how the Corporate Sustainability Reporting Directive (CSRD) relates to the European Green Deal, EU Taxonomy, and Sustainable Finance Disclosure Regulation (SFDR).

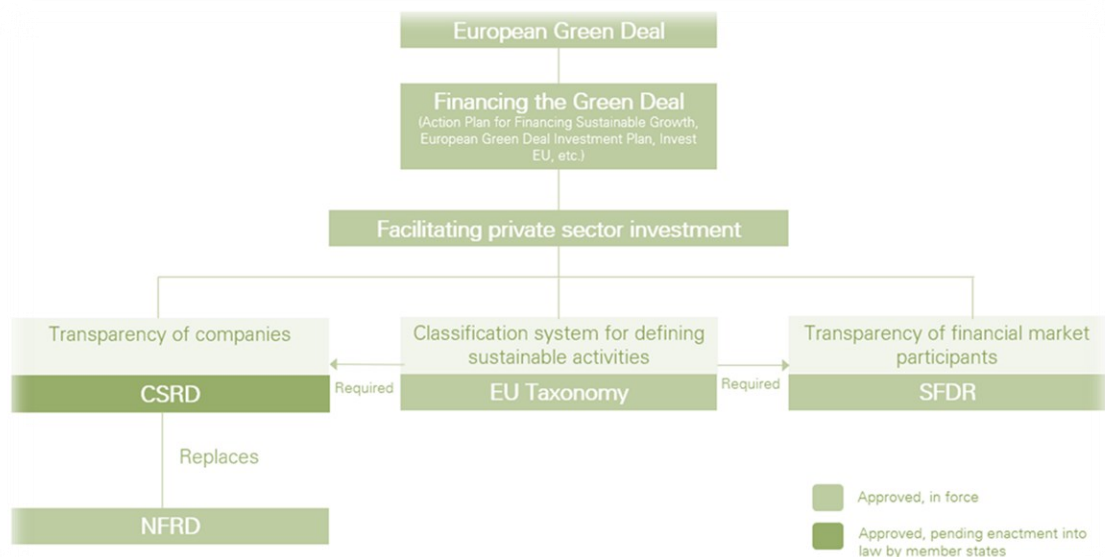


Figure 6. The relation of CSRD to the European Green Deal, EU taxonomy, and SFDR. (Onetrust, 2023)

It is also seen from Figure 6, that CSRD will replace Non-Financial Reporting Directive (NFRD), in compliance with which the case company had to previously include non-financial statements as an integral part of its annual public reporting obligations.

The aim of CSRD is to establish more comprehensive and standardized sustainability reporting requirements for companies operating in the EU, non-EU companies meeting certain thresholds for net turnover in the EU and companies with securities listed on a regulated EU market. Companies falling within the scope will be obligated to disclose information regarding the impact of sustainability-related factors, such as climate change, on their operations, as well as information on how their business model affects sustainability factors. The reporting requirements encompass environmental, social, human rights, and governance factors. Environmental aspects encompass climate-related emissions (including Scopes 1, 2, and 3 greenhouse gases), water and marine resources, circular economy, pollution, and biodiversity. (Harvard Law School Forum on Corporate Governance, 2023)

Data centers that haven't established clear and meaningful sustainability metrics and key performance indicators (KPIs) should take immediate action. Compliance with CSRD requires limited assurance (expected to transition to reasonable assurance), aiming to enhance the credibility of reported sustainability data and provide much-needed transparency in the data center industry. It is crucial for the data center community to move beyond setting goals and take swift action to implement policies and detailed plans in preparation for CSRD. Both the industry as a whole and individual companies must define sustainability strategies and establish processes and controls to generate the necessary qualitative and quantitative data for producing meaningful sustainability reports. (Shorten, 2022)

#### 4.3 Data Center Energy Consumption

It is important to start with a high-level understanding what is a data center, how it is designed, what kind of services it provides and how it is powered.

Lääkkölä (2015) describes a data center as “a dedicated facility where computing systems and networking, data storage and other associated equipment are physically located, maintained and operated.”

According to IBM (2023), a data center is “a physical room, building or facility that houses IT infrastructure for building, running, and delivering applications and services, and for storing and managing the data associated with those applications and services.” (IBM, 2023)

Kaivosoja (2016) defines fundamental components, that all data centers include, regardless of their size:

1. The Building: Building Security measures.
2. ICT Equipment: Server Computers that handle data processing, Data Storage Equipment for storing and retrieving data and networking equipment to facilitate communication between devices.
3. Electrical Infrastructure: Power Supply Equipment responsible for providing electrical power, Uninterruptible Power Supply (UPS) to ensure continuous power in case of outages and Backup Power sources for emergency situations.
4. Mechanical Infrastructure: Air Conditioning Equipment to regulate temperature, Cooling Equipment for maintaining optimum operating conditions, Humidity Control Equipment to manage moisture levels. (Kaivosoja 2016)

According to Lääkkölä (2015), data centers offer services to either the entity operating the data center or external clients. The services provided by data centers can be broadly categorized into the following areas:

1. Data storage
2. Data processing
3. Data transmission. (Lääkkölä 2015)

These services are commonly referred to as "Something-as-a-Service", where data center provide the underlying infrastructure and resources for the specific service. Infrastructure-as-a-Service (IaaS) allows customers to access virtual server instances, storage volumes, and internet connections. The hardware and physical system

architecture are virtually abstracted, and the customers can control and manage these resources through a web interface. Companies can purchase data center virtualized resources with either constant capacity or elastic capacity that can be scaled based on demand. This flexibility caters to various needs, such as setting up private cloud environments, offering Software-as-a-Service (SaaS) or Platform-as-a-Service (PaaS) solutions to their own customers. To deliver these services, data centers require computing servers, storage devices, and networking equipment. There is a wide range of hardware options available in each category, and it is the responsibility of the data center maintainer to select the most suitable options and configure them appropriately based on the specific service requirements. (Lääkkölä, 2015)

Portnoy (2016:15) states that cloud computing creates the concept of a virtual data center, which contains all the elements of the physical data center due to the ability to abstract the physical layer from the hardware. Resources in a virtual data center are offered on as-needed basis, which allows companies to speed up the deployments, not sacrificing scalability, resiliency, or availability.

According to IBM (2023), there are three most popular types of cloud services: IaaS, PaaS, and SaaS. “IaaS, or infrastructure as a service, is on-demand access to cloud-hosted physical and virtual servers, storage, and networking - the backend IT infrastructure for running applications and workloads in the cloud. PaaS, or platform as a service, is on-demand access to a complete, ready-to-use, cloud-hosted platform for developing, running, maintaining, and managing applications. SaaS, or software as a service, is on-demand access to ready-to-use, cloud-hosted application software.” (IBM, 2023)

Figure 7 below demonstrates the distinction among IaaS, PaaS, and SaaS.

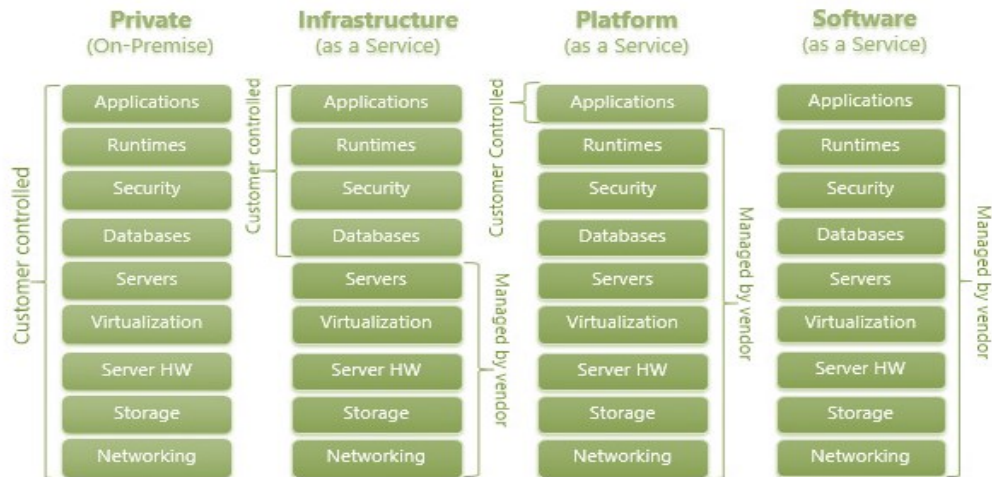


Figure 7. Difference between IaaS, PaaS, and SaaS. (Tech Trainees, 2018)

As it is seen from Figure 7, the control and management of resources are divided between the customer and the vendor (Tech Trainees, 2018). The cloud computing model allows businesses to host all their applications on the servers of cloud service providers (CSPs), instead of purchasing, installing, and maintaining their own physical hardware. It helps organizations to easily scale new digital solutions by enhancing their computing power quickly and cost-effectively. (McKinsey & Company, 2022)

The European cloud computing market is rapidly growing and by 2030, it is projected to reach 560 billion euros, which is illustrated in Figure 8 below.

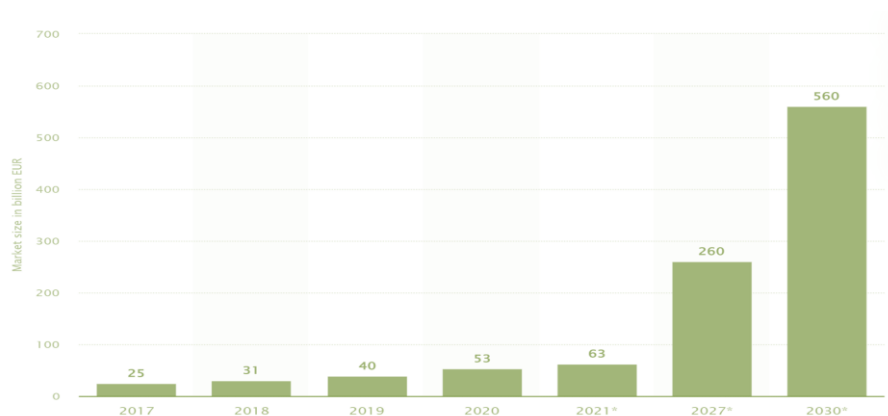


Figure 8. Cloud computing market size in Europe from 2017 to 2030 (in billion EUR). (Statista, 2023)

In comparison with the old data centre approach, the deployment in the cloud is significantly more environmentally friendly, since the technology sector is the largest purchaser of renewable energy sources. (Nature Climate Change, 2020)

Modern data centers, which are typically powered by electricity, are adopting alternative energy sources such as geothermal, wind and solar power to reduce carbon emissions and adhere to sustainable operational practices. Electric power is supplied from external power grids and is distributed within the data center to support infrastructure facilities, Information Technology (IT) equipment and other systems. To ensure uninterrupted power supply even in the events of power failures, data centers employ Uninterrupted Power Supplies (UPS). (Makris, 2017)

Figure 9 below illustrates data center infrastructure energy use, divided into infrastructure (of non-IT power) and the IT equipment power consumption.

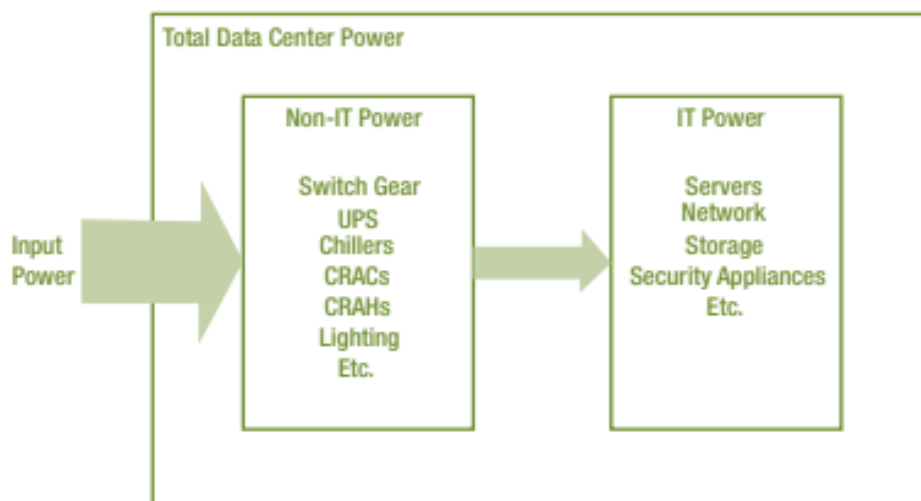


Figure 9. The power consumption of a data center. (Gough et al., 2015:3)

Non-IT power is seen as a factor, which contributes to inefficiency of the data center, as it does not directly aid in processing information. The purpose of infrastructure is to support IT equipment by maintaining appropriate environmental conditions and delivering uninterrupted power. Therefore, it is an overhead power usage. The IT equipment, on the other hand, directly contributes to information processing, making it closely tied to the data center's efficiency. (Gough et al., 2015: 3)

Makris (2017) suggests that the complex architecture of the data center consists of hardware and software elements. The hardware layer within a data center comprises various elements, including cooling systems, power distribution units, lighting equipment, servers, and networking equipment. The servers account for most of the energy consumption in the racks, with the processor being the primary component that consumes the highest amount of energy. Adoption of virtualization is one of the solutions for achieving energy efficiency in data centers.

Makris (2017) categorizes the software layer into Operating System/Virtualization layer and the Application layer. He states that “The first mainly refer to the host OS that is installed in the servers and the cloud deployment running on top of it. The second refers to the different type of applications running in the servers which vary depending on the industry and business cases”

This study is particularly focused on virtualization layer and specifically in energy consumption metrics related to this layer. These domains will be discussed in more detail in the following section.

#### 4.4 Energy Consumption Metrics for Virtualization Layer

Virtualization is the key underlying technology in cloud infrastructures. (Keller et al., 2010) “Virtualization is a disruptive technology, shattering the status quo of how physical computers are handled, services are delivered, and budgets are allocated.” (Portnoy, 2016: 1)

“A virtual machine (VM) is a virtual environment that functions as a virtual computer system with its own CPU, memory, network interface, and storage, created on a physical hardware system (located off- or on-premises).” (RedHat, 2022)

“A virtual machine (VM) can virtualize all of the hardware resources, including processors, memory, storage, and network connectivity. A virtual machine monitor (VMM), which today is commonly called a *hypervisor*, is the software that provides the environment in which the VMs operate.” (Portnoy, 2016:2)

According to Lee et al. (2018), “virtualization refers to the creation of a virtual version of a device or resource, such as server, storage resource, network or even an operating system wherever the framework divides the resource into one or more execution environments. In other words, virtualization is a framework or methodology of dividing the resources of a computer into multiple execution environments, by applying one or additional concepts or technologies such as hardware and software partitioning, time-sharing, partial or complete machine simulation, quality of service, emulation, and many others. (Lee et al., 2018: 39)

Figure 10 below illustrates a virtualization system.

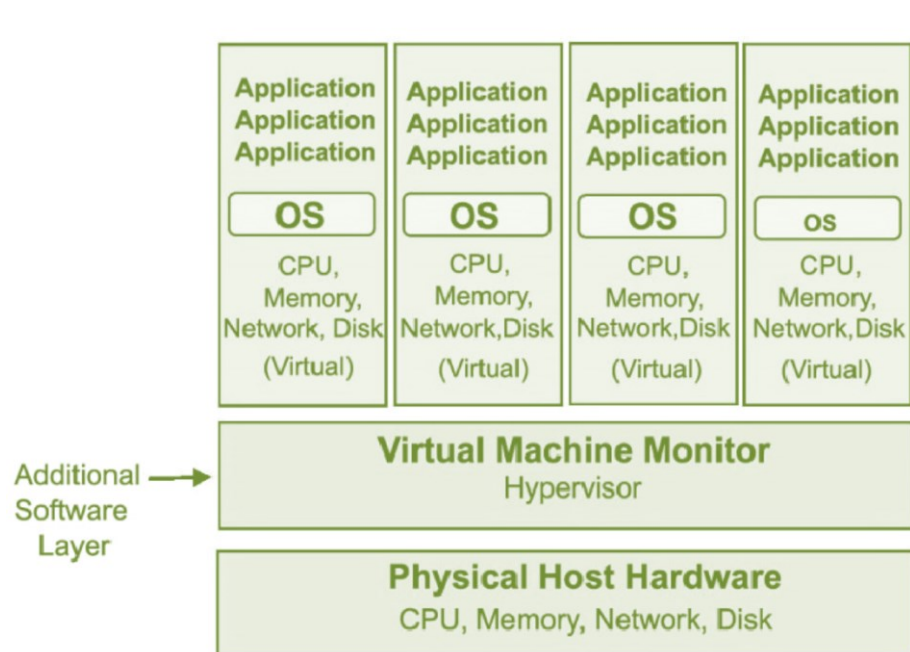


Figure 10. Simple representation of a virtualized system. (Lee et al., 2018: 39)

As it is seen from Figure 10, the hypervisor acts as an intermediary between the physical layer and virtualized servers or virtual machines, providing an abstraction of the underlying hardware resources.

Hypervisors, as foundation of virtual environments, not only provide a platform for running virtual machines, but also enable improved availability features and create additional options for provisioning and management. On the other hand, virtual machines are the engines, powering the applications. (Portnoy, 2016: 16)

The hypervisor manages various tasks such as storage Input/Output (I/O) requests, network I/O, memory processing, and central processing unit (CPU) workload for all the guests hosted on the physical server where the hypervisor is installed. (Portnoy, 2016: 28) The hypervisor considers compute resources such as CPU, memory, and storage as a shared pool that can be easily moved between current virtual machines or allocated to new ones. (RedHat, 2022)

By virtualizing servers, multiple virtual servers can be hosted on each physical server, improving hardware utilization. This eliminates the need to purchase additional physical resources and reduces power, space, and cooling requirements in data centers. (RedHat, 2022)

According to IBM (2023), in cloud computing, virtual machines are commonly available in two forms: single-tenant and multi-tenant configurations. “Public, or multi-tenant, virtual machines are virtual machines in which multiple users are sharing common physical infrastructure.” (IBM, 2023)

In a multi-tenant scenario, virtual machines (VMs) can be assigned to different customers or user accounts. Determining the energy consumption of individual customers in a scenario where several customers use one virtual machine can be challenging. Energy consumption in a data center is typically measured at the infrastructure level, such as the power consumed by the physical server hosting the virtual machine. It is difficult to accurately allocate the energy usage to specific virtual machines or customers within that shared environment.

According to Haghshenas et al. (2022), power consumption of a host in a virtualized environment is influenced by various factors, including CPU, memory, hard disk, I/O, and network utilization.

Central processing units are responsible for executing instructions and performing calculations in a computer. The energy consumed by a CPU depends on its workload and the operations it performs. It collaborates with other hardware elements and depends on software instructions to execute a variety of tasks and procedures. (Computerblog, 2023)

Similarly, RAM, also consumes energy. Random Access Memory (RAM), also referred to as random-access memory or device memory, functions as the temporary storage for all the data required by the processor for quick access during various tasks. It serves as the intermediary between the device's CPU (processor) and the data stored on its hard drive. When data, such as running a program or accessing a file, is retrieved, RAM facilitates the delivery of that data to the CPU. Additionally, RAM plays a fundamental role in servers. (Techradar, 2021)

The amount of energy consumed by RAM depends on its capacity, speed, and usage patterns. By measuring the power consumption of RAM modules, it is possible to track the energy consumed by the memory subsystem of an IT system.

Each server in a cloud environment has its own specifications, including processing speed, memory capacity, storage capacity, and energy consumption. The processing speed of a server is typically measured in Million Instructions Per Second (MIPS). Additionally, servers have Random Access Memory (RAM) and bandwidth for data transfer. The servers also have local disks to host Operating Systems (OS), while Network Attached Storage (NAS) is used for storing virtual machines (VMs) and facilitating live VM migration. In a cloud environment, numerous users can request a certain number of VMs, and each VM has its own CPU load, memory utilization, and network transfer rate. (Gul et al., 2020)

CPU, memory, disk, and network are individual metrics or dimensions that contribute to assessing the utilization and load of servers, they can refer as metrics or resource types used in Virtual Machine (VM) consolidation an allocation decision in cloud data centers. These metrics are used to determine the utilization and workload of physical servers hosting virtual machines (VMs) and make decisions regarding VM migrations and resource allocation. (Hieu et al., 2017)

The study, conducted by Hieu et al. (2017) highlights the limitations of approaches that consider only CPU utilization and proposes a more comprehensive approach that considers multiple resource dimensions, such as CPU, memory, disk, and network bandwidth. By considering multiple resource types and their utilization patterns, the proposed approach aims to achieve more accurate energy consumption evaluation and optimization.

In the research paper by Gul et al. (2020), two energy-aware techniques for VM consolidation consider energy consumption by both CPU and RAM. This indicates that the proposed techniques consider both CPU and RAM as metrics for energy consumption.

Cloud computing utilizes virtualized hardware to enable a physical machine (PM) to run multiple virtual machines (VMs) with diverse types and distributions of resources. Within a cloud, various applications are hosted on these VMs, and each VM on a PM handles different workloads that may change over time. (Gul, et al., 2020)

Therefore, due to the diverse types of resources, varying workloads, and dynamic nature of cloud computing, accurately measuring, and attributing energy consumption to individual VMs or virtualized resources is a complex task.

Povoa et al. (2013) states that “Power management is an expensive and important issue for large computational infrastructures such as datacentres. However, measuring energy consumption of scalable systems may be impractical due to both cost and complexity for deploying power metering devices on a large number of machines.”

Povoa et al. (2013) recommends utilizing resource utilization information, such as the usage of the processor, memory, disk operations, and network traffic, as indicators for power consumption estimation.

Virtualization-aware energy monitoring tools can track energy consumption at different levels, including at the host server level, individual VM level, or even at the granularity of specific hardware components assigned to each VM. They consider factors such as CPU utilization, memory usage, disk I/O, and network traffic, while considering the resource sharing and consolidation aspects of virtualization.

Measuring energy consumption in the virtualized layer presents numerous obstacles and constraints. These include challenges related to resource sharing among multiple virtual machines (VMs), as energy usage gets distributed across shared physical resources. Precisely attributing energy consumption to individual VMs demands advanced techniques for accurate allocation. The introduction of virtualization overhead, such as the hypervisor layer, further complicates energy measurement, as differentiating its consumption from the VMs' energy usage becomes problematic. Dynamic resource

allocation techniques employed in virtualized environments add another layer of complexity, as VMs continuously migrate or resize, resulting in varying energy consumption patterns that are difficult to track accurately. Additionally, high virtual machine density on a physical server can significantly impact energy consumption due to resource contention and higher utilization, posing challenges in accurately measuring energy usage in densely populated virtualized environments. Moreover, the absence of standardized approaches and tools for measuring energy consumption in virtualized environments makes it difficult to compare and benchmark it across different virtualization platforms.

The next chapter discusses the factors beyond virtualization and explores comprehensive metrics to evaluate the overall energy efficiency of data centers.

#### 4.5 Data Center Energy Efficiency Metrics

Data centers are complex systems, where various factors impact their overall efficiency. The first step for making Data Centers more environmentally friendly is to determine the energy efficiency metrics, which should be monitored and analyzed. There is a variety of energy efficiency metrics, each possessing its own set of advantages, drawbacks, and constraints. The right choice of metrics helps to better track the performance, make informed decisions to enhance operational efficiency, reduce energy consumption and carbon footprint and ultimately to cut costs.

Data centers are highly dynamic environments, and their initial design becomes outdated soon after they start operating. Energy consumption estimations are based on static designs and initial assumptions, failing to account for the dynamic live environment. As data center facilities gradually fill up and infrastructure and capacity expand over time, load changes and the efficiency of each subcomponent must be considered during assessment. (Dumitrescu et al., 2018)

According to Makris (2017), the metrics which allow to measure infrastructure energy efficiency in data centers include Power Usage Effectiveness (PUE), Power to Performance Effectiveness (PPE), Data Center Productivity (DCeP) and Data Center Infrastructure Efficiency (DCIE).

Pitalier (2018) states that PUE metric is widely used and publicly disclosed by many data centers, it is one of the most important indicators of Data Centers. While DCIE is another metric, which is used to express the same concept of Data Center energy efficiency.

Rădulescu et al. (2017) calls PUE “the classical metric for measuring energy efficiency in DCs.” Jaureguiualzo (2011) claims that PUE is “widely accepted for measuring energy efficiency in Data Centers” (DC). According to Van de Voort et al. (2017), “the most widely used performance metric is PUE”. Similarly, Gough et al. (2015), defines PUE-as “the accepted metric for infrastructure efficiency.”

Thus, this section analysis Power Usage Effectiveness value, as the key metric, explores its benefits, limitations, and restrictions, as well as other complimentary metrics to PUE.

#### 4.5.1 Defining Power Usage Effectiveness (PUE)

Power usage effectiveness (PUE) metric was first proposed by Christian Belady and Chris Malone in 2006. (Rădulescu et al., 2017) Since then it has become a globally recognized metric for reporting data center energy use efficiency.

Dumitrescu et al. (2018) suggests that the evaluation of PUE value becomes crucial before making decisions regarding operational, IT and facility changes to enhance data center efficiency. Any modification within the facility, such as commissioning a new generation of servers or adapting advanced UPS technology, can affect PUE, which serves as a monitoring parameter, enabling the identification of areas for improvement.

PUE calculates the energy efficiency of a Data Center as Figure 11 bellow illustrates.

$$PUE = \frac{\textit{Total Facility Energy}}{\textit{IT Equipment Energy}} \textit{ (kW)}$$

Figure 11. The formula for calculating Power Usage Effectiveness metric. (Pitalier, 2018)

PUE is defined as the ratio of *Total Facility Energy* to *IT Equipment Energy*, where *Total Facility Energy* is the total power used in Data Center and *IT Equipment Energy* is the power used for IT equipment. (Rădulescu et al., 2017)

Total facility energy refers to the energy that is specifically allocated to the data center. This includes the energy measured at the utility meter of a dedicated data center facility or the meter for a data center or data room within a facility that serves multiple purposes. On the other hand, IT equipment energy refers to the energy consumed by the equipment responsible for managing, processing, storing, or directing data within the computing area. (Avelar et al., 2012)

Belady and Seander (2018:16) state that “the energy efficiency of an IT data center can be measured as a ratio of power being used for HVAC (heating, ventilating, and air conditioning) or cooling and power for IT equipment in dependent of actual work being done or data being stored.”

Another definition of PUE is illustrated in Figure 12 below.

$$PUE_{f,p} = \frac{\text{Total\_facility\_energy}}{\text{IT\_equipment\_energy}}$$

Figure 12. PUE value calculation formula. (Rădulescu et al., 2017)

In the formula, presented in Figure 12,  $f$  stands for Data Point Collection Frequency: year, month or week and  $p$  indicated the Metering position: Basic (UPS output), Intermediate and Advanced (at IT equipment). (Rădulescu et al., 2017)

A metric for energy use within specified limits which partitions the data center into zones is called Partial PUE ( $pPUE$ ) and the formula for its computation is shown in Figure 13.

$$pPUE = \frac{\text{Total\_power\_within\_a\_zone}}{\text{IT\_equipment\_power}}$$

Figure 13. pPUE computation formula. (Rădulescu et al., 2017)

The pPUE metric takes into consideration all the infrastructure components and can be only calculated for zones with IT. (Rădulescu et al., 2017)

The computation formula for energy within the zone is illustrated below in Figure 14.

$$pPUE(\text{zone}) = \frac{(\text{Non\_IT\_power}(\text{zone}) + \text{IT\_equipment\_power}(\text{zone}))}{\text{IT\_equipment\_power}(\text{zone})}$$

Figure 14. pPUE (zone) computation formula. (Rădulescu et al., 2017)

The scheme of PUE computation is illustrated below in Figure 15.

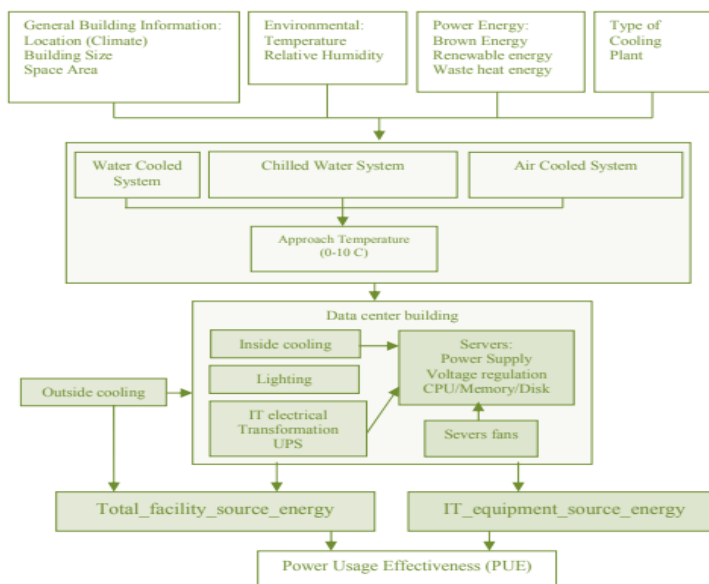


Figure 15. A general scheme for the PUE computation. (Rădulescu et al., 2017)

Power Usage Effectiveness value can range from 1(100% efficiency) to infinity. (Belady & Seader, 2018: 116) A lower PUE value (close to 1) indicates higher energy efficiency and reduced environmental impact, when performing Data Center using few energy on servers cooling. (Pitalier, 2018)

Figure 16 below shows the high-level visualization PUE subcomponents.

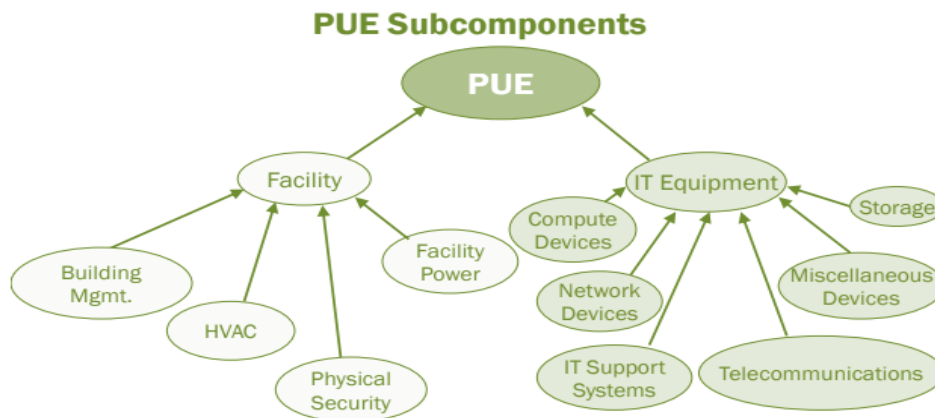


Figure 16. Overview of the subcomponents within a typical data center's facility and IT equipment. (Avelar et al., 2012: 18)

After defining the PUE value, its components and different computation approaches, it is important to discuss different levels of its measurement in detail, which is done further in this chapter.

#### 4.5.2 Power Usage Effectiveness (PUE) measurement levels

According to Avelar et al. (2012: 14) The Green Grid suggests a three-level approach for Power Usage Effectiveness assessment.

Table 5 illustrates the three-level approach to measure PUE value.

Table 5. High-level breakdown of The Green Grid's three-level approach to PUE measurement. (Avelar et al., 2012: 14)

	Level 1 (L1)	Level 2 (L2)	Level 3 (L3)
	Basic	Intermediate	Advanced
<b>IT Equipment Energy</b>	UPS Outputs	PDU Outputs	IT Equipment Input
<b>Total Facility Energy</b>	Utility Inputs	Utility Inputs	Utility Inputs
<b>Measurement Interval</b>	Monthly/Weekly	Daily/Hourly	Continuous (15 minutes or less)

As it is seen from Table 5 Level 1 is the Basic, Level 2 is Intermediate and Level 3 is Advanced.

According to Avelar et al. (2012: 15-16), to report Level 1, Level 2, or Level 3 measurements, it is necessary to utilize the designated measurement location specified for each respective level.

**Level 1 Basic:** The measurement of IT load at Level 1 is conducted at the output of the Uninterruptible Power Supply (UPS) equipment and can be obtained from the UPS front panel, a meter on the UPS output, or, in cases where multiple UPS modules are present, a single meter on the common UPS output bus. The incoming energy is measured at the utility service entrance, which supplies power to all the electrical and mechanical equipment responsible for powering, cooling, and conditioning the data center. Basic monitoring necessitates the collection of power measurements at least once a month, while energy measurements are recommended to be taken with the same frequency. Typically, this requires some level of human involvement to perform the measurements.

**Level 2 Intermediate:** The measurement of the IT load takes place at the output of the Power Distribution Units (PDUs) located within the data center. This measurement can typically be obtained from the front panel of the PDU or through a meter connected to the secondary side of the PDU transformer. Alternatively, individual branch circuit measurement can also be used for Level 2 monitoring. Like Level 1, the incoming energy is measured at the utility service entrance that supplies power to all the electrical and mechanical equipment required for data center operations. Intermediate monitoring

involves the regular collection of power measurements, ideally daily, and energy measurements are recommended to be taken with the same frequency. Unlike Level 1, Level 2 monitoring involves less human involvement as the data can be collected electronically through meters. Real-time logging of data is expected at Level 2, allowing for comprehensive trend analysis.

**Level 3 Advanced:** The measurement of the IT load at Level 3 is performed at the individual IT equipment level within the data center. This can be achieved through the use of metered rack PDUs, also known as plug strips, which monitor power consumption at the strip or receptacle level. Alternatively, the IT devices themselves can provide the necessary measurement. It is important to exclude non-IT loads from these measurements to obtain accurate results. Like Levels 1 and 2, the incoming energy is measured at the utility service entrance that supplies power to the entire data center infrastructure. Advanced monitoring requires more frequent data collection, with power measurements ideally taken once every 15 minutes or even more frequently, and energy measurements recommended to be collected with the same frequency. At Level 3, data collection should not rely on human activity; instead, automated systems are used to gather and record data in real-time. This advanced monitoring should support extensive trending and analysis of the collected data. The main challenge lies in collecting the data in a user-friendly format that meets the needs of various stakeholders and enables comprehensive aggregation for a complete overview of the data center's energy performance.

To ensure consistency and comparability in measurements, it is advised to take Level 1 and Level 2 measurements at around the same time of day. This helps to capture the data center's energy consumption when the load is relatively consistent compared to previous measurements. Additionally, when making week-to-week comparisons, it is important to keep the day of the week constant to ensure accurate and comparable results. By maintaining these practices, the measurement processes can provide reliable data for assessing changes and improvements in energy consumption over time.

Figure 17 below illustrates a standard data center layout and defines specific measurement points corresponding to the PUE measurement levels, listed in Table 5.

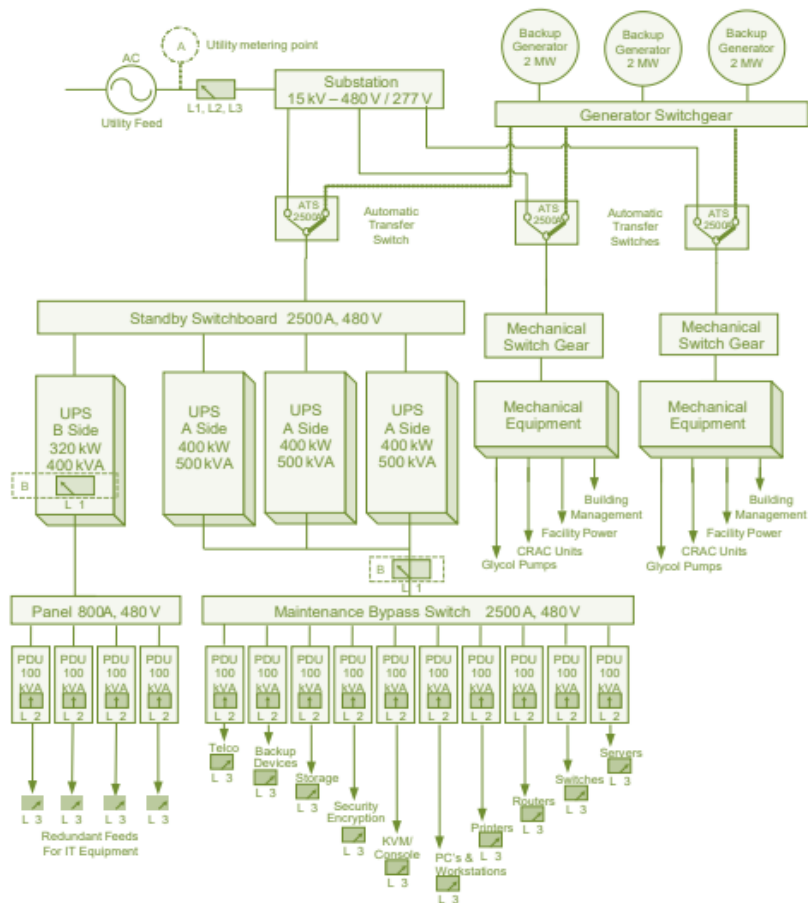


Figure 17. Graphical representation of the three PUE measurement levels for a typical data center. (Avelar et al., 2012: 15)

The Green Grid recommends an automated real time monitoring of PUE value with the data collected every 15 minutes or less. It is suggested that data centers should use the average PUE value measured over a one-year period, when reporting the metric. (Avelar et al., 2012: 21)

Table 6 shows details regarding the required and recommended measurement points and intervals.

Table 6. Guidance as to which measurement points and intervals are required and recommended for each PUE measurement level. (Avelar et al., 2012: 22)

Where do I measure?		Level 1 (L1)	Level 2 (L2)	Level 3 (L3)
How often do I measure?		Basic	Intermediate	Advanced
<b>IT Equipment Energy</b>	Required	UPS outputs	PDU outputs	IT equipment input
<b>Total Facility Energy</b>	Required	Utility inputs	Utility inputs	Utility inputs
	<b>Additional recommended measurements*</b>		UPS inputs/outputs Mechanical inputs	PDU outputs UPS inputs/outputs Mechanical inputs
<b>Measurement Intervals</b>	Required	Monthly	Daily	15 minutes
	<b>Additional recommended measurements*</b>	Weekly	Hourly	15 minutes or less

\* Recommended measurements are in addition to the required measurements. The additional measurement points are recommended to provide further insight into the energy efficiency of the infrastructure.

Rădulescu et al., (2017) explains the need in continuous or periodic measurement of PUE, for achieving relevant value.

Regardless the fact, that PUE metric is widely recommended, it also has industry-acknowledged drawbacks, which will be discussed further in this chapter.

#### 4.5.3 Critical Analysis of Power Usage Effectiveness (PUE)

The overall efficiency of a data center is determined by the amount of useful work produced in relation to the energy consumed. However, assessing useful work can be challenging due to the varying tasks performed by different data centers and their reliance on external factors. As a result, the Power Usage Effectiveness (PUE) metric gained popularity because it measures efficiency based on the ratio of energy available for useful work to energy lost as overhead (i.e., infrastructure). This metric was widely adopted in the industry after its introduction. Since energy consumption is a significant expense for data centers, regardless of their specific focus, reducing infrastructure

energy consumption is a common requirement to enhance efficiency. (Van De Voort et al., 2017: 6-7)

Rădulescu et al. (2017) defines several limitations of Power Usage Effectiveness metric. Notable critiques emphasize that PUE cannot effectively compare data centers of varying sizes, configurations, technologies, and geographic locations. It also fails to consider the regional climate where data centers are situated. For example, the comparison of a DC located in a Nordic country to one in a tropical region cannot be valid, as in colder climates the need for extensive cooling systems is reduced. Cooling systems consume around 30 percent of the energy in facility, while DC equipment consumes approximately 50 percent. Hence, a data center located in a tropical region may have a higher final PUE compared to a data center in a Nordic country, even if overall the first operates more efficiently.

The PUE metric has other limitations that need to be taken into consideration. First, it is crucial to measure PUE when the IT equipment is operating at full capacity. Otherwise, estimated values will lack practical meaning. Another drawback is that PUE does not account for the size of investments made to improve energy efficiency, which can impact the interpretation of the indicator. Moreover, virtualization can affect PUE as reductions in power consumption by IT equipment may not proportionally decrease the power used by infrastructure components, potentially leading to higher PUE values. Inefficient cooling infrastructure due to under-loaded conditions further compounds this issue. It is recommended to include server utilization factors in the PUE metric to provide more informative sustainability insights and avoid confusion when making energy-related decisions based on minimizing PUE. (Rădulescu et al., 2017)

According to Van De Voort et al. (2017: 8), one of the key concerns regarding the use of PUE as the primary performance metric is its failure to account for IT equipment efficiency. With advancements in technology, the computational power per watt has been increasing according to Moore's law (Moore, 1965). This means that the amount of useful work produced per watt can double every two years. Consequently, one of the most effective strategies for energy efficiency could be to regularly update and renew IT equipment.

PUE has a restricted focus on energy consumption and does not account for factors such as on-site energy generation or waste heat recovery, as highlighted by The Green Grid (2012).

These aspects, although significant, are not considered within the PUE metric's primary objectives. Additionally, PUE does not monitor or consider the specific energy source being utilized. (Van De Voort et al., 2017)

According to Van De Voort et al. (2017), to achieve a comprehensive evaluation of energy efficiency within the data center industry using performance metrics, it is necessary to expand beyond relying solely on PUE. Additionally, it is crucial to track the efficiency of individual components within the data center. While PUE is effective within its limitations, it is suggested that complementary metrics be used alongside PUE.

#### 4.5.4 Power Usage Effectiveness (PUE) complimentary metrics

According to Makris (2017), PUE does not cover all the elements of the IT infrastructure, leading to the need for redefining metrics that can measure the energy efficiency in a better context. Alternative metrics, such as Power Performance Effectiveness (PPE), Data Center Productivity (DCeP), and Data Center Infrastructure Efficiency (DCIE) are proposed to address the limitations of PUE and provide more comprehensive evaluation of energy efficiency in data centers.

Cappiello et al. (2013) states that the Power Usage Effectiveness (PUE) is a commonly used metric to assess the energy efficiency of a data center by measuring energy waste. However, while PUE provides a quick and intuitive way to classify data centers based on their greenness, it has limitations in identifying specific areas of improvement and interventions to mitigate energy waste. Recognizing this, the Green Grid Consortium has expanded the set of metrics to cover other factors such as IT system productivity, efficiency, and resource saturation. This broader set of metrics provides a more comprehensive understanding of data center performance beyond PUE alone, allowing for a more detailed assessment of critical elements and potential interventions.

PUE is effective within its constraints, but there are complementary metrics, which may provide a more comprehensive evaluation. Table 7 presents a synthesis of these metrics.

Table 7. Complementary metrics to PUE. (Rădulescu et al., 2017)

Name and Acronym	Unit	Type	Opt	Formulas
Carbon Usage Effectiveness CUE	$\frac{kgCO_2}{kWh}$	Min	0	$CUE = \frac{Total\_CO2\_Emmission_{DC}}{IT\_equipment\_Energy}$
Energy reuse ERF	Percent	Max	1	$ERF = \frac{Reuse\_energy}{Total\_Energy}; 0 \leq ERF \leq 1$
Water usage WUE	Liters/kWh	Min	0	$WUE = \frac{Water\_usage}{IT\_Energy}; 0 \leq WUE$
IT Equipment Utilization ITEU	Percent	Max	1	$ITEU = DC\_IT\_Equip\_util\_rate = \frac{Total\_measured\_energy\_IT\_equip}{Total\_specification\_energy\_IT\_equip}$
IT Equipment Energy Efficiency ITEE	Capacity/ kWh	Max	$\infty$	$ITEE = \frac{TotalIT\_equip\_ratedWorkcapacity}{Total\_rated\_energy\_ITequip}$
Green Energy Coefficient GEC	Percent	Max	1	$GEC = \frac{Total\_Green\_Energy_{DC}}{Total\_Energy\_Comsumtion_{DC}}$
Data Center Performance per unit Energy DPPE	Work/ kWh	Max	$\infty$	$DPPE = \frac{DC\_Work}{Carbon\_Energy} =$ $= ITEU * ITEE * \frac{1}{PUE} * \frac{1}{1 - GEC}$
Energy Reuse Effectiveness ERE	Percent	Min	0	$ERE = \frac{Total\_Energy\_Reused}{IT\_equipment\_Energy} =$ $= (1 - ERF) * PUE$
Data Center Infrastructure Efficiency DCiE	Percent	Max	1	$DCiE = \frac{1}{PUE}$
Data Center Productivity DCP	Useful work / Watt	Max	$\infty$	$DCP = \frac{Useful\_work}{Total\_facility\_Power}$

As it is seen from Table 7, each metric is associated with the formula for calculating the value. According to Rădulescu et al. (2017), there is a need for a globally accepted metric for energy efficiency in computer servers, as stakeholders show great interest. The implementation of the European Ecodesign Directive led to a detailed analysis of policy requirements, including energy use and performance. A new server metric based on Server Efficiency Rating Tool (SERT) was developed. In another study, a summary of air conditioning energy performance in 100 DCs was provided along with energy efficiency metrics and benchmarks. The state-of-the-practice of power measurements and energy efficiency techniques in the data center industry was investigated through interviews. Two new metrics, power variability (PVar) and infrastructure power adaptability (IPA), were proposed to analyze the relationship between IT and power infrastructure. Current approaches for assessing operational efficiency in DCs are mostly static and fail to reflect flexibility and optimization effects. New metrics were proposed for DC energy efficiency evaluation, including Adaptability Power Curve (APC), Grid Utilization Factor (GUF), and CO<sub>2</sub> avoided emissions (CO<sub>2</sub> Savings). While PUE has

increased the number of energy efficient DCs, it lacks performance indicators for IT efficiency, water usage, heat recovery, on-site energy generation, and carbon impact.

Despite its shortcomings, the result of the profound analysis, conducted by Rădulescu et al. (2017), shows that PUE remains an important metric for evaluating data center energy efficiency.

#### 4.6 Industry Standards and Best Practice

This section discusses industry standards for assessment of energy consumption in cloud environments and best practice. It covers Google Cloud Carbon Footprint reporting methodology and Cloud Carbon Emissions Measurement and Analysis Tool, which is an open-source software solution, based on the industry leading methodologies to convert cloud utilization data into estimated energy usage and carbon emissions and

There are several widely recognized frameworks and guidelines for assessment of energy consumption in cloud environments.

The Green Grid, a global consortium of technology companies, data center operators and other industry stakeholders, focuses on promoting resource-efficient and sustainable data centers. It promotes best practices, metrics, and standards for data center design, operation, and efficiency. The Green Grid conducts researches and collaborates with industry experts, provides educational resources to help organizations to optimize data center energy efficiency, reduce environment impact, and improve overall sustainability. The Green Grid provides a range of metrics tailored for data center energy efficiency, some of which are applicable to virtualized environments. These metrics include Power Usage Effectiveness (PUE), Data Center Infrastructure Efficiency (DCiE), and Water Usage Effectiveness (WUE). PUE quantifies the ratio of total data center energy usage to the energy consumed by IT equipment, while DCiE represents the inverse of PUE, highlighting the efficiency of IT equipment. WUE focuses on water consumption, which indirectly affects energy consumption. The consortium especially highlights the importance of Power Usage Effectiveness (PUE) in data centers. (Avelar et al., 2012)

Currently, there is a lack of widely accepted metrics for measuring and monitoring energy consumption in both data centers and cloud environments. To address this, the Standard Performance Evaluation Corporation (SPEC) has introduced a power and performance methodology. This methodology allows designers and implementers of performance benchmarks to incorporate a power component, providing a more comprehensive assessment of energy consumption. (Cappiello et al., 2013)

The European Code of Conduct for Data Centres is an initiative by the European Commission, which offers best practices for enhancing energy efficiency in data centers. It encourages data center operators to measure and report their energy efficiency using metrics such as PUE. (European Commission, 2022)

Additionally, cloud service providers (CSPs) like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform offer documentation and guidelines for optimizing energy usage within their respective cloud environments. These guidelines can provide insights into measuring and managing energy consumption at the virtualized resource level. Google Cloud's revealed high-level methodology for the assessment of will be reviewed further in this section in detail.

#### 4.6.1 Google Cloud Carbon Footprint Reporting Methodology

According to Google Cloud (2023), it has developed the Carbon Footprint Reporting Methodology to provide customers with insight into the greenhouse gas emissions linked to their purchases from Google Cloud. The approach involves examining the total carbon emissions generated by the computing infrastructure that supports Google's internal services. These emissions are then assigned to specific Google Cloud products and further allocated to individual customers based on their product usage. Through the Carbon Footprint report, customers receive personalized data on their gross greenhouse gas emissions, enabling them to evaluate the environmental effects of their Google Cloud usage and make efforts to mitigate those impacts. The energy consumption associated with Google Cloud products and services plays a crucial role in determining the greenhouse gas emissions calculated using this methodology. The methodology considers the energy usage of the computing infrastructure, encompassing elements such as data centers, servers, and other hardware components. By analyzing the energy consumption of the infrastructure, Google can estimate the overall carbon emissions produced during the operation of its services, with energy consumption serving as a

significant contributor to the total greenhouse gas emissions associated with Google Cloud products. The Carbon Footprint report incorporates this energy consumption data and assigns the corresponding emissions to each Google Cloud product, allowing customers to gain insights into the carbon footprint related to their specific product usage and take actions to reduce their environmental impact.

The Carbon Footprint reports adhere to the established and widely accepted Greenhouse Gas Protocol (GHGP) standards for carbon reporting and accounting, which offer comprehensive guidelines for accurate emission reporting. Carbon Footprint primarily uses the GHGP's location-based reporting method, which reflects all electricity generation sources in each location, excluding Google's renewable power purchases. It helps customers understand their impact on greenhouse gas emissions without considering Google's carbon-free electricity. In the future, Carbon Footprint will offer a market-based footprint option that aligns with GHGP's rules, allowing customers to calculate Scope 3 emissions inventories and assess their own products and services more effectively. The Carbon Footprint report includes emissions from various activities categorized under Scope 1, Scope 2, and Scope 3. Scope 1 emissions encompass on-site combustion of fossil fuels, such as diesel for backup power and natural gas for heating, as well as fuel usage in fleet vehicles. Scope 2 emissions include electricity use related to Google Cloud products, including compute and networking equipment, as well as ancillary services like cooling and lighting in both Google-owned and externally owned data centers (calculated based on location and market factors). Scope 3 emissions cover the upstream lifecycle emissions of data center equipment and buildings, business travel and commuting of Google data center employees. The report excludes emissions from certain activities. These exclusions consist of electricity generation losses "during transmission and distribution, extraction and transportation of fuels for grid electricity generation, lifecycle emissions of generation facilities and equipment, fugitive emissions from HVAC system coolants, emissions from small equipment deployments at internet service providers' partners, emissions from Google networking equipment outside data centers, and downstream end-of-life emissions of data center equipment and buildings". The report considers the first two exclusions potentially relevant but relatively small compared to reported emissions, while estimating the remaining four exclusions as not significant for Google Cloud customer reporting. Carbon Footprint reporting process is automated by Google Cloud. This section further outlines the method used by Google Cloud to perform these calculations. Google Cloud operates as a shared computing platform where compute resources like processing power, memory, storage, and

networking are shared among its customers. The platform is built around internal services, which are specific software functionalities running on Google's data center machines. These internal services are used by Google Cloud products, which are then consumed by customers as product units. (Google Cloud, 2023)

One significant source of greenhouse gas emissions for Google Cloud is electricity consumption. Data centers consolidate compute resources in shared buildings, which require electricity to power the computing equipment as well as lighting, cooling systems, power infrastructure, and other ancillary needs. The electricity used by Google Cloud comes from various generation plants operating on different grids worldwide. The greenhouse gas emissions associated with electricity generation depend on factors such as the type of fuel used (for instance, natural gas, coal, renewable sources like wind, solar, and hydro) and other factors unique to each grid. The generation sources on a grid can also change throughout the day. Determining the precise electricity usage and resulting carbon footprint of specific products and customers in Google Cloud poses a technical challenge. This is due to the complex nature of shared resources utilized to meet customer compute requirements. To address this, Google Cloud has developed new methodologies and assumptions to allocate and apportion the carbon footprint appropriately. These efforts enable Google Cloud to provide customer-specific footprint reports that accurately represent the environmental impact of each customer's cloud computing usage and product selections. The Carbon Footprint process begins by calculating energy usage based on compute usage and data center resource requirements. It then determines location-based carbon emissions from electricity usage and distributes those emissions among customers and their purchased products. Additionally, emissions from non-electricity sources are included in the calculations. To account for Google's clean electricity purchases, the market-based electricity footprint option matches these purchases with relevant data center loads, establishing regional market-based emission factors. These factors replace the location-based emission factors in the emission reports. (Google Cloud, 2023)

Energy usage allocation involves evaluating dynamic power (energy used when running a workload) and idle power (energy used when machines are idle). The dynamic power is allocated to internal services based on their CPU usage, while idle power is allocated based on resource allocation in the data center. Overhead energy use, such as power systems, cooling, and lights, is allocated to each machine and its users. Google's shared infrastructure services track the usage of other internal services, enabling the

reallocation of energy consumption based on relative usage. If there is insufficient usage data, internal costs are used for reallocation. After completing the calculations and allocations, power use is allocated to each internal service in each data center on an hourly basis. Greenhouse gas emissions are calculated by multiplying location-specific energy use by a grid electricity carbon emission intensity factor. The emissions are reported in kilograms of carbon dioxide equivalent (kg CO<sub>2</sub>e). The emission factor data is obtained from Electricity Maps, with country-specific factors from the International Energy Agency used when necessary. To determine emissions, the energy use for each internal service at each location is multiplied by the appropriate carbon emission intensity factor, resulting in a location-based electricity carbon footprint. For electricity footprint allocation to SKUs (customer-facing product units), the gross carbon footprint of each SKU is divided among them based on their usage and list prices. Different carbon intensities in each location are also considered. This allocation is solved using a series of equations that satisfy specific principles. Further allocation of electricity footprints to customers involves dividing each SKU's carbon footprint by its total usage in a region to establish a per-usage carbon intensity factor. This factor is then multiplied by each customer's SKU usage in that region to determine the per-SKU, per-region, per-customer footprint. These footprints are aggregated into customer-specific Google Cloud product footprints at a monthly granularity. Validation "is performed to ensure that the aggregation of customer electricity carbon footprints equals the total Google Cloud location-based electricity carbon footprint. The market-based electricity footprint is calculated by aligning Google's clean electricity purchases with data center loads according to GHGP rules". Annual market-based emissions are calculated based on actual clean electricity generation and usage at each site. Scaling factors are used to adjust location-based emissions to market-based emissions in regions where clean electricity contracts are in place. The market-based scaling factor is updated annually, and the reports represent the renewable energy activity from the previous year. In regions without clean electricity contracts, location-based and market-based carbon emissions are identical. Energy consumption calculations involve recording the internal services and power use by machines on an hourly basis. Dynamic power is allocated based on CPU usage, while idle power is allocated based on resource allocation. Overhead load is allocated to machines, and the power required by shared infrastructure services is allocated based on their usage by higher-level internal services. (Google Cloud, 2023)

According to Google Data Centers (2023), Google takes a comprehensive approach to measuring Power Usage Effectiveness (PUE) to accurately assess its data center energy efficiency. They include the performance of their entire global data center fleet, not just the newest and best facilities. Measurements are taken continuously throughout the year, accounting for variations in different seasons. Google's PUE calculations also incorporate all sources of overhead, ensuring a comprehensive efficiency metric. While they could report lower numbers by adopting looser interpretations of the Green Grid's PUE measurement standards, Google chooses a higher standard. They believe it's important to measure and optimize all aspects of their sites, rather than just a part of them. As a result, Google reports a comprehensive trailing twelve-month (TTM) PUE of 1.10 across all their large-scale data centers, considering stable operations, all seasons, and all sources of overhead. (Google Data Centers, 2023)

#### 4.6.2 Cloud Carbon Emissions Measurement and Analysis Tool

Cloud Carbon Footprint is an open-source project, sponsored by Thoughtworks Inc. Freely available Cloud Carbon Emissions Measurement and Analysis tool, developed within the frames of Cloud Carbon Footprint project, offers transparency and resources for measuring, tracking, and minimizing the environmental impact of user's cloud operations. By employment of established methods, cloud utilization data is transformed into estimates of energy consumption and carbon emissions. This allows to generate metrics and projections of carbon reduction, which can be shared with employees, investors, and other relevant stakeholders. (Cloud Carbon Footprint, 2023)

"The term open-source more generally refers to a community-based approach to creating any intellectual property (such as software) via open collaboration, inclusiveness, transparency, and frequent public updates." (IBM, 2023)

The Cloud Carbon Emissions Measurement and Analysis tool is hosted on GitHub platform, "code hosting platform for collaboration and version control." (W3schools, 2023)

The code and methodology of Cloud Carbon Footprint are openly accessible and transparent, continuously enhanced through feedback and evaluations from experts within the open-source and sustainability communities.

Figure 18 below shows contributions to Cloud Carbon Emissions Measurement and Analysis tool development, excluding merge commits and bot accounts.

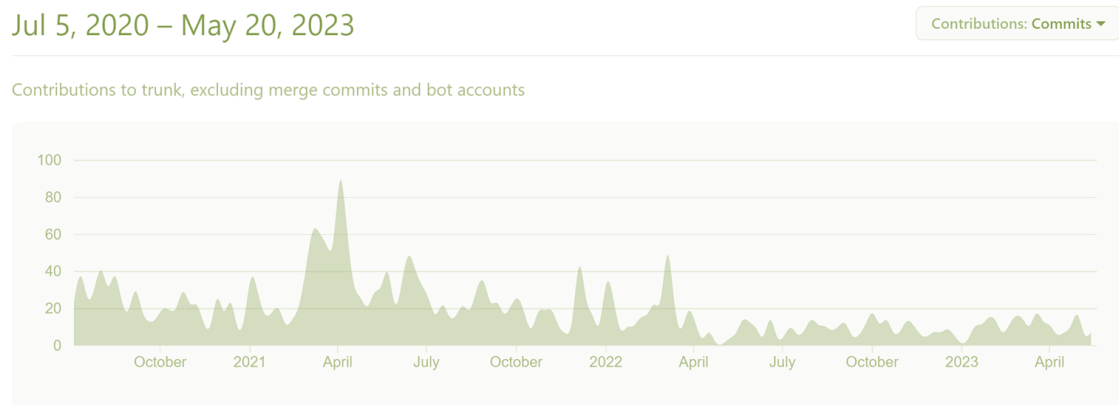


Figure 18. Number of contributors for the development of an application. (GitHub, 2023)

According to Cloud Carbon Footprint (2023), the Cloud Carbon Emissions Measurement and Analysis application retrieves utilization information (such as computing, storage, and networking) from leading cloud service providers and uses it to determine the approximate amount of energy consumed (measured in Watt-Hours) and greenhouse gas emissions in terms of carbon dioxide equivalents (metric tons CO<sub>2</sub>e). The resulting visualizations are presented in a user-friendly dashboard, allowing developers, sustainability advocates, and other individuals involved in an organization to analyze the data and make informed decisions. It provides estimates for energy and carbon emissions across all cloud usage, including embodied emissions from manufacturing, and allows users to analyze emissions by cloud provider, account, service, and time period. The tool also offers actionable recommendations for reducing cost and carbon emissions specifically for AWS and Google Cloud, along with projected savings and the real-world impact. With the ability to measure usage at a granular service level and utilize actual server utilization for hyperscale data centers, it visually monitors energy usage and carbon footprint through graphs and charts, enabling the export of metrics in CSV format for easy sharing with stakeholders. Additionally, Cloud Carbon Footprint provides multiple integration options to incorporate energy and carbon metrics into existing data sets, pipelines, monitoring systems, or dashboards. Its open and extensible nature allows for potential future expansion to include other cloud providers, on-premises infrastructure, or co-located data centers. (Cloud Carbon Footprint, 2023)

To estimate the energy and carbon emissions associated with cloud provider usage, a process of classifying each row of usage is followed. To classify the usage types, first the pricing or usage unit is examined, such as hours or seconds for AWS and usage units for GCP and Azure. If the unit is hours or seconds, it is typically categorized as a Compute usage type. If it is byte-seconds or Gigabytes-Months, it is likely to be related to Storage, and if it involves bytes or Gigabytes, it is probably associated with Networking. Other units are generally disregarded. Additionally, the classification is validated by examining the more detailed usage type, which helps determine if it falls under Compute, Storage, Networking, or Memory categories. If a row's classification is unknown, there is a reclassifying procedure. If a row is unsupported, the application disregards it. After classifying the usage row, the application considers the associated usage amount when estimating energy and carbon emissions. To calculate the energy consumption, the Kilowatt-hour/Cost coefficient is used. For each "known" type of usage, an average coefficient is constructed by dividing the kilowatt-hour per usage amount (or cost for AWS) for each unique combination of service and usage unit. The coefficient is then applied by multiplying it with the usage amount (or cost for AWS) for any unidentified rows. This enables the estimation of the energy consumption for unknown rows. When encountering an unknown row, a specific process is followed. If there is a known ratio of usage amount to kilowatt hours for the same service and usage unit, the usage amount is multiplied by this ratio. This calculation helps to estimate the kilowatt hours consumed. Subsequently, the value is converted to CO<sub>2</sub>e (carbon dioxide equivalent) using Google's published grid emission data and the percentage of carbon-free energy specific to each data center. The methodology suggests that once the kilowatt hours for compute, storage, and networking are estimated, the next step is to multiply these values by the cloud provider's Power Usage Effectiveness (PUE). (Cloud Carbon Footprint, 2023)

According to Carbon Cloud Footprint (2023), PUE is a measure of energy efficiency for data centers, where a score of 1 indicates that all consumed energy is utilized solely for powering the servers without any waste on cooling. The PUE information is obtained from publicly available data provided by the cloud providers. For instance, Google Cloud Platform (GCP) discloses their PUE score. Organizations are progressively utilizing the power usage effectiveness (PUE) measure to assess the energy efficiency of prospective cloud service providers. (Cloud Carbon Footprint, 2023)

#### 4.7 Conceptual Framework of Thesis

Table 8 shows the Conceptual Framework of this thesis for the assessment methodology for energy consumption of virtualized resources in data centers. The Conceptual Framework consists of three major themes discussed earlier in this work: *Energy Consumption Metrics for Virtualization Layer*, *Data Center Energy Efficiency Metrics*, and *Industry Standards and Best Practice: Methodology for Energy Consumption Assessment from Cloud Usage*. By integrating these themes, the Conceptual Framework provides a structured approach to guide the development of assessment methodology for energy consumption of virtualized resources in data centers.

Table 8. Conceptual framework of this thesis.

Energy Consumption Metrics for Virtualization Layer	Data Center Energy Efficiency Metrics	Industry Standards and Best Practice: Methodology for Energy Consumption Assessment from Cloud Usage
<b>Sources</b>		
<p>(Portnoy, 2016) (Lee et al., 2018) (Haghshenas et al., 2022) (Gul et al., 2020) (Hieu et al., 2017) (Povoa et al., 2013)</p>	<p>(Makris, 2017) (Rădulescu et al., 2017) (Avelar et al., 2012) (Belady and Seander, 2018) (Dumitrescu et al., 2018) (Jaureguiualzo, 2011) (Van de Voort et al., 2017) (Gough et al., 2015) (Pitalier, 2018)</p>	<p>(Cloud Carbon Footprint, 2023) (Google Cloud, 2023) (Google Data Centers, 2023) (Avelar et al., 2012) (European Commission, 2022) (Cappiello et al., 2013)</p>
<b>Answering the questions</b>		
<p>1.What are the energy consumption metrics used for measuring virtualized resources in data centers?</p> <p>2.How do these metrics capture the energy usage of virtualized resources?</p> <p>3.What are the limitations and challenges associated with measuring energy consumption in the virtualized layer?</p>	<p>1.What are the commonly used energy efficiency metrics in data centers?</p> <p>2.How are these metrics defined and measured?</p> <p>3.How do these metrics contribute to understanding and improving energy consumption in data centers?</p>	<p>1.What are the industry standards and best practices for assessing energy consumption in cloud environments?</p> <p>2.How can these best practices be adapted and applied to assess the energy consumption of virtualized resources in data centers?</p> <p>3. How can the insights from industry best practices contribute to advancing the current methodology for energy consumption assessment in virtualized data centers?</p>

As it is seen in Figure 19, the questions presented in each block of the Conceptual Framework cover key aspects related to the assessment of energy consumption by virtualized resources within data centers. The questions address the metrics, methodologies, industry standards and best practices. This knowledge provides the

basis necessary to develop the assessment methodology of energy consumption by virtualized resources in data centers for further usage in the case company.

First element of the Conceptual Framework explores the concept of virtualization, its roles as underlying technology in cloud infrastructure, as well as its benefits. Various factors that influence power consumption in virtualized environments are explored, including CPU, memory, hard disk, I/O, and network utilization. These metrics provide insights into resource utilization and workload of virtual machines hosted on physical servers. These metrics capture the energy usage of virtualized resources by monitoring the utilization and activities of different components. The utilization of CPU reflects the workload and operations performed by the processor. The memory usage indicates the energy consumed by RAM modules. Disk I/O and network traffic measurements help assess the energy consumption related to storage and network operations. By tracking these metrics, it becomes possible to estimate energy consumption of virtualized resources. However, the need for comprehensive approaches that consider multiple resource dimensions is emphasised. This part of the Conceptual Framework also covers the complexity of measuring energy consumption in dynamic and diverse cloud environments, particularly in multi-tenant scenarios where energy usage is shared among customers or user accounts. It is challenging to accurately attribute energy consumption to individual virtual machines due to resource sharing among multiples VMs. The hypervisor layer complicates even more the energy measurement, as it is problematic to differentiate its consumption from the virtual machine's energy usage. Energy consumption can be impacted by the high virtual machine density on physical servers, so it becomes challenging to accurately measure energy usage in densely populated virtualized environments. It is also pointed out in this block, that the lack of standardized approaches and tools for measuring energy consumption in virtualized environments constrain benchmarking across virtualization platforms.

Second element of the Conceptual Framework focuses on data center energy efficiency metrics. One of the widely recognized and used in the industry metrics is Power Usage Effectiveness (PUE). Other commonly used energy efficiency metrics, which provide insights to the overall energy efficiency and effectiveness of data center operations, include Power Usage Effectiveness (PUE), Data Center Infrastructure Efficiency (DCiE), and Energy Reuse Effectiveness (ERE). PUE is calculated by dividing the total power consumed by the data center by the power consumed by the IT equipment alone. A lower PUE indicates higher energy efficiency, with the ideal value being 1.0. DCiE is the

reciprocal of PUE, calculated by dividing the power consumed by the IT equipment by the total power consumed by the data center. Higher DCiE values indicate better energy efficiency, with a perfect efficiency rating being 100%. ERE measures the extent to which waste heat generated by the data center is reused or recycled. It is calculated by dividing the total energy consumed by the IT equipment by the energy reused within the data center or elsewhere. A higher ERE indicates better utilization of waste heat. By regularly monitoring these metrics, data center can track the energy consumption patterns and identify opportunities for optimization, compare energy efficiency with industry standards and best practices, implement energy-saving measures to reduce overall energy consumption and operational costs, as well as explore opportunities for waste heat reuse or recycling.

Third element of Conceptual Framework discovers industry standards for assessing energy consumption in cloud environments, including widely recognized frameworks and guidelines, and best practices, such as Google Cloud Carbon Footprint reporting methodology and an open-source Cloud Carbon Emissions Measurement and Analysis Tool. The Green Grid is a global consortium, which promotes sustainable data centers, offers best practices and standards. Metrics like PUE, DCiE, and WUE assess energy and water efficiency in virtualized environments. The Standard Performance Evaluation Corporation (SPEC) introduced a power and performance methodology to assess energy consumption in data centers and cloud environments. The European Code of Conduct for Data Centres provides energy efficiency best practices for data centers and clouds. Cloud providers like AWS, Azure, and Google Cloud offer guidelines for optimizing energy usage in their environments. Google Cloud adheres to GHGP standards and assigns emissions to specific products and customers based on usage. It calculates energy usage and carbon emissions based on compute and resource requirements. It considers clean energy purchases and provides location-based and market-based footprint options. The Cloud Carbon Emissions Measurement and Analysis Tool measures and tracks cloud environmental impact. It estimates energy consumption, carbon emissions, and offers visualizations and integration options.

The existing knowledge and best practice findings contribute to the development of assessment methodology for energy consumption by virtualized resources in DCs, which is discussed in Section 5 of this thesis, dedicated to the Proposal Building stage.

## **5 Building Proposal for Assessment Methodology of Energy Consumption by Virtualized Resources in Data Centers for the Case Company**

The Proposal Building stage, covered in this chapter, is based on the findings from the Current State Analysis, Conceptual Framework, built on Existing Knowledge and Best Practice, and the Data Collection 2, which includes internal cooperation and discussions with the stakeholders. This chapter consists of three sections, an overview of the proposal building stage, findings derived from Data Collection 2 and Initial Proposal.

### **5.1 Overview of the Proposal Building Stage**

Proposal Building stage of this thesis consists of several steps, which are discussed in this section. The development efforts in the proposal building were based on weaknesses associated with the accuracy of the reports found during the Current State Analysis of the sustainability reporting process. The focus for the improvement was kept on the two stages of the reporting process: *Data Evaluation* and *Data Processing*, on which the metrics for the KPIs are selected and the calculation methodology is chosen. The emphasis of this study was done on refining the metrics and assessment methodology for potentially more accurate calculation of energy consumption by virtualized resources in data centers and ensuring that the calculations align with industry standards and best practices. The selected scope of the thesis was the rational guiding the choice of the main themes for the Existing Knowledge and Best Practice review, which lead to the choice of major elements for Conceptual Framework, relevant for Proposal Building.

First, the Initial Proposal for the assessment methodology of energy consumption by virtualized resources in data centers was created, based on the insights from the cooperation with the stakeholders on the Current State Analysis stage and relevant findings from the Existing Knowledge and Best Practice. The Power Usage Efficiency (PUE) metric, which is included to the Proposal, was introduced to the stakeholders, as a part of the Initial Proposal. The metric was thoroughly reviewed by the Head of Sustainability department and Service Owner, Sustainability Hub and immediately incorporate to the leveraged in the case company assessment methodology of energy consumption by virtualized resources in data centers, which is based on volumes of allocated resources, such as CPU, RAM, and storage. Co-operation with the stakeholders helped to shape the Initial Proposal.

Second, the Data Collection 2 was carried out during comprehensive discussion and exchange of ideas regarding the proposal. Stakeholders were encouraged to provide the feedback, share their thoughts, and raise any concerns or suggestions they had. This step aimed to ensure that the proposal considered diverse perspectives and expertise of the stakeholders.

Third, the Initial Proposal was updated based on the feedback received from the stakeholders during Data Collection 2 round. The feedback was carefully analyzed, and relevant suggestions and recommendations were incorporated into the proposal.

Figure 19 illustrates the logical steps that were taken to build the Proposal.

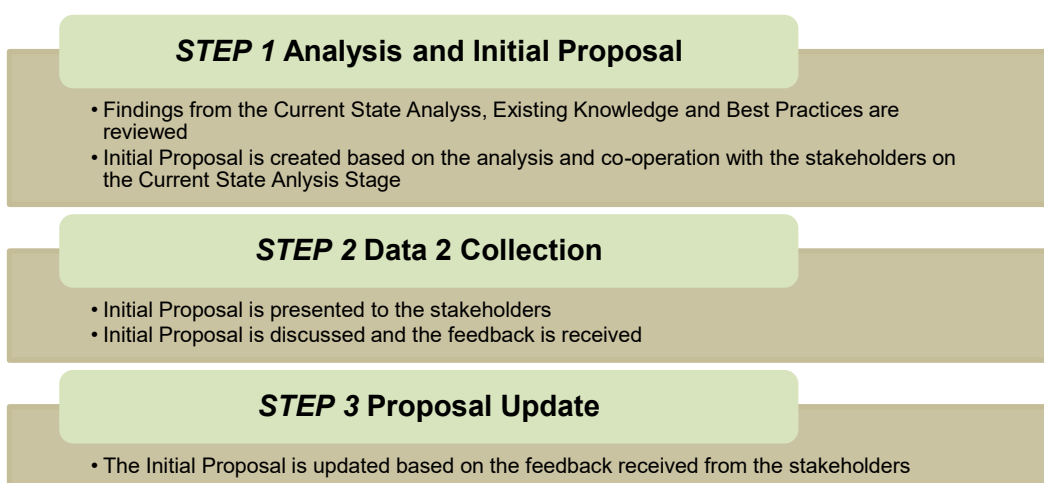


Figure 19. Proposal Building steps.

## 5.2 Findings from Data 2

This section presents the significant contributions and suggestions provided by stakeholders for the development of the Initial Proposal. These inputs are derived from three sources: Data Collection 1, which consists of findings from the Current State Analysis, input from existing knowledge and best practices, on which Conceptual Framework is built, and Data Collection 2, which includes suggestions gathered from the key stakeholders during the co-creation process.

The stakeholders' suggestions are summarized in Table 9, which relates them to the selected focus areas from the CSA and the inputs from literature and best practices. The

suggestions cover several aspects such as the selection of metrics, the choice of calculation methodology, reviewing of best practices, and ensuring transparency to customers.

Table 9. Key stakeholder suggestions (findings of Data 2) for Proposal building in relation to findings from the CSA (Data 1) and the Conceptual framework.

#	Key focus area from CSA (from Data 1)	Input from literature (CF)	Suggestions from stakeholders for the Proposal, summary (from Data 2)	Description of their suggestion (in detail)
1	<i>Data Evaluation</i> stage of the Sustainability Reporting process: The selection of the metrics for the assessment of energy consumption by virtualized resources in Data Centers.	<p><i>Energy Consumption Metrics for Virtualized Layer:</i></p> <p>CPU utilization, memory usage, disk I/O, and network traffic.</p> <p><i>Data Center Energy Efficiency Metric:</i></p> <p>Power Usage Efficiency (PUE).</p> <p><i>Adoption of Industry Standards and Best Practices:</i></p>	a) Suggested by the thesis researcher PUE metric and presented open-source application for Cloud Carbon Emissions Measurement and Analysis, as well as described energy assessment methodology for cloud usage should be reviewed	<p>When the PUE metric was presented in the frames of assessment methodology leveraged by Clod Carbon Footprint project, Respondent 4 commented:</p> <p>"I agree with methodology when it comes to operational emissions, but I'm not really sure what PUE actually is, but this could be elaborated."</p> <p>Reviewing the same assessment methodology description and pointing at PUE, Respondent 3 mentioned:</p> <p>"What does it give to that?"</p>
2	<i>Data Processing</i> stage of the Sustainability Reporting process: The choice of calculation methodology for energy consumption by virtualized resources in Data Centers.	<p>The Green Grid suggests using PUE metric, which is also used by Google Data Centers and incorporated in the energy consumption assessment methodology by open-source Cloud Carbon Emissions Measurement and Analysis Tool.</p>	b) Measurement approaches should be transparent to the customers	Respondent 4 suggests: "The solution shouldn't be complicated. Customers must very easily understand how we make measurements."

As seen from Table 4, the input from stakeholders for the first focus area calls for a review of the PUE metric, the open-source application for Cloud Carbon Emissions Measurement and Analysis, and the energy assessment methodology for cloud usage. Stakeholders are seeking a better understanding of the PUE metric and its benefits in assessing energy consumption. For the second focus area, stakeholders emphasize the need for transparency in measurement approaches, allowing customers to easily comprehend how energy consumption is measured.

These suggestions from stakeholders, aligned with the conceptual framework elements and the findings from the current state analysis, have been taken into consideration

during the proposal building process. By addressing these recommendations, the proposal, which will be discussed in the next section, aims to provide a more comprehensive and transparent approach to assessing energy consumption by virtualized resources in data centers.

### 5.3 Initial Proposal

The Initial Proposal aims to advance the assessment methodology for energy consumption by virtualized resources in data centers. The proposal is based on three key elements: Energy Consumption Metrics for the Virtualization Layer, Data Center Energy Efficiency Metrics, and Adaptation of Industry Standards and Best Practices, which are based on the main themes selected for literature review on the stage of the Existing Knowledge and Best Practice. These topics in their turn were chosen based on the weaknesses identified during Current State Analysis stage and the need for an advanced, more accurate approach to assess energy consumption by virtualized resources in data centers.

#### 5.3.1 Energy Consumption Metrics for Virtualization Layer

Measuring energy consumption in virtualized environments is a complex task due to various factors such as resource sharing, virtualization overhead, dynamic resource allocation, and high virtual machine density. An enhanced methodology for assessing energy consumption in data centers can address these challenges by incorporating multiple resource dimensions, including CPU utilization, memory usage, disk I/O, and network traffic. By considering these multiple dimensions, the methodology provides a more holistic assessment of energy consumption patterns and captures the diverse types of resources and workloads in a cloud environment. This approach enables a more accurate evaluation of energy consumption trends and facilitates granular analysis by individual resource dimensions. By measuring the energy consumption of each dimension separately, such as CPU, RAM, disk, and network, the methodology gains insights into their specific contributions. This granular analysis allows tracking energy usage patterns, identifying energy-intensive areas, and optimizing resource allocation.

However, in a multi-tenant scenario where multiple customers share a virtual machine, accurately attributing energy consumption to individual customers becomes challenging.

Energy consumption in data centers is typically measured at the infrastructure level, making it difficult to allocate energy usage to specific virtual machines or customers in a shared environment. Nevertheless, the approach can still be used to measure the overall energy consumption of the shared virtualized resource by analyzing CPU utilization, memory usage, disk I/O, and network traffic. To address the challenge of energy allocation in multi-tenant scenarios, advanced techniques and algorithms can be employed. These techniques consider factors such as CPU utilization, memory usage, network traffic patterns, and other relevant metrics to estimate the energy consumption of individual customers within the shared environment. Achieving precise energy consumption allocation in such scenarios may require sophisticated monitoring and measurement mechanisms.

Another challenge highlighted by the study is the differentiation of energy consumption between the hypervisor layer and virtual machines (VMs). The current methodology does not account for this distinction, leading to inaccuracies in energy measurement. To overcome this, an improved methodology should incorporate techniques or mathematical algorithms that allow for the separation of hypervisor overhead from VM energy consumption. This distinction is crucial for accurately attributing energy usage to specific virtualized resources. Moreover, the proposed methodology should consider dynamic resource allocation and VM migrations, which impact energy consumption patterns. By accounting these dynamic changes, such as VM migrations and resizing, the methodology can provide a more realistic assessment of energy consumption in virtualized environments. In overall to enhance the methodology incorporating multiple resource dimensions, addressing the challenges of energy allocation in multi-tenant scenarios, differentiating between hypervisor and VM energy consumption, and considering dynamic resource allocation and VM migrations is suggested. These improvements can lead to a more comprehensive and accurate assessment of energy consumption in data centers.

### 5.3.2 Data Center Energy Efficiency Metrics

This study suggests that the assessment methodology of energy consumption by virtualized resources in data centers, used in case company, focuses only on the energy consumed by virtualized resources, without taking into consideration other factors that contribute to overall energy consumption in data centers. This limited perspective does not allow to capture the holistic picture of energy usage. This study suggests eliminating

this drawback in the methodology by incorporating the Power Usage Effectiveness (PUE) metric into the current methodology.

PUE has gained significant recognition as a widely accepted industry standard for measuring data center energy efficiency. It was introduced by the Green Grid, a global consortium focused on promoting energy efficiency in data centers. This endorsement demonstrates the credibility and widespread acceptance of PUE within the industry. Major players in the data center industry, including Google, have adopted PUE as a key metric for assessing energy efficiency. Google, known for its commitment to sustainability, utilizes PUE to evaluate and improve the energy efficiency of its data centers. This demonstrates the practical applicability and effectiveness of PUE as a performance indicator. PUE is also integrated into assessment methodologies and tools designed to evaluate the environmental impact of data centers. The open-source Cloud Carbon Emissions Measurement and Analysis application, explored in this study, leverages PUE as a crucial component in its assessment methodology. This further validates the relevance and value of PUE in evaluating energy consumption. PUE is extensively referenced and recommended in academic studies and master's theses related to data center energy efficiency. These studies highlight PUE as a key metric for evaluating energy consumption and guiding energy optimization strategies in data centers. The inclusion of PUE in academic research reinforces its significance and the need for its integration into assessment methodologies. PUE is also recognized and suggested by initiatives such as the European Code of Conduct for Data Centres. These initiatives provide guidelines and best practices for improving energy efficiency in data centers, and their endorsement of PUE further underscores its importance as a metric for energy assessment.

By incorporating PUE, which is a ratio of the total energy consumed by the data center to the energy consumed by the IT equipment, a more comprehensive and holistic perspective of energy usage and efficiency can be achieved. Including the PUE metric provides insight into the energy consumed by supporting infrastructure components such as cooling systems, power distribution, and networking equipment. These factors are essential contributors to the overall energy consumption of a data center but are often overlooked in the current methodology. By considering PUE, the methodology can account for the energy efficiency of the entire data center infrastructure, enabling a more accurate assessment of energy consumption. Incorporating the PUE metric allows for a better understanding of the energy efficiency of the data center, identifying areas of

improvement and optimization beyond just virtualized resources. It helps data center operators and stakeholders make informed decisions regarding infrastructure upgrades, cooling optimizations, and equipment replacements, leading to improved energy efficiency and reduced environmental impact.

While this study was primarily focused on exploring the PUE metric in greater detail for the assessment of energy consumption by virtualized resources in data centers, it recognizes the potential contributions of other complementary metrics briefly discussed in this thesis in the section “Power Usage Effectiveness (PUE) complimentary metrics”. This demonstrates an awareness of the wider landscape of energy efficiency assessment and acknowledges the ongoing efforts to develop and refine metrics in the field. Considering other metrics alongside PUE can lead to a more comprehensive assessment, helping identify new areas for improvement and enabling informed decisions regarding resource allocation and optimization.

### 5.3.3 Adoption of Industry Standards and Best Practices

This Proposal highlights the importance of adopting industry standards and best practices for assessing energy consumption by virtualized resources in data centers. A global consortium, the Green Grid, plays a significant role in promoting resource-efficient and sustainable data centers. The Green Grid enables organizations to optimize energy efficiency, reduce environmental impact, and improve overall sustainability by offering best practices, metrics, and standards for data center design, operation, and efficiency. Metrics such as Power Usage Effectiveness (PUE), Data Center Infrastructure Efficiency (DCiE), and Water Usage Effectiveness (WUE) provided by the Green Grid are specifically tailored for data center energy efficiency, including virtualized environments.

To address the lack of widely accepted metrics for measuring energy consumption, the Standard Performance Evaluation Corporation (SPEC) introduced a power and performance methodology. This methodology allows designers and implementers of performance benchmarks to incorporate a power component, enabling a more comprehensive assessment of energy consumption. By utilizing SPEC's power and performance methodology, organizations can gain deeper insights into the energy efficiency of their virtualized resources and make informed decisions to optimize energy usage.

The European Code of Conduct for Data Centres, an initiative by the European Commission, offers valuable best practices for enhancing energy efficiency in data centers. This code of conduct encourages data center operators to measure and report their energy efficiency using metrics such as PUE. By aligning with this code of conduct, organizations can demonstrate their commitment to sustainability and contribute to the European Union's energy efficiency goals. Consistency in data center sustainability reporting and benchmarking among industry peers can be achieved by adopting this code of conduct.

Cloud service providers (CSPs) like Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform provide documentation and guidelines for optimizing energy usage within their respective cloud environments. These guidelines offer insights into measuring and managing energy consumption at the virtualized resource level. Leveraging the expertise of CSPs in energy efficiency, organizations can align their sustainability reporting with industry-leading practices and work towards minimizing their environmental footprint. It is important to note that the guidelines and resources provided by these CSPs are subject to updates and advancements as technology and best practices evolve. It is recommended to refer to the official documentation and resources provided by each cloud service provider for the most up-to-date guidelines and recommendations on measuring energy consumption at the virtualized level.

The adoption of industry standards and best practices is crucial for effective data center sustainability reporting. The Green Grid, SPEC, the European Code of Conduct for Data Centres, and CSPs' guidelines provide comprehensive frameworks and metrics for measuring and monitoring energy consumption in virtualized environments. By incorporating these industry standards into the assessment methodology, the case company can ensure better transparency and benchmark its performance.

As it was discussed in the Current State Analysis, the methodology for assessing energy consumption by virtualized resources in data centers is based on attributing purchased volumes of CPU, RAM, or storage to energy consumption. The approach considers the electricity generation sources in each location and calculates emissions accordingly. It also leverages market-based method, which considers emissions from the specific electricity sources chosen by companies. However, this approach has limitations as it lacks granular analysis, real-time monitoring, and consideration of critical resources such as disk I/O and network traffic.

In comparison, Google Cloud's Carbon Footprint Reporting Methodology provides a more comprehensive and sophisticated approach. To advance the case company's methodology, several key elements from Google's approach can be incorporated. First, a more granular analysis should be conducted, capturing detailed usage patterns of individual resource dimensions like CPU, RAM, disk I/O, and network traffic. This would provide a more accurate understanding of energy consumption. Real-time monitoring should also be implemented to ensure up-to-date and accurate data on energy consumption patterns. This can be achieved using monitoring tools and technologies that track resource utilization and energy consumption in real-time. Furthermore, critical resources such as disk I/O and network traffic should be considered in the methodology. These resources significantly contribute to energy consumption and should be measured and analyzed accordingly to provide a comprehensive assessment.

Automation of the analysis process is another important aspect to improve efficiency and accuracy. By automating data collection, calculation, and reporting, the case company can streamline the process, save time, and reduce the chances of errors. However, as the results of the Currents State Analysis showed, the automation of the data processing is already under the development.

In a multi-tenant environment, advanced techniques and algorithms should be explored to accurately allocate energy consumption to individual customers. Factors such as CPU utilization, RAM usage, network traffic patterns, and other relevant metrics can be considered to estimate the energy consumption of each customer within the shared environment. Additionally, differentiating between the energy consumption of the hypervisor layer and the virtual machines (VMs) would provide a more accurate measurement and attribution. This distinction allows for a better understanding of energy usage and the ability to allocate consumption to specific virtualized resources. Incorporating these aspects from Google's methodology will enhance the accuracy, effectiveness, and granularity of the case company's energy consumption assessment.

The Cloud Carbon Footprint project offers an open-source Cloud Carbon Emissions Measurement and Analysis tool that allows users to measure, track, and minimize the environmental impact of their cloud operations. The tool retrieves utilization information from cloud service providers, estimates energy consumption and carbon emissions, and presents the data in a user-friendly dashboard. It provides metrics, projections, and recommendations for reducing cost and carbon emissions, specifically for AWS and

Google Cloud. The methodology involves classifying usage types, estimating energy consumption using coefficients, and converting energy consumption to carbon dioxide equivalents (CO<sub>2</sub>e) using grid emission data and the percentage of carbon-free energy. The tool is transparent, continuously improved through community feedback, and allows for integration with existing systems. Comparing this methodology with the case company's approach, there are several ways the case company's methodology can be advanced. The Cloud Carbon Footprint tool measures usage at a granular service level and utilizes actual server utilization for hyperscale data centers. The case company's methodology currently analyses the purchased volume of CPU, RAM or storage. To improve accuracy and detect trends, the case company can consider measuring and analyzing usage and cost data at a more granular level, such as hourly measurement of RAM consumption. The Cloud Carbon Footprint tool offers multiple integration options to incorporate energy and carbon metrics into existing data sets, pipelines, monitoring systems, or dashboards.

The case company can enhance its methodology by integrating energy consumption and carbon emission data into its existing reporting and analysis systems, and as it was discussed in the Current State Analysis, prototypes are already in use and under the development. One more important aspect, that the Cloud Carbon Footprint tool includes embodied emissions from manufacturing in its estimates for energy and carbon emissions. The case company can enhance its methodology by considering embodied emissions as well. By accounting for the emissions associated with the production and disposal of hardware and infrastructure used in virtualized resources, the case company can provide a more comprehensive assessment of the environmental impact. The Cloud Carbon Footprint methodology also incorporates Power Usage Effectiveness (PUE), a measure of energy efficiency for data centers, to estimate energy consumption. By considering the energy efficiency of the data centers hosting the virtualized resources, the case company can refine its estimates and provide more accurate assessments of energy consumption and carbon emissions.

#### 5.4 Summary of the Initial Proposal

The Initial Proposal aims to improve the assessment methodology for energy consumption by virtualized resources in data centers. It consists of three key elements: Energy Consumption Metrics for the Virtualization Layer, Data Center Energy Efficiency Metrics, and Adoption of Industry Standards and Best Practices.

Figure 21 below illustrates that the Adaptation of Industry Standards and Best Practices is a central element of the Proposal, representing its overarching influence on the other two elements. This visual representation indicates that the industry standards and best practices shape and guide the choice of Energy Consumption Metrics for the Virtualization Layer and Data Center Energy Efficiency Metrics.

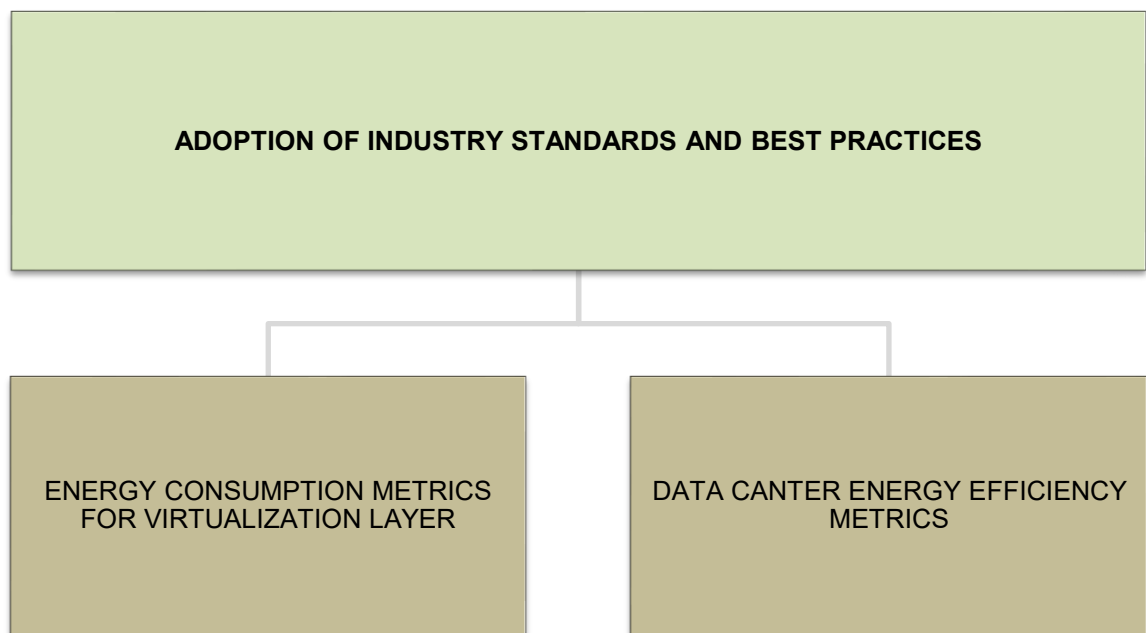


Figure 20. The presentation of key elements of the Proposal.

The first element, Energy Consumption Metrics for the Virtualization Layer, focuses on suggesting energy consumption metrics specifically tailored to virtualized resources. These metrics aim to capture the energy usage of virtualized resources accurately and comprehensively. By developing metrics specific to the virtualization layer, the methodology aims to provide a more granular understanding of energy consumption within this critical component of data center infrastructure.

The second element, Data Center Energy Efficiency Metrics, recommends commonly used energy efficiency metric that is widely used in data centers. This metric offers a holistic and standardized approach to measuring and evaluating the energy efficiency of data center operations. By incorporating this metric into the assessment methodology, data center can gain insights into the overall energy performance and effectiveness of the facilities, which helps to identify areas for improvement and implement energy-saving strategies.

The third element, Adaptation of Industry Standards and Best Practices, focuses on leveraging established industry standards and best practices for assessing energy consumption in cloud environments. These practices are adapted and applied to assess the energy consumption of virtualized resources in data centers. By incorporating insights from industry best practices, the proposed methodology aims to enhance the current approach to energy consumption assessment in virtualized data centers, ensuring a more comprehensive and accurate evaluation of energy usage.

The Initial Proposal addresses challenges in measuring energy consumption in virtualized environments, such as resource sharing, virtualization overhead, and dynamic resource allocation. It incorporates multiple dimensions like CPU utilization, memory usage, disk I/O, and network traffic to provide a holistic assessment. Advanced techniques are proposed to allocate energy consumption in multi-tenant scenarios and differentiate between hypervisor and VM energy usage. The inclusion of the Power Usage Effectiveness (PUE) metric enhances the methodology by considering the overall energy consumption of the data center infrastructure, including cooling systems and networking equipment. Industry standards and best practices from organizations like the Green Grid, SPEC, and the European Code of Conduct for Data Centres are recommended for comprehensive energy efficiency assessment. Insights from Google Cloud's Carbon Footprint Reporting Methodology and the Cloud Carbon Footprint project provide further enhancements, including granular analysis, real-time monitoring, consideration of critical resources, automation, and embodied emissions. By incorporating these improvements, the proposal aims to achieve a more accurate, comprehensive, and sustainable assessment of energy consumption by virtualized resources in data centers.

Next Chapter 6 is dedicated to the validation of the Proposal.

## 6 Validation of the Proposal

This chapter is dedicated to the validation process of the Final Proposal for advancement of assessment methodology for energy consumption by virtualized resources in data centers. Further recommendations for the enhancement of the assessment methodology are outlined at the end of this section.

### 6.1 Overview of the Validation Stage

Data Collection 3 was conducted to seek the evaluation by experts with extensive knowledge in the field, to gain valuable insights and perspectives on the strengths and weaknesses of the Initial Proposal for its potential improvement and the implementation of the thesis outcome. Stakeholders' feedback and expert judgment provided a thorough assessment of the proposal. The Initial Proposal aimed to address the challenges associated with measuring energy consumption by virtualized resources in data centers. It introduced the Power Usage Efficiency (PUE) metric as a key component, which was promptly incorporated into the assessment methodology by the case company. The collaboration with stakeholders, including the Head of Sustainability department and the Service Owner of the Sustainability Hub, played a vital role in shaping the proposal and validating the PUE metric's value. Additionally, the proposed assessment methodology considered multiple dimensions, such as CPU utilization, memory usage, disk I/O, and network traffic, to provide a holistic assessment of energy consumption. The proposed enhancements included advanced techniques for multi-tenant scenarios, differentiation between hypervisor and virtual machine energy usage, and the integration of industry standards and best practices. Insights from Google Cloud's Carbon Footprint Reporting Methodology and the Cloud Carbon Footprint project provided valuable input for granular analysis, real-time monitoring, and consideration of critical resources. The stakeholders expressed their satisfaction with the inclusion of Google's methodology as a best practice reference. This is illustrated in the following quote by the Respondent 4 "I think we can actually really benefit from that." They recognized its potential as a benchmark for the case company's own assessment methodology. It is important to note that while these proposed improvements are promising, they require further evaluation to determine their practicality and impact. Continued collaboration with stakeholders and experts will be crucial in validating and refining these enhancements.

The stakeholders' acceptance of the Initial Proposal provided further validation of the proposed improvements and the potential impact they can have on energy consumption assessment in virtualized environments. Their endorsement serves as an affirmation of the proposal's alignment with the needs and goals of the case company. It is justified in the following statement made by the Respondent 4:

Your thesis has been a good start for us. Also, because we have a lot of standards and documentation that we can go back and look into from this document. I really appreciate what you have done. (*Respondent 4*)

As the initial direction and framework have received validation and approval, a foundation for future research and development has been made, with the proposed improvements offering possibilities for advancing energy consumption assessment in virtualized environments.

## 6.2 Final Proposal

The Final Proposal outlines the key advancements that will enhance the accuracy, granularity, and comprehensiveness of the assessment methodology for energy consumption by virtualized resources in data centers. The findings from the CSA, Existing Knowledge, and Best Practice, as well as stakeholders' feedback to the Initial Proposal serve as the base for the Final Proposal.

Figure 21 below illustrates the summary of the Final Proposal development process.

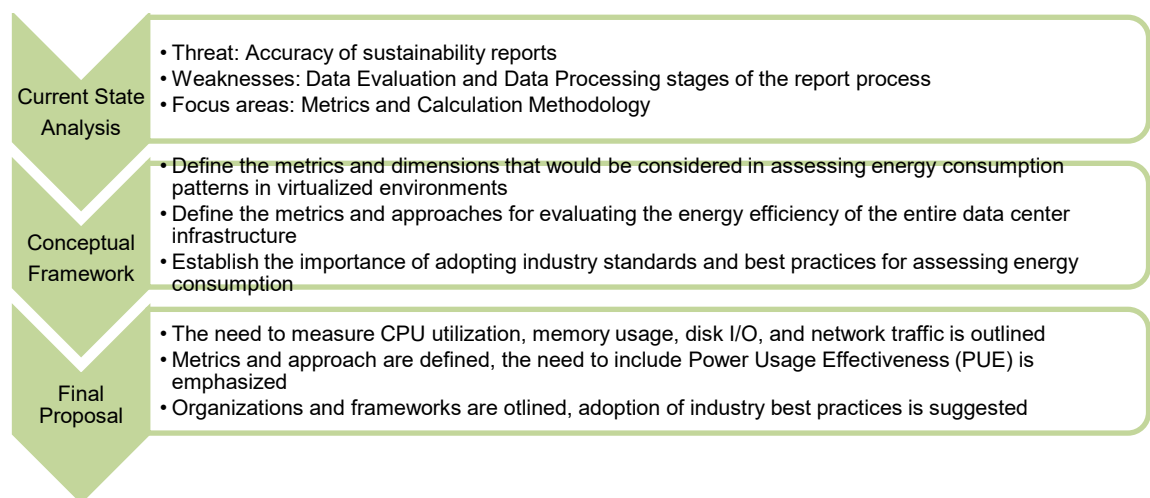


Figure 21. Summary of the Final Proposal development process.

First, the proposed methodology incorporates multiple resource dimensions, including CPU utilization, memory usage, disk I/O, and network traffic, to provide a more holistic assessment of energy consumption patterns in virtualized environments. By measuring the energy consumption of each dimension separately, the methodology gains insights into their specific contributions, enabling granular analysis and optimization of resource allocation. Additionally, advanced techniques and algorithms are suggested to be employed to accurately allocate energy consumption in multi-tenant scenarios and differentiate between hypervisor overhead and virtual machine (VM) energy consumption. It is suggested to consider dynamic resource allocation and VM migrations to provide a realistic assessment of energy consumption.

Second, to capture the holistic picture of energy usage, the proposed methodology should incorporate the Power Usage Effectiveness (PUE) metric into the assessment methodology. PUE, a widely accepted industry standard for measuring data center energy efficiency, will provide insights into the energy consumed by supporting infrastructure components such as cooling systems, power distribution, and networking equipment. By considering PUE, the methodology will account for the energy efficiency of the entire data center infrastructure, leading to a more accurate assessment of energy consumption. The inclusion of PUE will enable data center operators and stakeholders to make informed decisions regarding infrastructure upgrades, cooling optimizations, and equipment replacements, resulting in improved energy efficiency and reduced environmental impact.

Third, the proposed methodology emphasizes the adoption of industry standards and best practices for assessing energy consumption by virtualized resources in data centers. The Green Grid, SPEC, the European Code of Conduct for Data Centres, and cloud service providers' guidelines provide comprehensive frameworks and metrics for measuring and monitoring energy consumption in virtualized environments. By incorporating these industry standards into the assessment methodology, case company can ensure transparency and benchmark its performance against industry-leading practices. To further enhance the methodology, key elements from Google Cloud's Carbon Footprint Reporting Methodology are suggested to be incorporated. This includes conducting a more granular analysis of individual resource dimensions, implementing real-time monitoring, considering critical resources such as disk I/O and network traffic, automating data processing, employing advanced techniques for energy allocation in multi-tenant environments, and differentiating between hypervisor and VM

energy consumption. Additionally, the methodology suggests considering embodied emissions from the manufacturing of hardware and infrastructure used in virtualized resources, providing a more comprehensive assessment of the environmental impact. Integration with existing reporting and analysis systems will ensure seamless incorporation of energy consumption and carbon emission data.

### 6.3 Recommendations

This thesis proposal serves as an initial step towards advancing the assessment methodology for energy consumption by virtualized resources in data centers. It is important to acknowledge that this proposal alone does not provide a comprehensive solution to all the existing challenges. However, it sets the groundwork for future study.

During the research process, the utilization of the Power Usage Effectiveness (PUE) metric was initiated to assess energy efficiency in the case company's data center operations. Recognizing the significance of measuring and monitoring energy consumption, PUE was incorporated as a key performance indicator. Initially, the manual data collection process for PUE values on a quarterly basis was implemented. However, it was acknowledged that this approach posed limitations in terms of real-time tracking and prompt identification of areas for improvement. To address this challenge, it is recommended that the case company to follow the Green Grid's guidelines for the measurement of PUE in the future. These guidelines emphasize real-time monitoring and more frequent updates of PUE values, aligning with industry best practices. By transitioning to automated monthly data collection for PUE, the case company can enhance its energy efficiency assessment methodology, enabling more accurate and up-to-date insights into the energy consumption patterns. It is also important to mention, that while measuring PUE is the initial step in gaining a better understanding of a data center's efficiency, additional investigation is required to determine the most effective approach for further improvement. Extending the assessment scope beyond PUE is proposed to improve data centres' performance evaluation.

Moreover, it is recommended to the case company to stay updated with the advancements in power models, mathematical representations or algorithms that estimate the power consumption of various components and resources within a data center environment. These models are designed to provide insights into the energy

usage patterns and trends of virtualized resources, such as virtual machines, containers, or workloads. Power models consider factors such as CPU utilization, memory usage, disk I/O, and network activity to estimate the power consumed by these resources. As power models evolve, they offer more accurate and fine-grained measurements of energy consumption. Incorporating machine learning and AI techniques into power models can significantly enhance their capabilities. By leveraging historical power consumption data and workload characteristics, machine learning-based power models can predict and estimate real-time energy consumption. These models can identify optimization opportunities and recommend energy-efficient configurations for virtualized resources. Embracing these advanced techniques can lead to more effective energy management practices within data centers and ultimately result in a better understanding of how to accurately measure energy consumption by virtualized resources. By leveraging the capabilities of evolving power models, the case company can gain deeper insights into the energy usage patterns and trends of their virtualized resources. This improved understanding enables more precise measurement and assessment of energy consumption, allowing for better resource allocation and optimization.

The integration of dynamic power management strategies with power models is another recommended approach. Dynamic power management techniques allow for real-time optimization of energy consumption by dynamically adjusting resource allocation, workload placement, and power states. By adapting to changing workloads, data centers can achieve higher energy efficiency without compromising performance objectives. Guided by accurate power models, dynamic power management enables data centers to optimize energy consumption on the fly. The combination of dynamic power management with power models enables a more fine-grained and responsive approach to energy optimization. As the power models provide insights into the energy consumption patterns of virtualized resources, dynamic power management techniques can leverage this information to make intelligent decisions in real-time. This dynamic adjustment of resources and power states ensures that energy consumption is optimized at every moment, resulting in enhanced efficiency and reduced wastage.

By incorporating dynamic power management strategies guided by accurate power models, the case company can enhance their measurement of energy consumption by virtualized resources.

Collaboration with industry consortia, research institutions, and other stakeholders is essential to drive the advancement of power models. By actively participating in collaborative efforts, the case company can share knowledge, exchange best practices, and contribute to the development of standardized power modelling frameworks. Standardization and industry collaboration are crucial for establishing consistent practices and metrics for energy consumption assessment. Standardized frameworks and guidelines for measuring and reporting energy consumption in virtualized environments will enable comparability and benchmarking across different data centers.

It is highly recommended to follow the lead of hyperscalers' such as Amazon Web Services (AWS), Microsoft Azure, and Google Cloud Platform and other cloud service providers. These industry giants have established themselves as pioneers in data center sustainability and energy efficiency. Following the footsteps of hyperscalers' will enable the case to stay aligned with industry trends, drive innovation, and achieve greater energy efficiency in their own data center operations.

In addition to following the practices of hyperscalers, it is also recommended to stay updated with the emergence of open-source applications that facilitate accurate reporting of energy consumption by virtualized resources. As organizations strive for transparency and accountability in their data center operations, it becomes crucial to accurately measure and report energy consumption associated with the virtualized resources purchased by customers. Open-source tools dedicated to this purpose can provide specialized functionalities for monitoring and analyzing energy usage at a granular level, allowing for precise reporting and benchmarking of energy efficiency. By actively exploring and adopting these open-source applications, the case company can enhance its assessment methodology and ensure accurate reporting of energy consumption, promoting transparency and sustainability in the utilization of virtualized resources.

## 7 Conclusion

This section of the thesis provides a concise overview and assessment of the research, starting from the stated objective and progressing through to the final results.

### 7.1 Executive Summary

The case company of this thesis is headquartered in Europe and operates on a global scale, providing IT services, consulting, and cloud services through large-scale data centers. The case company places high demands on sustainability and continuously works towards the reduction of climate impact and aligns sustainability practices with international standards. This thesis addressed the increasingly complex landscape of Environmental, Social, and Governance (ESG) reporting, focusing on the business challenge faced by the case company in complying with the Corporate Sustainability Reporting Directive (CSRD) regulations. The Corporate Sustainability Reporting Directive (CSRD), entered into force in January 2023, mandates comprehensive reporting of ESG metrics by companies. The directive requires high-quality sustainability information, robust data collection, and verifiable reporting, with mandatory external audits. It also includes reporting of greenhouse gas emissions (Scope 1, 2, and 3). The CSRD has both internal and external impacts on the case company, necessitating improvements in data collection, measurement, and reporting systems. Due to the lack of standardized methodologies and established practices for assessment of energy consumption by virtualized resources in data centers, this study was initiated with the objective to propose opportunities for improving the calculation methodology for energy consumption by virtualized resources in data centers.

An action research approach was chosen to provide practical solutions by collaborating with practitioners. Qualitative data was gathered through unstructured and semi-structured interviews with stakeholders in three data collection rounds. Data 1 involved interviews and internal documentation reviews for the Current State Analysis. Data 2 included group interviews to gather suggestions for the proposal. Data 3 focused on validation and approval of the proposed methodology. The current state analysis identified strengths in regulatory compliance, comprehensive reporting of GHG emissions and energy consumption, understanding of customer needs, availability of resources, sustainability competencies, data source understanding, and decision-

making support. Weaknesses included data availability, real-time accessibility, manual processes, and the absence of standardized reporting practices. Opportunities were found in enhancing data dynamics, automating reporting, providing consulting services, and improving access to KPIs through BI dashboards. The selected focus areas for improvement were data evaluation and data processing, specifically refining metrics, and assessment methodology for accurate calculation of energy consumption by virtualized resources in data centers, aligning with industry standards. The topics for the literature review were environmental sustainability frameworks and initiatives, energy consumption metrics for virtualization layer, data center energy efficiency metrics, and industry standards for energy consumption assessment from the cloud. As the outcome, the conceptual framework was built on three key elements: energy consumption metrics for virtualization layer, data center energy efficiency metrics, and industry standards for energy consumption assessment from the cloud.

The Final Proposal represents an advancement in the assessment methodology for energy consumption by virtualized resources in data centers. Based on the findings from the CSA, Existing Knowledge, and Best Practice, as well as valuable feedback from stakeholders, the proposal introduces key improvements to enhance accuracy, granularity, and comprehensiveness. The proposed methodology incorporates multiple dimensions of resource usage, such as CPU utilization, memory usage, disk I/O, and network traffic, enabling a holistic assessment of energy consumption patterns. By measuring the energy consumption of each dimension separately, the methodology provides insights into their specific contributions, facilitating granular analysis and optimization of resource allocation. Furthermore, advanced techniques and algorithms are recommended for accurate energy allocation in multi-tenant scenarios and differentiation between hypervisor overhead and virtual machine (VM) energy consumption. Dynamic resource allocation and VM migrations are also considered to provide a realistic assessment of energy consumption. The proposal also incorporates the Power Usage Effectiveness (PUE) metric to capture overall energy usage and facilitate informed decision-making for infrastructure upgrades and optimizations. By adopting industry standards and best practices, including insights from Google Cloud's Carbon Footprint Reporting Methodology, which suggests granular analysis, real-time monitoring, consideration of critical resources, automation, advanced techniques for energy allocation in multi-tenant environments, differentiation between hypervisor and VM energy consumption, and consideration of embodied emissions from hardware and infrastructure manufacturing.

The thesis was validated through expert evaluations and stakeholder feedback. The proposed improvements, including the incorporation of the Power Usage Efficiency (PUE) metric and advanced assessment techniques, were positively received by stakeholders. Their endorsement validates the proposal's alignment with the needs and goals of the case company. Continued collaboration with stakeholders and experts will be crucial for further validation and refinement of these enhancements. The stakeholders' acceptance provides a solid foundation for future research and development in advancing energy consumption assessment in virtualized environments. The implementation of thesis proposal for advancement of assessment methodology for energy consumption in virtualized environments can enhance the accuracy and reliability of sustainability reports provided to customers and offer more precise and comprehensive information regarding environmental footprint. This, in turn, strengthens the company's credibility, increasing customer trust, demonstrating the company's commitment to transparency and responsible business practices.

## 7.2 Thesis Evaluation

The initial objective of this thesis was to propose opportunities for improving the currently leveraged calculation methodology for energy consumption by virtualized resources in data centers. The focus of the study was made on the metrics which could impact on the assessment of the energy consumption. The outcome of the study is aligned with the primarily goal, as the new metric was incorporated to the methodology immediately, once it was introduced in the Initial Proposal. However, while the incorporation of the new metric is a step forward, it is important to continue reviewing and evaluating other suggested improvements to ensure a comprehensive and robust methodology. This recognition highlights the dynamic nature of the field and the ongoing need for continuous refinement and enhancement. This study serves as a foundation for further exploration and adaptation, as new practices and technologies will emerge.

Limited access to real-world data was the biggest limitation of this research. The practices in this field are still evolving and not yet firmly established. The assessment of energy consumption by virtualized resources in data centers is a complex area, with ongoing advancements in technology, infrastructure, and operational practices. Now, there is a lack of standardized practices and methodologies that are universally accepted and endorsed by regulators or industry bodies.

Additionally, the thesis researcher has come to recognize the importance of obtaining a holistic understanding of the individuals and their respective roles and responsibilities involved in the project before delving into the intricacies of the challenge at hand, which was the case in the first phase of the research and led to uncertainty during the initial interviews. This experience made the thesis researcher recognize the importance of the establishment of a clear overview of the stakeholders' roles and responsibilities, their expertise, and the dynamics of their involvement to effectively address the research objectives.

There were no resource constraints, such as time or human resources, that could have affected the comprehensiveness and depth of the study. On the contrary, the involvement of stakeholders from various business divisions, units, and roles provided valuable insights and expertise that enriched the study. Their involvement allowed for a comprehensive exploration of different dimensions and considerations, ensuring that the research outcomes captured a broad range of insights and considerations.

### 7.3 Closing Words

It is crucial to acknowledge the significance and relevance of the topic addressed in this thesis. The current landscape indicates a growing emphasis on sustainability and environmental considerations, leading to the establishment of more regulations and frameworks in this domain. While regulations exist, the practical implementation and adoption of energy consumption assessment practices remain a challenge for many organizations. However, the case company has taken the first step towards addressing the challenge of energy consumption assessment. By acknowledging the importance of energy consumption assessment and taking proactive measures to address it, the case company has positioned itself as a leader in this evolving landscape.

This complex topic presents tremendous opportunities for exploration and innovation. The assessment of energy consumption in virtualized environments stimulates ongoing research, driving technological advancements, and shaping the establishment of best practices. Being involved in this research effort is a great honor, as it allows me to contribute to the advancement of knowledge in this important field. I am deeply grateful for the opportunity to be a part of this journey to the stakeholders, who supported and guided me throughout the entire process. It is my hope that this thesis can serve as a catalyst for further research and practical implementations in the case company.

## References

- Avelar, V., Azevedo, D., & French, A. (2012). PUE™: A Comprehensive Examination of the Metric. White paper. Available from: [https://datacenters.lbl.gov/sites/default/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric\\_v6.pdf](https://datacenters.lbl.gov/sites/default/files/WP49-PUE%20A%20Comprehensive%20Examination%20of%20the%20Metric_v6.pdf) (Accessed 4 February 2023).
- AuditBoard. (2022). AuditBoard website. Beyond ESG: ISSB Consolidation Heralds a New Era in Corporate Reporting and Assurance. Available from: <https://www.auditboard.com/blog/beyond-esg-issb-consolidation-heralds-a-new-era-in-corporate-reporting-and-assurance/> (Accessed 25 January 2023).
- Belady, C., & Seader, C. (Eds.). (2018). The Green and Virtual Data Center (pp. 116-117). CRC Press. Retrieved from O'Reilly.
- Boston College Center for Corporate Citizenship. (2023). Boston College Center for Corporate Citizenship website. The CSRD: An Overview of Coming ESG Requirements for Companies. Available from: <https://ccc.bc.edu/content/ccc/blog-home/2023/04/what-is-csrd-overview-esg-reporting-requirements-for-companies.html> (Accessed 14 May 2023).
- Cappiello, C., Datre, S., Fugini Paco Melia, M. G., Pernici, B., & Plebani, P. (2013). Monitoring and assessing energy consumption and CO2 emissions in cloud-based systems. Retrieved from ACM Digital Library.
- Cloud Carbon Footprint. (2023). Cloud Carbon Footprint website. Methodology. Available from: <https://www.cloudcarbonfootprint.org/docs/methodology/> (Accessed 30 January 2023).
- Coleman, A. (2007). Leaders as researchers: supporting practitioner enquiry through the NCSL research associate programme. *Journal of Educational Management Administration & Leadership*. Vol. 35, Issue 4, pp. 479-497.
- Computerblog. (2023). Computerblog website. What is a CPU? Available from: <https://dharmendracomputerblog.wordpress.com/2023/05/20/what-is-a-cpu/> (Accessed 10 April 2023).
- Digitaltrends. (2022). Digitaltrends website. What is CPU usage, and how to fix high CPU usage. Available from: <https://www.digitaltrends.com/computing/what-is-cpu-usage-and-how-to-fix-it/> (Accessed 10 April 2023).
- Drax. (2022). Drax website. Why and how to report your scope 1, 2 and 3 emissions. Available from: <https://energy.drax.com/insights/reporting-scope-emissions/#:~:text=There%20are%20no%20legal%20requirements%20to%20report%20on,are%20also%20hard%20for%20companies%20to%20directly%20influence.> (Accessed 15 April 2023).
- Dumitrescu, C., Plesca, A., Dumitrescu, L., Adam, M., Nituca, C., & Dragomic, A. (2018). Assessment of Data Center Energy Efficiency: Methods and Metrics. Retrieved from IEEE.

- Ecochain. (2022). Ecochain website. CSRD - Guide: What to report & How to comply. Available from: <https://ecochain.com/knowledge/complying-with-the-csrd-frequently-asked-questions/> (Accessed 4 February 2023).
- EPA. (2023). United States Environmental Protection Agency website. ISO 14064, International Standard for GHG Emissions Inventories and Verifications. Available from: <https://www3.epa.gov/ttnchie1/conference/ei16/session13/wintergreen.pdf> (Accessed 23 May 2023).
- European Commission. (2022). European Commission website. 2022 Best Practice Guidelines for the EU Code of Conduct on Data Center Energy Efficiency. Available from: <https://e3p.jrc.ec.europa.eu/publications/2022-best-practice-guidelines-eu-code-conduct-data-centre-energy-efficiency> (Accessed 25 May 2023).
- European Commission. (2023). European Commission website. Corporate sustainability reporting. Available from: [https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting\\_en](https://finance.ec.europa.eu/capital-markets-union-and-financial-markets/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en) (Accessed 5 April 2023).
- GitHub. (2023). European GitHub website. Cloud-Carbon-Footprint/Contributors. Available from: <https://github.com/cloud-carbon-footprint/cloud-carbon-footprint/graphs/contributors> (Accessed 20 May 2023).
- Gough, C., Steiner, I., & Saunders, W.A. (2015). Energy Efficient Servers: Blueprints for Data Center Optimization.
- Google Cloud. (2023). Google Cloud website. Carbon Footprint reporting. Available from: <https://cloud.google.com/carbon-footprint/docs/methodology> (Accessed 20 May 2023).
- Google Data Centers. (2023). Google Data Centers website. Efficiency. Available from: <https://www.google.com/about/datacenters/efficiency> (Accessed 25 May 2023).
- Green Business Bureau. (2022). Green Business Bureau website. Scope 1 2 3 Emissions Explained: Understanding the GHG Protocol's Emission Classification System. Available from: <https://greenbusinessbureau.com/green-practices/energy/scope-1-2-3-emissions/> (Accessed 27 January 2023).
- Greenhouse Gas Protocol. (2023). Greenhouse Gas Protocol website. Diagram of Scopes and Emissions Across the Value Chain. Available from: [https://ghgprotocol.org/sites/default/files/ghgp/standards\\_supporting/Diagram%20of%20scopes%20and%20emissions%20across%20the%20value%20chain.pdf](https://ghgprotocol.org/sites/default/files/ghgp/standards_supporting/Diagram%20of%20scopes%20and%20emissions%20across%20the%20value%20chain.pdf) (Accessed 27 January 2023).
- Greenhouse Gas Protocol. (2023). Greenhouse Gas Protocol website. The Greenhouse Gas Protocol Initiative. Available from: [https://ghgprotocol.org/sites/default/files/standards\\_supporting/psp-draft-1.pdf](https://ghgprotocol.org/sites/default/files/standards_supporting/psp-draft-1.pdf) (Accessed 15 January 2023).
- Greenhouse Gas Protocol. (2023). Greenhouse Gas Protocol website. About Us. Available from: <https://ghgprotocol.org/about-us> (Accessed 23 May 2023).

- Gul, B., Khan, I.A., Mustafa, S., Khalid, O., Hussain, S.S., Dancey, D., & Nawaz, R. (2020). CPU and RAM Energy-Based SLA-Aware Workload Consolidation Techniques for Clouds. PDF. Retrieved from IEEE.
- Haghshenas, K., Pahlevan, A., Zapater, M., Mohammadi, S., & Atienza, D. (2018). Magnetic: Multi-Agent Machine Learning-Based Approach for Energy Efficient Dynamic Consolidation in Data Center. PDF. Available from: <https://core.ac.uk/download/pdf/211998492.pdf> (Accessed 14 March 2023).
- Harvard Law School Forum on Corporate Governance. (2023). Harvard Law School Forum on Corporate Governance website. EU Finalizes ESG Reporting Rules with International Impacts. Available from: <https://corpgov.law.harvard.edu/2023/01/30/eu-finalizes-esg-reporting-rules-with-international-impacts/> (Accessed 2 February 2023).
- Hieu, N.T., Di Francesco, M., & Ylä-Jääski, A. (2017). Virtual Machine Consolidation with Multiple Usage Prediction for Energy Efficient Cloud Data Centers. PDF. Retrieved from IEEE.
- IBM. (2023). IBM website. What is a data center? Available from: <https://www.ibm.com/topics/data-centers> (Accessed 10 May 2023).
- IBM. (2023). IBM website. What are IaaS, PaaS and SaaS? Available from: <https://www.ibm.com/topics/iaas-paas-saas> (Accessed 12 May 2023).
- IBM. (2023). IBM website. What is open-source software? Available from: <https://www.ibm.com/topics/open-source> (Accessed 10 May 2023).
- IBM. (2023). IBM website. Multi-tenant vs. single tenant. Available from: <https://www.ibm.com/topics/virtual-machines> (Accessed 23 May 2023).
- Jauregui-alzo, E. (2011). PUE: The Green Grid metric for evaluating the energy efficiency in DC (Data Center). Measurement method using the Power Demand. PDF. Retrieved from IEEE.
- Kaivosoja, J. (2016). Green Data Centers: Evaluating Efficiency. Master's Thesis, Aalto University. Available from: [https://aaltodoc.aalto.fi/bitstream/handle/123456789/23186/master\\_Kaivosoja\\_Juho\\_2016.pdf?isAllowed=y&sequence=1](https://aaltodoc.aalto.fi/bitstream/handle/123456789/23186/master_Kaivosoja_Juho_2016.pdf?isAllowed=y&sequence=1) (Accessed 23 March 2023).
- Keller, E., Szefer, J., Rexford, J., & Lee, R.B. (2010). NoHype: Virtualized Cloud Infrastructure without the Virtualization. PDF. Available from: <https://www.cs.princeton.edu/~jrex/papers/isca10.pdf> (Accessed 10 April 2023).
- Leavy, P. (2014). The Oxford Handbook of Qualitative Research. New York: Oxford University Press. Retrieved from ProQuest Central.
- Lee, D. N., Kumar, R., Nguyen, G.N., and Chatterjee, J. (2018). Cloud Computing and Virtualization. Retrieved from ProQuest Central.
- Lääkkölä, R. (2015). Data Center Degrowth - an Experimental Study. Master's Thesis, Aalto University. Available from: [https://aaltodoc.aalto.fi/bitstream/handle/123456789/16238/master\\_L%C3](https://aaltodoc.aalto.fi/bitstream/handle/123456789/16238/master_L%C3)

[%A4%C3%A4kk%C3%B6l%C3%A4 Riku 2015.pdf?sequence=1](#)  
(Accessed 10 March 2023).

- Makris, T. (2017). Measuring and Analyzing Energy Consumption of the Data Center. Master's thesis, Aalto University. Available from: [https://aaltodoc.aalto.fi/bitstream/handle/123456789/27933/master\\_Makris\\_Theodosios\\_2017.pdf?sequence=1](https://aaltodoc.aalto.fi/bitstream/handle/123456789/27933/master_Makris_Theodosios_2017.pdf?sequence=1) (Accessed 2 April 2023).
- McKinsey & Company. (2022). McKinsey & Company website. What is cloud computing? Available from: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-cloud-computing> (Accessed 17 February 2023).
- McKinsey Digital. (2023). McKinsey Digital website. The green IT revolution: A blueprint for CIOs to combat climate change. Available from: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-green-it-revolution-a-blueprint-for-cios-to-combat-climate-change#/> (Accessed 16 February 2023).
- Microsoft. (2023). Microsoft website. Calculating My Carbon Footprint. Available from: <https://www.microsoft.com/en-us/sustainability/emissions-impact-dashboard?activetab=pivot:mostpopular> (Accessed 2 February 2023).
- Nature Climate Change. (2020). Nature Climate Change website. Hiding greenhouse gas emissions in the cloud. Available from: <https://www.nature.com/articles/s41558-020-0837-6> (Accessed 3 April 2023).
- Net Zero Climate. (2023). Net Zero Climate website. What is Net Zero. Available from: <https://netzeroclimate.org/what-is-net-zero/> (Accessed 23 May 2023).
- Onetrust. (2023). Onetrust website. Ultimate guide to the EU CSRD ESG regulation for business. Available from: <https://www.onetrust.com/blog/ultimate-guide-to-eu-csrd-esg-regulation-for-businesses/> (Accessed 26 April 2023).
- Pitalier, Q. (2018). Cloud services environmental impact evaluation: toward the better knowledge of the benefits and consequences of our usage of the cloud. Master's thesis, Aalto University. Available from: [master Pitalier Quentin 2018.pdf \(aalto.fi\)](#) (Accessed 12 April 2023).
- Portnoy, M. (2016). Virtualization Essentials. John Wiley & Sons, Incorporated. Retrieved from ProQuest Central.
- Povoa, L.V., Bignatto, P.W., Monteiro, C.E., and Mueller, D. (2013). A model for estimating energy consumption based on resources utilization. Retrieved from IEEE.
- PRI. (2022). PRI website. CSRD and ESRS: how EU corporate sustainability reporting is evolving. Available from: <https://www.unpri.org/pri-blog/csr-and-esrs-how-eu-corporate-sustainability-reporting-is-evolving/10539.article> (Accessed 30 January 2023).
- Rădulescu, C.Z., Rădulescu, D.M. and Sipică, A. (2017). An Analysis of the Power Usage Effectiveness metric in Data Centers. Retrieved from IEEE.

- RedHat. (2022). RedHat website. What is a Virtual Machine (VM)? Available from: <https://www.redhat.com/en/topics/virtualization/what-is-a-virtual-machine> (Accessed 23 May 2023).
- Sreejesh, S., Mohapatra, S and Anusree, M R. (2014). Business Research Methods. Switzerland: Springer International Publishing. Available from: <https://link-springer-com.ezproxy.metropolia.fi/book/10.1007/978-3-319-00539-3> (Accessed 15 March 2023).
- Statista. (2023). Statista website. Cloud computing market size in Europe from 2017 to 2030. Available from: <https://www-statista-com.ezproxy.metropolia.fi/statistics/1260032/european-cloud-computing-market-size/?locale=en> (Accessed 12 February 2023).
- Science Based Targets Network. (2023). Science Based Targets Network website. Our mission. Available from: <https://sciencebasedtargetsnetwork.org/our-mission/> (Accessed 23 May 2023).
- Shorten, C. (2022). CSRD in the Data Centre sector – what does it mean for you? Available from: <https://www.linkedin.com/pulse/csrd-data-centre-sector-what-does-mean-you-colm-shorten/> (Accessed 20 April 2023).
- Techradar. (2021). Techradar website. What is RAM and how much do I need? Available from: <https://www.techradar.com/news/what-is-ram> (Accessed 4 March 2023).
- Tech Trainees. (2018). Tech Trainees website. SaaS vendor Workday also finally offers PaaS. Available from: <https://www.techtrainees.com/saas-vendor-workday-also-finally-offers-paas/> (Accessed 2 March 2023).
- Wolters Kluwer. (2023). Wolters Kluwer website. Road to Corporate Sustainability Reporting Directive (CSRD). Available from: <https://www.wolterskluwer.com/en/expert-insights/esg-corporate-sustainability-reporting-directive> (Accessed 15 April 2023).
- World Resources Institute. (2023). World Resources Institute website. Greenhouse Gas Protocol. Available from: <https://www.wri.org/initiatives/greenhouse-gas-protocol> (Accessed 12 March 2023).
- W3schools. (2023). W3schools website. What is GitHub? Available from: [https://www.w3schools.com/whatis/whatis\\_github.asp](https://www.w3schools.com/whatis/whatis_github.asp) (Accessed 21 May 2023).