# UNIVERSITY OF TWENTE.

**MSc. Business Information Technology**

University of Twente

Department of Electrical Engineering, Mathematics and Computer Science

# TOWARDS SUSTAINABLE DATA CENTERS: An Analytical Model Of Circular-Economy

*Authors:*
Andrei-Victor Gorgan (ID: S2252309)

*Supervisors:*
Ana-Lucia Vârbănescu
Devrim Yazan

Date: July 5th, 2024

# Preface

I started my academic path by doing a bachelor's degree in Computer Science, which humbled me in so many ways that it is not even possible to describe them, I learned that while many things on the surface may be simple and trivial, if you look at a fine (or large) enough scale, they become enormous complexities. My bachelor thesis, aimed at the fine-workings between GPU and CPU decompression have shown me the passion I have for the cloud as well as high performance computing.

I then decided not to continue with Computer Science per se but venture into the domain of business where I developed a passion for Enterprise Architecture (EA) and also discovered the complex and intriguing problems that sustainability has to offer. Therefore, for my master thesis I decided to combine my two passions, Enterprise Architecture, Sustainability, and the cloud environment of computer science. And as a personal goal, i wanted to know what can I bring to businesses to help them become more sustainable?

In this 6-month journey that I embarked together with my supervisors, I had a lot of ups and downs, head scratches, and debates; however, I am grateful to both of my supervisors for the time that they have invested in me and my project, the attention to details that I have been scrutinized over and over again. I am thankful to both of them for the endless doors that they have opened for me and I promise both of them that I will continue to be as stubborn as I was until now when defending my ideas. Furthermore, I am thankful to the handful of people who took their time to complete and participate in my interview for this thesis, answers without which I would not be able to gain the insights that I have gained.

I am grateful to Eldir Tommassen and the whole team of ActFact B.V who helped and guided me to understand the intricacies behind data center architecture, the motives for redundancy, practical implication of sustainability and much more. Last but not least, I want to thank my family and friends for supporting me through all the highs and lows that this thesis has put me through.

I strive to believe that my thesis and academic work will help shape the future better.

Andrei-Victor Gorgan
Hengelo, 5th of July, 2024

**Abstract**

In the modern world of scarce resources and fast-paced development, sustainability and efficiency are growing concerns. This is evident in the rapid development and adoption of cloud computing. However, computing power in the context of cloud data centers uses a lot of resources and generates continuous waste such as heat, but also electronic waste.

This study aims to bridge existing gaps in the research on the circular economy in data centers and propose directions for future exploration through a systematic review of the literature (SLR). The primary objective is to develop an updated Total Cost of Ownership (TCO) model that includes the cost of decommissioning hardware. This addition provides a more accurate and holistic view of the financial impacts associated with data center operations, from acquisition and usage to end-of-life management. By integrating return-on-investment (ROI) into the model, the study ensures that sustainability practices do not compromise business profitability, demonstrating that environmentally friendly strategies can align with financial goals.

W designed and performed a comprehensive review of the literature and interviews with industry stakeholders to validate the proposed model. In addition, we examined case studies to evaluate the effectiveness of the model in real-world scenarios. This comprehensive approach ensures that the findings are academically rigorous and practically applicable, offering actionable insights for industry professionals.

The updated TCO model not only addresses financial aspects but also incorporates environmental and social costs, emphasizing the importance of sustainable practices in data center operations. By incorporating these broader considerations, the model provides a robust framework to evaluate and improve sustainability in data centers, promoting a balance between technological advancement and environmental responsibility. The newly designed model serves two purposes: the multitude of components we contribute to a better understanding of the contributing factors of data center sustainability, more specifically electronic waste, while on the other hand, the operational aspect of the model provides field professionals such as data center managers and operators with the tools required to make informed choices over the investments of the data center.

***Keywords—*** Sustainability, Data Center, Circular Economy, Business, Recycling, Total Cost of Ownership, Decommissioning, Hardware Acquisition, Sustainability Practices

# Contents

# Acronyms

**A | C | D | F | G | H | I | K | O | P | R | S | T | U | V | W**

**A**

**AVL** Approved Vendor Lists. 49

**C**

**CPU** Central Processing Unit. 21, 46, 50, 53, 62

**CRM** Critical Raw Materials. 49, 50, 67, 86

**D**

**DLC** Direct Liquid Cooling. 31

**F**

**FOSS** Free and Open-Source Software. 68

**G**

**GPU** Graphical Processing Unit. 50

**H**

**HDD** Hard Disk. 50, 62

**HPC** High Performance Computing. 31, 33

**I**

**IaaS** Infrsatructure as a Service. 36

**K**

**KPI** Key Performance Indicator. 21

**O**

**ODD** Optical Disc Drive. 50

**OEM** Original Equipment Manufacturer. 58

**P**

**PaaS** Platform as a Service. 36

**PSU** Power Supply Unit. 50, 62

**PUE** Power Usage Efficiency. 19, 45, 47, 76, 77

**R**

**RAM** Random Access Memory. 21, 50, 53

**REE** Rare Earth Elements. 50

**ROI** Return On Investment. 12, 24

**S**

**SaaS** Software as a Service. 36

**SDG** Sustainable Development Goals. 53, 54, 84, 87

**SoC** System On a Chip. 53

**SPUE** Server Power Usage Efficiency. 77

**SSD** Solid State Drive. 50

**T**

**TCO** Total Cost of Ownership. 10, 11, 13, 16, 18, 19, 24, 27, 29, 37, 40, 60, 63, 70, 72, 74, 83–87

**U**

**UNEP** United Nations Environment Programme. 56, 57, 59

**V**

**VLAN** Virtual Local Area Network. 78, 81

**W**

**WEEE** Waste of Electrical and Electronic Equipment. 58, 65

# 1 Introduction

Cloud computing has significantly transformed the IT landscape, surpassing the capabilities of the initial Personal Computer (PC) era by leveraging vastly increased raw processing power. Unlike cluster computing, which involves linking multiple servers to work together as a single system, cloud computing delegates resources from on-premises setups to massive data centers. These data centers take on the responsibility of ensuring efficiency in both their operations and those of their customers. This shift requires more complex decision making for data center managers than highlighting the critical importance of optimizing data center operations for sustainability and economic viability. This evolution has facilitated the growth of various online services and platforms, making high-performance computing accessible to a broader audience. However, despite the somewhat whimsical nomenclature, cloud computing infrastructure is firmly grounded in physical data centers located around the world. These data centers are crucial for delivering the services we rely on daily, but they also consume substantial resources, both in terms of electrical energy by generating heat required for cooling down the infrastructure, and materials used in the production of IT equipment, such as server components, cabling, labor, necessitating dedicated research and innovation to address the energy efficiency and computational power required by our modern digitized society.

| Amortized Cost | Component | Sub-Components |
|---|---|---|
| 45% | Servers (Machines) | CPU, memory, storage systems |
| 25% | Infrastructure | Power distribution and cooling |
| 15% | Power draw | Utility costs |
| 15% | Network | Links, transit, equipment |

**Table 1:** Cost breakdown of data centers [30]

The increase in demand for online platforms has been driven by several key factors. Initially, there was a massive increase in digital services due to a general trend towards online activities. This shift was further accelerated by the COVID-19 pandemic, which significantly increased the use of videoconferencing, streaming, online gaming, and e-commerce as people adapted to lockdowns and social distancing measures [52, 68, 38]. The COVID pandemic forced companies to adopt remote work, improve their online presence and improve digital infrastructures, leading to a 20% increase in total internet usage and a rise in the share of e-commerce in global retail trade from 14% in 2019 to 17% in 2020 [74, 68].

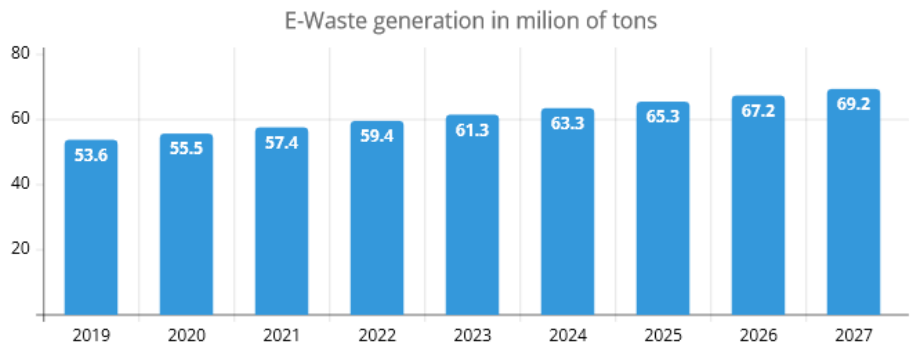



**Figure 1:** E-Waste data

Moreover, the growing popularity of AI applications such as ChatGPT, DALL-E, and Google's Bard (Gemini) has further increased the demand for robust data center infrastructure to support intensive computational requirements [38]. Thus, the increased demand for data centers is driven by the expansion of digital services, pandemic-induced digital acceleration, and the growing AI sector, emphasizing the need for efficient and sustainable infrastructure [52, 68, 74, 38].

For example, in their 2023 yearly developer report, JetBrains indicates that a substantial 48% of respondents now host their applications in a cloud environment [40]. This widespread adoption reflects the cloud's role as a cornerstone of modern IT strategies, offering scalability and flexibility that traditional on-premises solutions cannot match. However, the financial implications of setting up and maintaining data centers are significant, with initial setup costs that can reach millions of dollars and high operational expenses.

The breakdown of data center costs and their respective components, detailed in Table 1, highlights the substantial investment required [30].

Moreover, the environmental impact of this accelerated development, deployment, and operation of data center infrastructure is a growing concern. E-waste, or electronic waste, has been escalating steadily over the years. This increase is driven by the rapid obsolescence of hardware and the constant demand for newer, more powerful technology to keep up with advancing applications. Figure 1 illustrates this troubling trend, showing the persistent rise in electronic waste, which underscores the need for effective recycling and waste management strategies within the industry. This issue of electronic waste is exacerbated by the projected growth in data center demand, making sustainable practices more critical than ever.

To address these environmental concerns, it is essential that data centers adopt a holistic approach when calculating their Total Cost of Ownership (TCO). This approach should include not only the direct financial costs, but also the environmental costs associated with electronic waste, embodied carbon, and operational energy consumption. Incorporating these factors into the design and scaling strategies for data centers, can help promote more sustainable industry practices. For example, effective management of electronic waste can involve measures such as extending the life cycle of hardware, investing in energy efficient technologies, and implementing comprehensive recycling programs [61].

Furthermore, sustainable data centers should explore renewable sources for their operational energy requirements, improve cooling efficiency to reduce energy consumption, and design modular and scalable data center architectures that can adapt to future technological advancements with minimal environmental impact. The future of cloud computing and data center infrastructure is intrinsically linked to sustainable practices. By acknowledging and addressing the environmental and financial challenges associated with data centers, the IT industry can ensure that it continues to innovate and grow while minimizing its ecological footprint. This balance between technological advancement and environmental responsibility is crucial for the long-term viability and success of cloud computing and its associated industries.

## 1.1 Research Questions

This research not only addresses existing deficiencies, but also proposes actionable solutions to promote a more sustainable and efficient future for data center operations. We aim to identify any existing gaps in the research on the circular economy in data centers and propose solutions for research through a Systematic Literature Review (SLR). The goal is to provide a comprehensive understanding of the current landscape, pinpointing where further research and innovation are needed to improve sustainability practices in data centers.

To this end, we pose the following research question:

***"What Total Cost of Ownership model can be employed to evaluate sustainable growth strategies within data center environments?"***

In response to this question, our aim is to define an updated Total Cost of Ownership (TCO) model that includes the cost of decommissioning hardware in data centers. This consideration is critical because it addresses the full lifecycle of data center hardware, from acquisition through operation to end-of-life management. By incorporating decommissioning costs, the updated TCO model offers a more accurate and comprehensive understanding of the financial impacts associated with data center operations. This holistic approach ensures that all phases of hardware management are accounted for, promoting more sustainable and economically viable practices.

Furthermore, the Return On Investment (ROI) of a data center is taken into account to ensure that sustainability practices do not adversely affect the profitability of the business. Balancing environmental sustainability with economic performance is essential, as businesses need to remain financially viable while adopting greener practices. By evaluating ROI alongside sustainability efforts, this research aims to demonstrate that environmentally friendly practices can be integrated into business strategies without compromising profitability. This alignment of sustainability and profitability is crucial for the long-term success and acceptance of green practices in the industry.

To achieve our research goal, we formulate four research questions that provide the needed milestones towards an answer to our main question.

RQ1: What are the key challenges and benefits of implementing circular economy principles in data center operations?

RQ2: How do maintenance costs, including service costs, software licensing, and repair costs, contribute to the overall TCO?

RQ3: What are good practices for hardware acquisition that align with sustainability goals?

RQ4: How do different sustainability practices (reduce, reuse, recycle) affect the operational costs and efficiency of data centers?

By answering these questions, our study advances the understanding of sustainability in data center operations. A systematic review of the literature provides a solid foundation for this work: By evaluating current research, identifying trends, and highlighting gaps, we identify promising new directions for research (RQ1). We then focus on evaluating feasible circular economy practices and the development of an inclusive TCO model (RQ2, RQ3), considering ROI to demonstrate the environmental and economic viability of sustainable practices (RQ4). This comprehensive approach ensures that the findings are grounded in existing knowledge, while pushing the boundaries of current understanding, and offering both theoretical assessments and practical solutions for industry practitioners to promote a more sustainable future for data centers.

## 1.2 Methodology

Our research proposes an improved analytical model to quantify the cost and possibilities of data centers with respect to the decommissioning of old hardware. To answer *RQ1, RQ2, and RQ3*, we start with an extensive literature review, collecting sustainable practices from various ICT domains, and pursue their evaluation in the context of sustainable data centers. As such, our study follows an iterative design process, as shown in Figure 2.

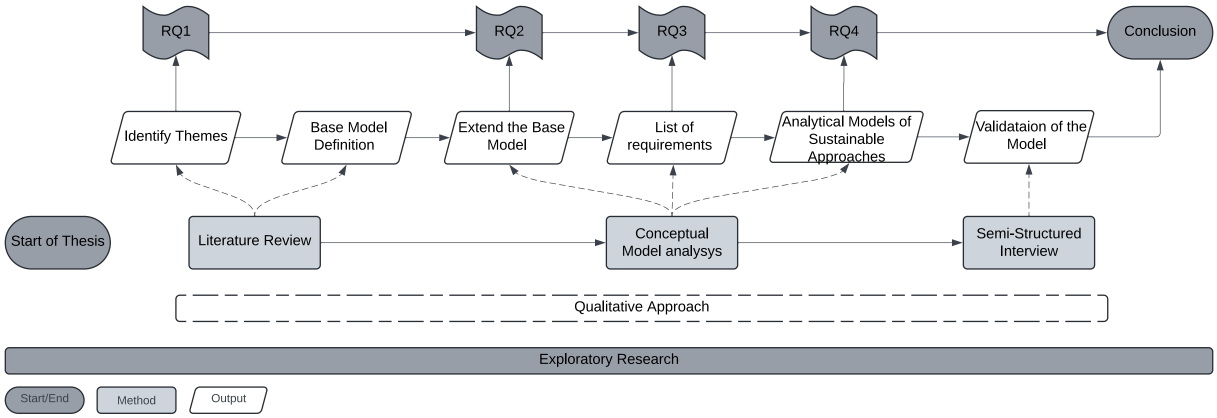



**Figure 2:** Iterative Research Design

Our answer to ***RQ1***starts with the TCO model by Hardy et al. [34] as a base, and extends it by considering the costs generated by the management of electronic waste through both the generation and recycling of electronic waste.

To answer ***RQ2***, we start from the EcoUp scaling strategy, defined by Yan et al. in [75]. We will consider an additional step in the calculation of the portfolio and server cost, taking into account the cost of decommissioning.

To answer ***RQ3 and RQ4***, we carry out interviews with the stakeholders involved in order to validate the proposed model. We also investigate several case studies to further validate the performance and usability of the model.

To ensure the successful completion of this research, we developed and followed the plan illustrated by Table 2.

## 1.3 Thesis Structure

The introduction of this report establishes the critical importance of incorporating sustainability and circular economy principles into data center operations, emphasizing the necessity for a comprehensive Total Cost of Ownership (TCO) model that encompasses not only financial aspects but also environmental and social costs. We present a review of existing literature on TCO models in the second section, focusing on the acquisition and expansion of server machines and examining sustainability considerations relevant to server machines in data centers, particularly during the decommissioning process.

The research then identifies existing gaps in the areas of data center sustainability, TCO models, and decommissioning processes, highlighting the need for further investigation to develop more holistic and sustainable practices. Guiding the study are primary

| Phase | Action | Description |
| --- | --- | --- |
| Interviews | Create Interviews | Develop interview questions to collect information from stakeholders involved in the operations of the data center. |
| | Contact Companies | Identify and reach out to companies that are willing to participate in the interviews. |
| | Conduct Interviews | Execute the interview process, collecting valuable data from various stakeholders. |
| Model Development | Create Initial Model | Construct the initial version of the Total Cost of Ownership (TCO) model, integrating aspects of sustainability and circular economy. |
| | Improve Model | Refine the model based on preliminary findings and feedback from initial interviews. |
| | Analyze Interviews | Analyze the interview data to extract meaningful insights that will inform further model improvements. |
| Case Studies and Validation | Case Study Validation | Apply the model to real-world scenarios through detailed case studies to validate its effectiveness. |
| | Validate Model | Confirm the robustness and reliability of the model by evaluating its performance in different contexts and case studies. |

**Table 2:** Phases of the Research

research questions that form the foundation for investigating relevant literature and developing approaches to address these queries. The design of interview questions is described in section 3, emphasizing assessments of individual components, hardware acquisition, decommissioning costs, and sustainability strategies. Further, the literature review examines the integration of sustainability and circular economy principles into data center architecture, critiquing the traditional TCO model for its failure to address environmental impacts and introducing new cost category such as *”Circular Economy Initiatives”*.

Section 4 presents a summary of the modifications done to the existing TCO categories, incorporating additional costs related to services, software licensing, and repairs to provide a more comprehensive understanding of the financial implications associated with data center operations. Best practices for hardware acquisition aligned with sustainability objectives are provided, emphasizing the importance of thorough requirement elicitation to understand hardware needs and their connection to sustainability objectives. The sustainability impacts on data center operations are discussed, highlighting the benefits of adopting circular economy principles and introducing strategies for reducing environmental impact through design, reuse, and recycling. Lastly, section 8 proposes enhancements to the existing TCO model , introducing new additions that focus on a comprehensive approach to costs, including the integration of sustainability metrics.

The application of the developed TCO model in real-world scenarios is discussed, covering the validation process for the new model and evaluating its effectiveness through various case studies.

We conclude in section 10 summarizes the main findings of the study, highlighting key insights gained from the investigation, discussing its limitations, and suggesting areas for future investigation to further enhance data center sustainability.

# 2 Literature Review

This section discusses the literature study on each of the following models: TCO, acquisition and expansion of server machines, and finally the sustainability considerations of server machines in data centers, especially when considering the decommissioning process.

## 2.1 Literature Review Method

A literature review assists in the development of theories using existing research and identifying areas where further study is necessary [73]. A concept-centric approach is chosen because, according to Webster and Watson, an author-centric method does not efficiently synthesize the literature. As a result, this review explores the current literature and theories related to data center sustainability, aiming to highlight key concepts and ultimately propose a conceptual framework that integrates and expands existing research [73]. The review addresses the research questions RQ1 and RQ3.

A comprehensive literature review should include all relevant scientific studies on the topic, without limiting the scope to specific research methodologies, journals, or geographic regions [73]. To identify all significant contributions, a systematic three-step process is employed. First, a primary query for the Scopus database is tested and refined. The main goal is to capture the core body of literature on data center sustainability and total cost of ownership. The query below was determined to provide the most thorough results, yielding a total of 1067 publications:

> TITLE-ABS-KEY((”data center” OR ”datacenter” OR ”data-center”) AND ((”sustainability” OR ”decommissioning”) OR (”Total Cost Of Ownership” OR ”TCO”)))

The publications retrieved from the query were initially evaluated based on their titles and abstracts. Articles that did not meet the research objectives were excluded, followed by a thorough full-text review to determine which articles would be included in the study. Secondly, as recommended by Webster and Watson, forward and backward search (snowballing) was used to identify additional publications, improving the review. Moreover, due to the interdisciplinary nature of data centers and sustainability, contributions come from various fields such as industry, finance, computer science, and electrical engineering, all contributing to the gray literature.

Thirdly, to ensure no relevant publications related to the broad domain of Total Cost of Ownership (TCO) were missed, a more general query was created. Out of a total of 734 results, the query was refined using strict exclusion criteria. To maintain the relevance of this research, papers published before ***2012*** were excluded. Additionally,

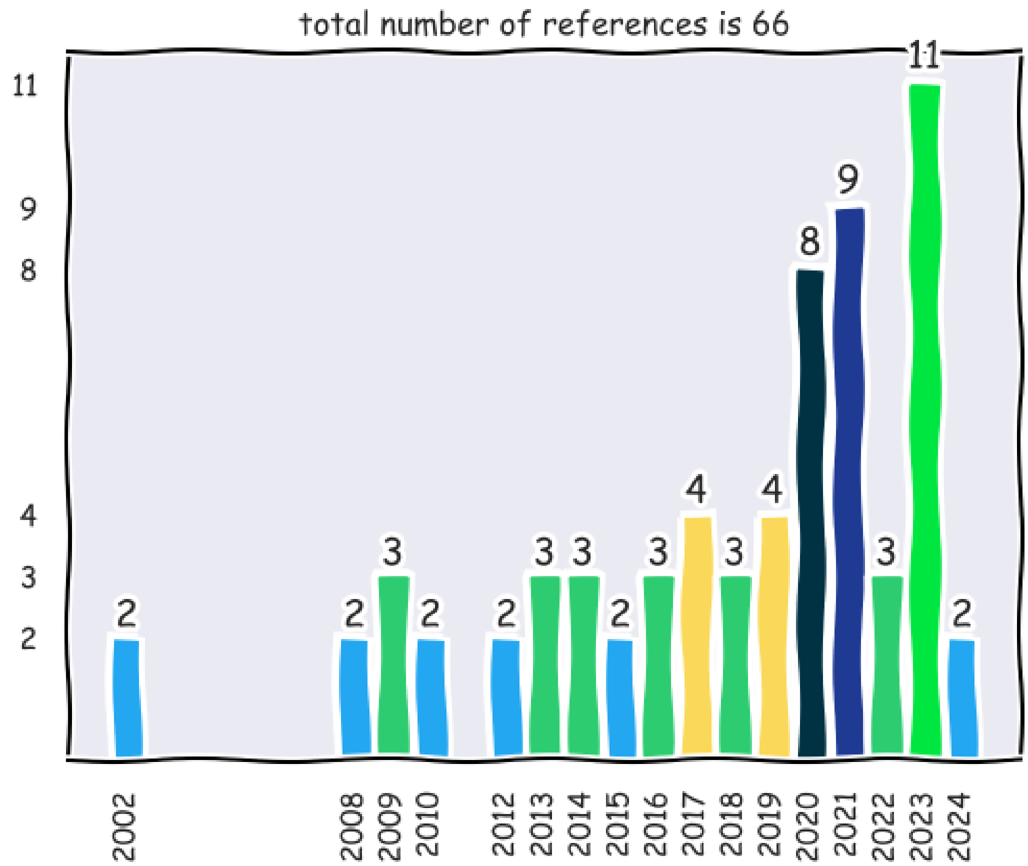



**Figure 3:** Distribution of chosen papers and year of publication

only publications in English, including articles, book chapters, reviews, and conference proceedings, were considered. Finally, literature from unrelated disciplines such as social science, physics, and astronomy was excluded, resulting in a total of 414 results. The query is provided below. Literature retrieved from this query is also included in the results of snowballing.

> TITLE-ABS-KEY((”data center” OR ”datacenter” OR ”data-center”) AND ((”sustainability” OR ”decommissioning”) OR (”Total Cost Of Ownership” OR ”TCO”)) AND NOT (cool* OR power*) AND cost*)

Furthermore, Figure 3 displays the distribution of the chosen articles according to the year of publication.

## 2.2 Total Cost of Ownership

The Total Cost of Ownership (TCO) is a financial estimate that includes all direct and indirect costs associated with the acquisition, use, maintenance, and disposal of a product or asset throughout its life cycle. It goes beyond the initial purchase price and includes various expenses that occur throughout the ownership period. TCO typically includes:

| | |
|---|---|
| **Initial Cost** | Purchase price, installation expenses, and any initial setup costs. |
| **Operating Costs** | Costs incurred during the use of the product or asset, such as energy consumption, maintenance, repairs, and consumables. |
| **Support Costs** | Expenses related to customer support, warranties, training, and any additional services required. |
| **Downtime Costs** | Costs associated with any disruptions or downtime in operations, including potential losses in productivity. |
| **Disposal Costs** | Expenses related to decommissioning, disposal, or recycling at the end of the product's life. |

**Table 3:** Total Cost of Ownership types

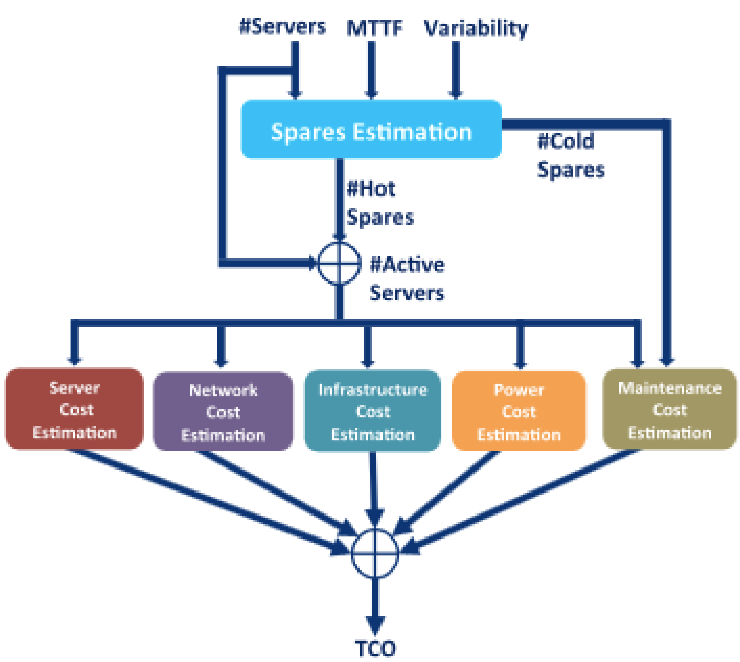



**Figure 4:** Total Cost of Ownership model [34]

The concept of TCO delineates a comprehensive approach to grasp all relevant supply

chain expenses associated with the contracting of a specific supplier for a particular product or service. The TCO covers the general expenses incurred during the acquisition, utilization, management, maintenance, and disposal phases of the item or service in question [51]. It should be noted that TCO does not require exact calculations for every cost element but focuses on key cost considerations and factors that could influence the decision-making process, as elaborated later on [19].
Businesses often use the TCO analysis to make informed decisions about investments by considering not only the upfront cost but the general expenses associated with owning and operating a particular asset or product.

Hardy et al. proposed a five-step TCO model overview of the data center [34]. The model, created in 2013, describes the total cost of ownership of a data center. A unique consideration of this model is the differentiation between hot and cold server spares. Although hot spares contribute to the total number of active servers, they need to be included in the total cost estimate for servers, network, power, and maintenance. Meanwhile, the cold spares are not active and only contribute to the cost of maintenance. Figure 4 illustrates this difference. Furthermore, the model considers the cost of maintaining and operating a data center per month. An overview of the main equation is shown below:

$$TCO = C_{infrastructure} + C_{server} + C_{network} + C_{power} + C_{maintenance} \tag{1}$$

### 2.2.1 Infrastructure Cost

The model provides the cost of infrastructure as:

$$C_{infrastructure} = \frac{C_{building} + C_{cooling}}{D_{datacenter} * 12} \tag{2}$$

Where $C_{building}$ represents the total cost of acquiring land and building the warehouse, which takes into account the total area required for the server racks and the area required for the cooling equipment. $C_{cooling}$ consists of the cost of equipment required for the cooling of servers and other hardware within the data center. $D_{datacenter}$ represents the expected depreciation of the entire data center building in years. However, [13], builds on the initial paper and introduces a new aspect, respectively, Power Usage Efficiency (PUE). As a result, the cost of cooling depends on the efficiency of the servers present in the data center. An updated cooling cost is represented in the equation 3 where $C_{cooling/watt}$ is the cost of cooling a watt of energy and $P_{server/networking}$ is the power consumed by servers and networking equipment, respectively.

$$C_{cooling} = (SPUE \times P_{server} + P_{networking}) \times C_{cooling/watt} \tag{3}$$

### 2.2.2 Server Cost

To begin with, most servers include a number of spares, servers that do not run, but are available to be quickly spun up in case an emergency takes place. The number of hot spares is determined based on the variability factor (VF), which means that when all available servers are required, VF = 0.

$$N_{hotspare} = \frac{N_{servermodulesreq}}{1 - VF} - N_{servermodulesreq} \tag{4}$$

Cold spares exist, we don't think of how you estimate of them, but they form the $C_{coldspares}$. Therefore, the cost of a server is the product of the number of servers required and the cost of a server module. This is then multiplied by the depreciation of the server in years and multiplied by 12 to account for monthly costs.

$$C_{svrtype} = \frac{N_{srvmodules} \times C_{srvmodule}}{D_{srv} \times 12} \tag{5}$$

However, considering that multiple configurations can exist within one data center, we consider the cost of k configurations and the fact that servers are usually not paid in full but are credited and acquire interest ($\alpha$%) over time.

$$C_{server} = \sum_{i=0}^{k} C_{srvtype,i} \times (1 + \alpha_i\%) \tag{6}$$

### 2.2.3 Networking Cost

The networking cost is composed of an overview of the hardware equipment required by each rack individually, assuming a number of switches and cables required per rack. If the racks are not homogeneous, the cost is the sum of all different variations of the rack. Furthermore, the cost contains a depreciation value, considering that the equipment will depreciate in a few years.

$$C_{network} = \sum_{i=1}^{k} \frac{N_{i\_racks} \times C_{i\_equipemnt}}{D_{i\_equipement} \times 12} \tag{7}$$

Equation 7 shows the cost calculation for a nonhomogeneous data center, where $i$ represents a different rack configuration, with up to $k$ possible configurations.

### 2.2.4 Maintenance Cost

To ensure the well-working of the servers, monthly maintenance is required. Maintenance includes the cost of replacement components, cold spares, and the salary costs of personnel. Given that data centers are equipped with cold spares, we take this as the number of servers that need to be maintained. $D_{srv}$ is the depreciation value of a server over a year.

$$C_{maintenance} = \frac{N_{coldspares} \times C_{srvmodules}}{D_{srv} \times 12} + (N_{racks} \times C_{salaryperrack}) \tag{8}$$

## 2.3 Load-dependent upgrading of servers

Whilst Total Cost of Ownership is an important Key Performance Indicator (KPI) for a data center, it is also of the utmost importance that hardware re-inspection is done according to the needs of the data center and those of end customers [48, 66]. The two most commonly used strategies in data centers are scaling up and scaling out or some combination of the two; however, recent research shows that improvements can be made to this [75, 14, 58].

### 2.3.1 Scaling Up

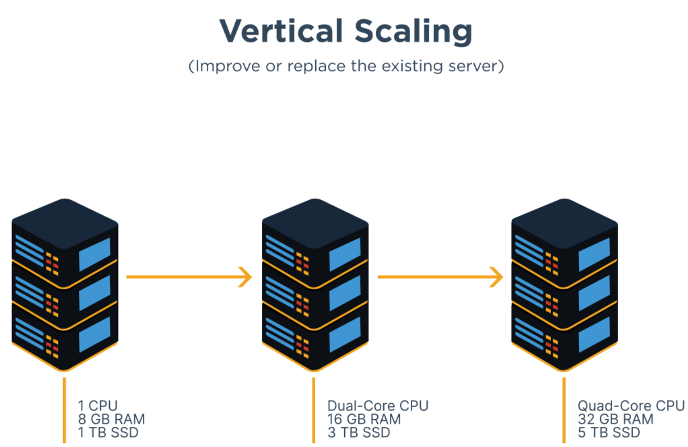



**Figure 5:** Example model of vertical scaling in a data center [44]

Scaling up, also known as vertical scaling, is an architectural strategy in which you increase the capacity and performance of an individual machine by adding more resources to it [50, 49]. Instead of distributing the workload across multiple smaller machines, you enhance the capabilities of a single server.

In a scale-up model, you typically add more powerful components to a single server. This could include upgrading the Central Processing Unit (CPU), adding more RAM, or expanding storage capacity. It is particularly effective when dealing with tasks that benefit from a single powerful processing unit, such as complex computations or database transactions that may not easily parallelize [64].

Managing a single larger server can be simpler than dealing with multiple nodes in a distributed system. There is less complexity in terms of coordination and communication between different machines. Scaling up is often easier in the short term because it involves upgrading or adding components to an existing system. This can be a more straightforward process than designing and implementing a distributed system for scale-out [65, 53, 33].

Although scaling up can provide a quick boost in performance, there is a limit to how much you can scale a single machine. Eventually, you may reach the maximum capacity of the hardware, and further scaling may become impractical or cost-prohibitive.

### 2.3.2 Scaling Out

In IT, a scale-out architecture model, also known as horizontal scaling, refers to an approach where the system's capacity and performance are increased by adding more hardware or nodes to the existing infrastructure.

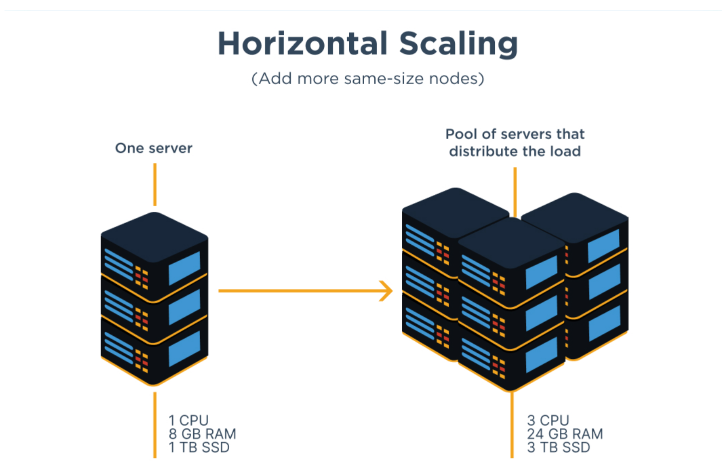



**Figure 6:** Example of horizontal scaling in a data center [44]

In a scale-out model, the system's capacity is increased by adding more nodes or machines to the network. Each new node contributes to the overall capacity of the system, distributing the workload among multiple units [26, 31]. To ensure optimal utilization of resources, a load balancing mechanism is typically used. The system can easily adapt to changing workloads by adding or removing nodes as needed. This provides flexibility and cost-effectiveness, since resources can be scaled up during periods of high demand and scaled down during periods of low demand [33, 31].

An advantage of a scale-out model is increased resilience to hardware failures. If one node fails, the others can continue to operate, minimizing downtime [53]. Furthermore, scale-out architectures often use standard, off-the-shelf hardware components, making it easier and more cost-effective to acquire and replace components. Often, it is also considered a cost-effective scaling method: Instead of investing in a single powerful machine, organizations can start with a smaller infrastructure and gradually add more nodes as the workload grows [71].

Common examples of scale-out architectures include distributed file systems, web server clusters, and cloud computing environments. In these systems, the addition of new nodes allows the infrastructure to handle increased workloads and provide improved performance.

### 2.3.3 The alternative

Whilst the architectural strategies of scaling up and/or out have been proven to work in a plethora of applications, they fail to excel in data centers. In the case of data centers, one machine usually serves a specific application, and applications are not run on all configurations of servers, either because it is not needed or because of architectural constraints. This forms the basis of EcoUp, a framework developed to help data centers upgrade machines according to application needs [75].

To achieve this, the framework (presented in Figure 7) displays 4 steps: *Predicting Performance*, *Weighing Applications*, *Estimating Cost* and *Rating Cost Efficiency*. All of which are presented in the following paragraphs.

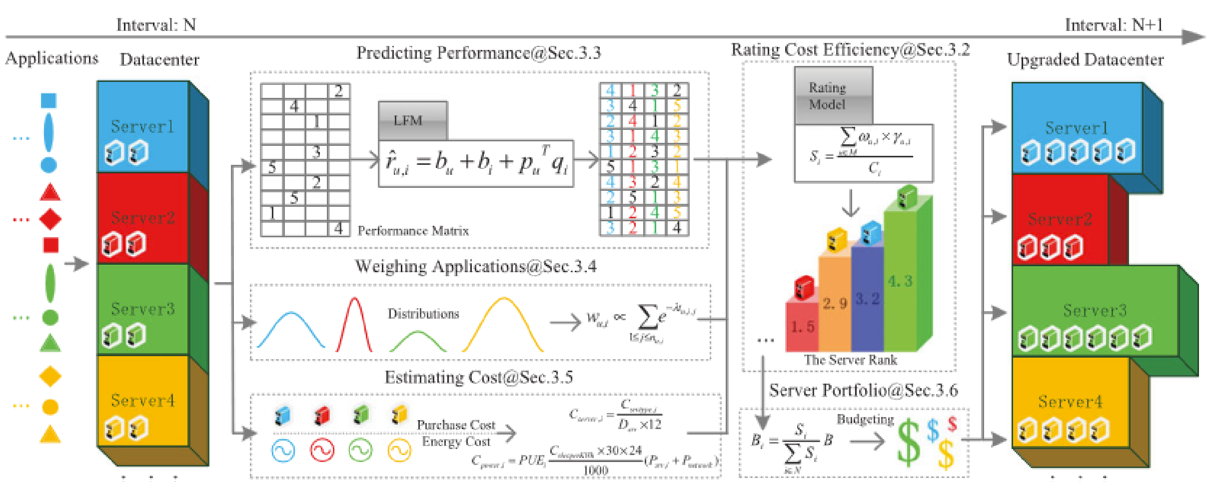



**Figure 7:** Framework for choosing the servers to use [75]

**Predicting Application usage**

$$\hat{r}_{u,i} = b_u + b_i + p_u + {}^T q_i \tag{9}$$

In order to perform an analysis, data is required on the applications. Realistically, a data center cannot run one application on all available server configurations. However, from the data recorded, a matrix $R_{U\times I}$ of the configurations of the I server and the U applications can be created and populated using equation 9. The prediction function considers the application characteristics $p_u$ multiplied by each characteristic of the server $q_i$. Moreover, due to some servers running multiple applications and or other biases in server configurations, as well as application biases such as more memory, the additions of $b_u$ and $b_i$ are made. All of these parameters can be computed from the historical logs of the servers.

Another aspect to consider is that of run-time. It is often the case that not all applications run at the same time and for the same duration. For example, an application such as a database may have short bursts of high-power usage and, for the rest of the time, stay quiet. An Exponential Decay of time can then be used to weigh applications.

$$w_{u,i} \propto \sum_{1\leq j\leq n_{u,i}} e^{\lambda t_{u,i,j}} \tag{10}$$

Where $n_{u,i}$ is the total number of application U on server I within the specified time window $t_{u,i}$. Therefore, the more an application is used, the more weight it has; meanwhile, the higher the time window, the less impact the application has on the performance of the server.

**Rating Cost Efficiency**

$$S_i = \frac{\sum_{u\in M} w_{u,i} \times r_{u,i}}{C_i} \tag{11}$$

Combining the cost of running a server ($C_i$) with the performance index of the said server, a cost efficiency score can be computed. Using the score, one can better rank server configurations for the applications used within the data center.

## 2.4 Revenue of old hardware

Every commerce technology requires an investment, often referred to as Total Cost of Ownership. Return On Investment (ROI) is the other half of the equation: It is the result you get, the return on investment you made. With legacy platforms, Total Cost of Ownership can spiral out of control, and ROI is increasingly difficult to achieve. Therefore, servers should only be upgraded once either Total Cost of Ownership cannot be minimized or the hardware cannot sustain an increase in ROI.

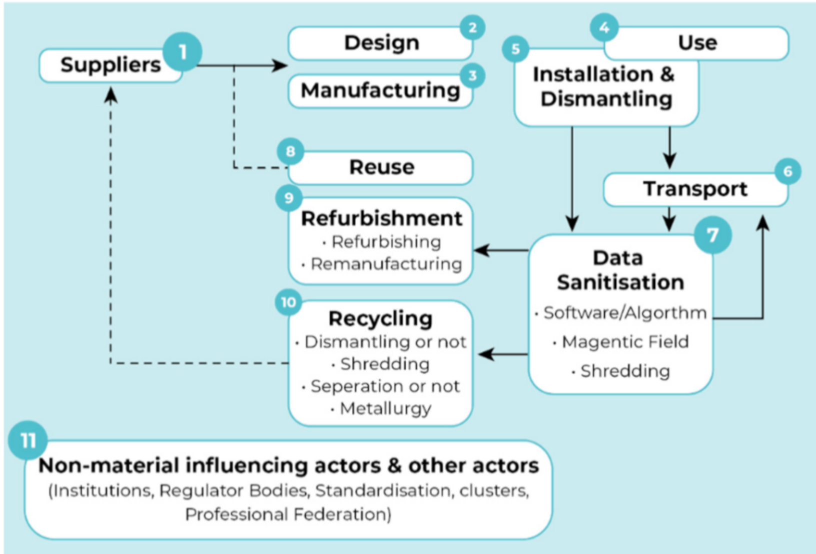



**Figure 8:** Steps to reduce E-Waste in servers [6]

Servers typically have a short refresh cycle, which is on average 3-5 years but sometimes as little as one year, due to various factors, including maintenance and lease contracts [7]. The recovery of value and materials present in waste electrical and electronic equipment (WEEE) is critical to preventing environmental damage, preventing depletion of resources, and utilizing energy efficiently [36]. When dealing with the disposal of data center servers, four key scenarios can be distinguished. It is preferable to dismantle servers rather than shred them, as this allows the sorting of materials before sending them to shredding facilities. In the realm of reuse, the central consideration revolves around whether servers are reused within the region or exported for reuse. Exporting for reuse can result in valuable materials leaving the region, with no guarantee of recovery at the end of their subsequent lifecycle. Regional reuse, while more environmentally friendly, poses challenges due to the free trade of servers. It is common for multiple brokers to have had ownership of a server before it reaches the final end-user, making it impractical to maintain control over the materials. [36, 7]

Regarding material breakdown, servers contain a high proportion of steel, aluminum, and plastic; three of the main materials for industrial greenhouse gas emissions worldwide. Even though current legislation imposes proper disposal and recycling of e-waste, some materials are non-recyclable, and recyclable ones will not be fully recoverable. Table 4 shows the number of recoverable materials following End-of-Life (EoL) scenarios 2 and 3, where scenario 2 represents servers that are recycled after some parts are manually separated and treated (such as batteries, hard drives, etc.) and scenario 3 represents servers that are recycled without any previous treatment [7, 57].

| Component / Material | Amount in server (g) | Recycling EoL 2(g) | Recycling EoL 3 (g) |
|---|---|---|---|
| Aluminium | 1,263 | 1,185.1 | 1,149.07 |
| Brass | 7 | 6.65 | 4.9 |
| Copper | 806.56 | 747.98 | 483.72 |
| Steel | 14,861 | 13,996.18 | 13,970.21 |
| Other metals | 216 | 151.11 | 151.11 |
| Zinc | 96 | 67.2 | 57.6 |
| ABS | 360 | 266 | 266.4 |
| EVA | 75 | 0 | 0 |
| HDPE | 210 | 27.76 | 0 |
| PBT | 240 | 0 | 0 |
| PC | 289 | 0 | 0 |
| PCABS | 324.28 | 0 | 0 |
| PCFR40 | 51 | 0 | 0 |
| PCGF | 52 | 0 | 0 |
| PUR | 2 | 0 | 0 |
| PVC | 145 | 0 | 0 |
| Styrofoam | 1,026 | 0 | 0 |
| Rubber | 35 | 0 | 0 |
| Other materials | | | |
| Cables | 31 | 7.4 | 7.4 |
| Electronics | 3,966 | 596.12 | 444.32 |
| Paper | 3,629 | 0 | 0 |
| Neodymium magnets | 68 | 0 | 0 |
| Batteries | 44.6 | 20 | 20 |
| Total | 27,797 | 17,142 | 16,555 |

**Table 4:** Material decomposition of a server

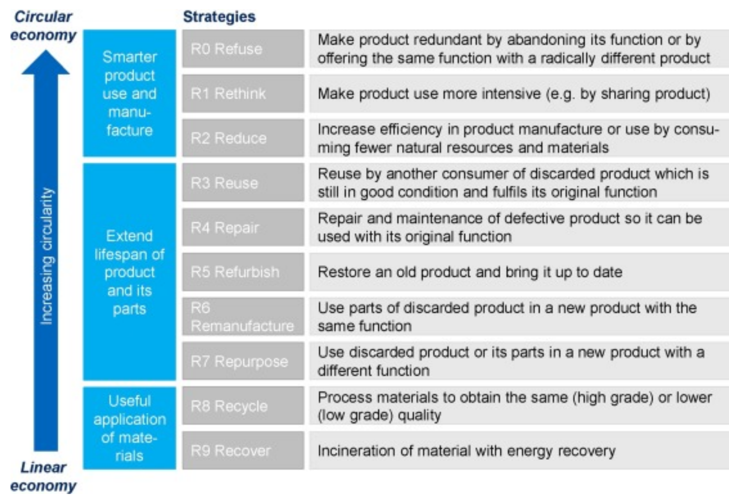



**Figure 9:** The 10-R framework [47]

Furthermore, the concept of circular strategy can play a role in the development of data centers. The fundamental principle of a circular economy involves expanding the obligations of manufacturers and the responsibility of end users. Within this framework, resources are retained within the system for extended periods, maximizing their value. Eventually, at the end of their life cycle, components are reclaimed for reuse or recycling. Figure 9 displays the Kirchherr model of different circularity movements, as well as classifying the extent to which materials are recovered from the grave [47].

## 2.5 Research Gaps

Whilst the work of Hardy et al. provides a good starting point for modeling the Total Cost of Ownership of a data center [34], it does not touch on aspects such as the effect of different loads on the data center. Furthermore, the environmental impact is considered to be crude, using an *Equivalent Performance Coefficient* rather then an established metric such as TFLOPS.

Meanwhile, there are strategies to upgrade and scale the data center, such as scaling horizontally or vertically. Although EcoUp tries to close the gap towards a more sustainable data center by only replacing required hardware for the use cases of the data center, it does not touch on the costs of server decommission [75].

Lastly, there is research in the area of server revenue; however, it has never been tied to the Total Cost of Ownership (TCO) of the server [58]. Circular business practices have been linked to data centers, especially in terms of energy efficiency, however, the decommissioning of server and the cost associated with e-waste are not included in existing TCO frameworks.

# 3 Interview Method

The section on the interview method describes the utilization of interviews to collect detailed experiences and strategies from data center professionals regarding server management, with a focus on upgrading, decommissioning, and refreshing machines. The interview framework consisted of 32 questions categorized into four sections: Individual Assessment, Hardware Acquisition Costs, Hardware Decommissioning Costs, and Circular Sustainability Strategies, addressing roles, hardware purchasing behaviors, decommissioning methods, and sustainability practices. Data were gathered using Microsoft Forms, with optional follow-up interviews. Participants, chosen from known industry contacts and LinkedIn, were based in Europe and held positions such as Data Center Manager, Director, and Supervisor. The results showed diverse opinions on the significance of differentiating between hot and cold spare servers, with smaller businesses often not distinguishing due to cost concerns. Power efficiency was not a priority for some, but all participants agreed on the essential importance of cooling. Budget constraints influenced hardware purchasing decisions for most participants, who prioritized performance, reliability, and compatibility. Sustainable and certified processes were employed for hardware disposal, with a preference for standardized, replaceable components. The study highlighted strategic hardware lifecycle management, proper planning, capacity alignment, and reliability, emphasizing the importance of reuse, repair, refurbishment, and recycling as key sustainability practices.

## 3.1 Interview design

Interviews were chosen as the main method of data collection due to their ability to capture nuanced experiences, emotions, and challenges faced by practitioners. Given that there is no solution to fit all sizes of business and requirements, capturing the experience and thoughts of current practitioners is of great importance to design a model capable of overcoming current issues. Therefore, the primary objective of the interviews was to determine daily experiences and current strategies with regard to upgrading, decommissioning, and refreshing server machines in the context of data center usage.

### 3.1.1 Question Design

The interview consists of 32 questions that are divided into 4 main sections: *Individual assessment, Hardware Acquisition Costs, Hardware Decommissioning Costs, Circular Sustainability Strategies*. The personal evaluation part focused on determining the organization, position, and quantity of physical machines each participant was overseeing. This is crucial for linking various sustainability approaches to the number of managed servers. Additionally, participants are asked to provide their contact information..

The second section aims to identify current habits and processes of participants when purchasing new hardware for data centers. The paper of Hardy mentions a specific differentiation between hot and cold spare machines [34], making it an interesting aspect to evaluate in practice. Furthermore, it is important to know if budgeting, usage, and software requirements are real criteria for choosing server hardware.

The third section evaluates the current techniques used by professionals (participants) to decommission server hardware from the data center(s) they manage. The frequency of decommissioning, the costs associated with current methods, and a look at possible future improvements are the focus points of these questions. The purpose of this section is to establish current practice among data center professionals with respect to the decommissioning process [16, 7].

The final section focuses on the section on circular sustainability. This section presents participants with current theory as well as a series of similar questions to establish the current position of the respondent with respect to the 9R sustainability practices proposed by the United Nations [20].

### 3.1.2 Data Collection Procedures

A questionnaire was created using the Microsoft Forms platform and distributed through accessible links to all participants by email. Each participant was also offered the option to be contacted at a later date for an interview in person or by video conferencing. The completion time of the questionnaire was approximately 30 minutes. After making improvements to the Total Cost of Ownership model, interviewees who indicated their willingness to participate in the in person interview were contacted. In person interviews took $\approx$ 30 minutes in which interviewees were presented with findings and updates to the model.

## 3.2 Participant selection

Participants were selected based on known relationships in the data center industry or by persons found on Linked-In. The search criteria on Linked-In were individuals who were marked as having or having had one of the following roles: ***Datacenter Manager, Datacenter Director, Datacenter Supervisor***. Furthermore, individuals who had skills in Infrastructure as a Service (IaaS) were also considered, although this returned a smaller number of individuals.

Furthermore, we limited the number of participants to be within Europe, we also selected from a broad range of companies, all the way from small local data centers to

| Role | Population | Percentage of population |
|---|---|---|
| CEO | 1 | 12.5% |
| HPC Engineer | 1 | 12.5 % |
| System Administrator | 3 | 37.5% |
| Infrastructure Manager | 4 | 50% |

**Table 5:** The matrix of questionaire questions

large giants. The participants were grouped according to their role, and their distribution can be seen in Table 5.

## 3.3 Interview Findings

The interview was concluded in the end with a total of 8 responses. Following the responses of the participants, they have been grouped among the number of machines managed. This has been done to gain a better insight into the different approaches given the size of the enterprise they need to support.

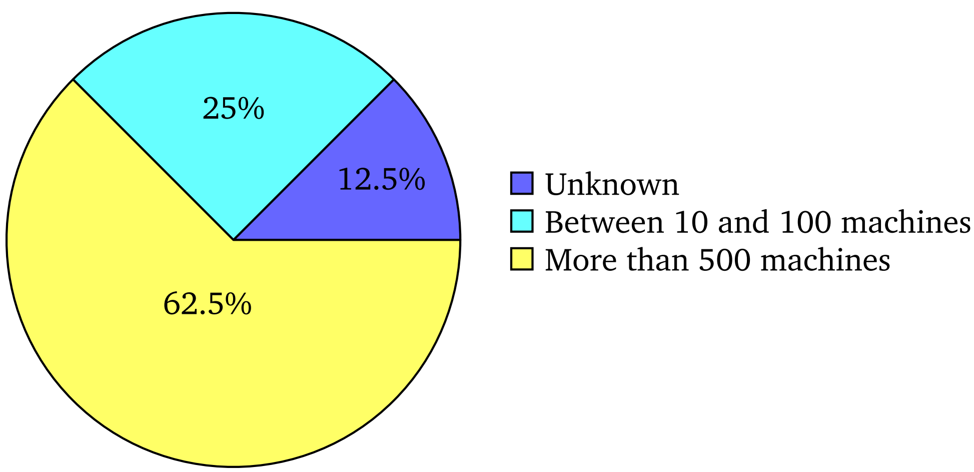



**Figure 10:** Size of the enterprise managed

### 3.3.1 Hot vs Cold

One of the questions aimed to understand the importance of warm and cold spare servers. The question was answered with five responses against any important implications between the cold and the hot spares, and 3 claiming that it is important to make a clear distinction between the two. Furthermore, one of the participants explained that in the case of small businesses, there is often no distinction between hot and cold. Although a server may not perform intensive tasks, a business intends to use it so that it is not deemed a *"careless expense"*, therefore not turning it off and turning it into a cold spare.

### 3.3.2 Power efficiency

Out of the eight participants, three have claimed that power efficiency is not important to them. This is an interesting insight as it shows that the industry, while on a whole it moves towards higher efficiency levels and improved sustainability, still has to raise awareness of why it is important to look out for efficiency of the hardware running.

In a separate question, participants are asked about the importance of cooling. Here all eight participants agree on a consensus that it is evident that cooling and power efficiency are critical in data center operations, significantly influencing total costs. Running 24/7, these facilities incur substantial power expenses, which requires high efficiency to minimize costs. High Performance Computing (HPC) nodes, particularly those with dense configurations and high-end GPU accelerators, present unique cooling challenges. Traditional air cooling is not only inefficient, but also unsustainable for racks consuming 60-100 kW, leading to a preference for Direct Liquid Cooling (DLC).

Although DLC is not yet standard across the industry, it is common in HPC, with some organizations pioneering efficient cooling designs with high inlet temperatures. Efficient cooling is vital, as it accounts for about 30% of a data center's energy use and directly impacts CPU performance. The push for high-density server configurations further emphasizes the need for advanced cooling solutions. Despite the focus on performance, incorporating the latest cooling technologies ensures long-term cost savings and operational efficiency, supporting high-performance systems over extended production periods of 5 to 7 years.

## 3.4 Purchasing new hardware

One of the questions that the participants faced was about the use of a set budget to purchase, repair, or improve hardware. Six out of eight participants have agreed that there are budget limitations that they need to adhere to. The remaining two participants have claimed that there are no monetary setbacks to improve their hardware. However, this shows that the larger majority of participants would be interested in minimizing costs in order to follow the budgeting scheme with which they have been given to work.

Furthermore, seven participants acknowledge that they aim to size their hardware in line with the software applications they expect to be running. This is a relief and shows that the industry, although not necessarily adapting a data-driven approach such as

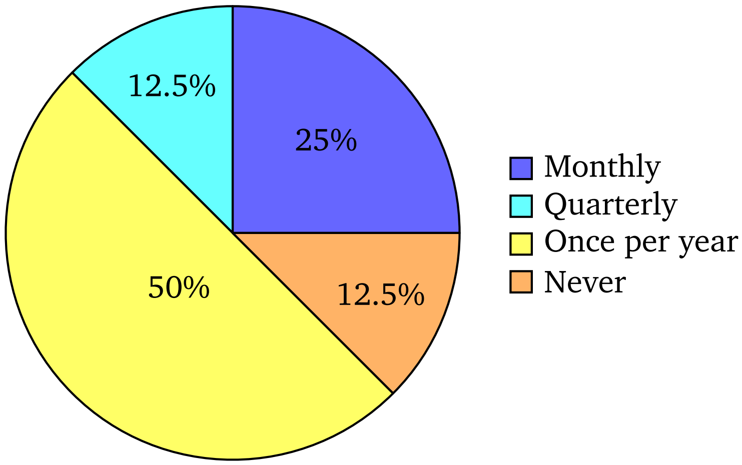



**Figure 11:** Frequency of hardware replacement

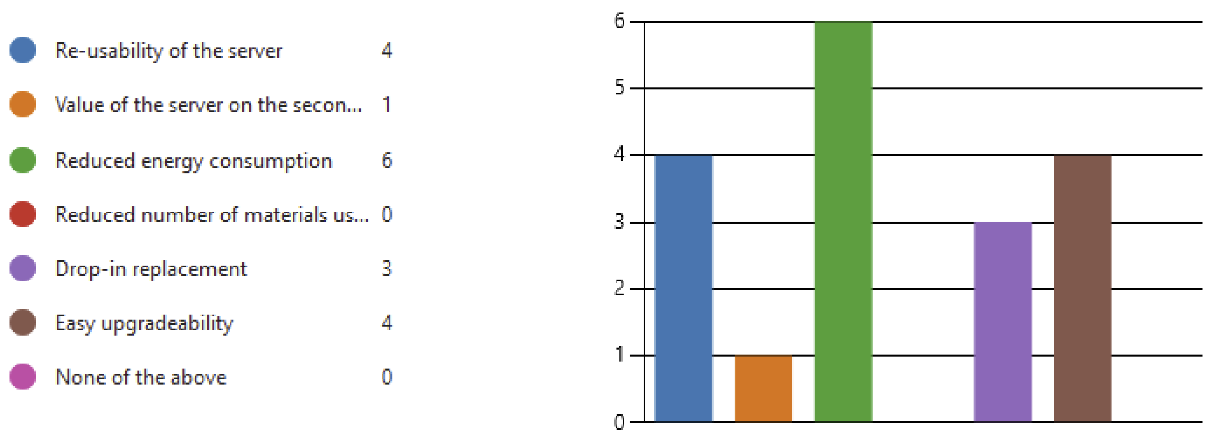



**Figure 12:** Interests shown when asked about the requirements of the server.

EcoUp [75], pays attention to the requirements with which it is faced and minimizes waste by sizing accordingly and refusing to invest in over-speced solutions. Figure 12 displays the interest show by the participants when purchasing new hardware. When asked about specific requirements that they consider when purchasing hardware, some of the participants mention:

*"Primarily price / performance and the expected service life. We consider the performance per watt and the reliability of the server. Compatibility with existing infrastructure and software."*

## 3.5 Decommissioning hardware

Based on the responses to the survey, the handling and disposal of old hardware in the data centers are managed through sustainable and certified processes. Specialist certified companies are typically engaged in recycling or reusing the hardware to the fullest extent possible. Many respondents indicated that their organizations require vendors to responsibly manage the transition and removal of old machines when delivering new ones. This includes sustainable recycling and dismantling efforts, often requiring vendors to be certified for these processes. For example, some vendors can profit from refurbishing parts due to their licensing agreements, while others may incur additional costs, affecting the overall pricing of new equipment. Organizations aim to maximize the service life of their equipment, repurposing it for less critical tasks or development environments before finally recycling it. The disposal process often involves brokers who guarantee data wiping and proper recycling, providing certificates for these actions. For security reasons, some hardware is broken down or shredded. Ultimately, leftover components are handled by recycling companies, ensuring that all data carriers are securely wiped or destroyed.

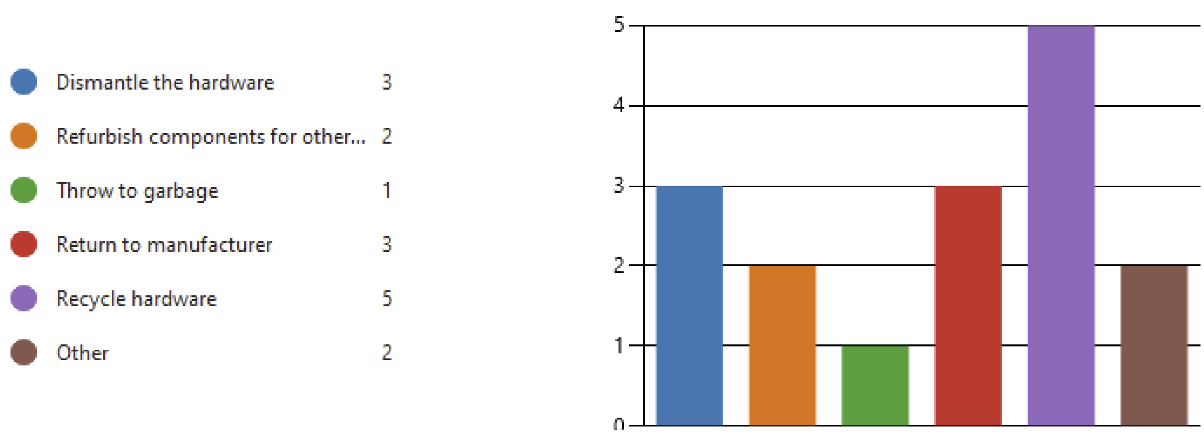



**Figure 13:** Methods applied by interviewees to decommission hardware

Furthermore, when asked if vendors should focus on using standardized, replaceable, and upgradeable components, six of the respondents agree with such a statement. One of the participants mentions that while he supports such a push for standardization, HPC clusters are often built from the ground up with engineering challenges in mind and cannot always adhere to the standards imposed. Packaging is also mentioned here. One of the participants argues that they often need to handle large amounts of packaging material which is difficult to recycle, such as plastic bags or styrofoam. The packaging materials then need to be sorted and then prepared for recycling, which drives up the cost of installing the new hardware. They claim that many of such packaging additions are not required for the products they use and are just old practices.

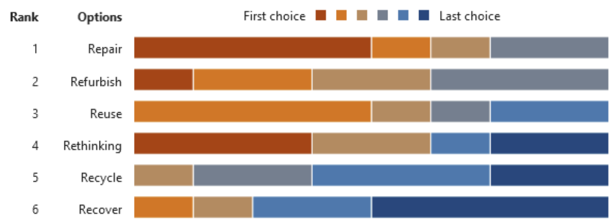



**Figure 14:** Distribution of votes for sustainable options

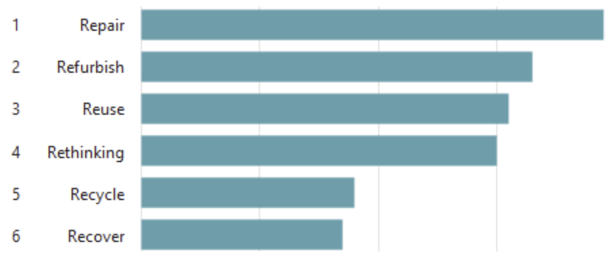



**Figure 15:** Final count of the sustainable options

### 3.5.1 Sustainable Data Centers

Based on interviews with various experts, a strategic approach to managing hardware throughout its lifecycle in data centers was evident. Figure 14 shows the distribution of votes in multiple categories, while Figure 15 shows the total number of votes relative to the other categories. An interviewee shared, ”During the lifetime in our environment, we use a lineup of choices. When acquiring a new service, we rethink what is needed for the service.” Another expert emphasized that, ”We make the vendor responsible for a number of these things from the outset and verify the vendor’s certification and credibility. We also have a support contract for parts to be replaced, etc..”

Proper planning and capacity alignment with expected workloads were highlighted as significant factors. ”By proper planning and aligning capacity with expected workload, we can have a great impact,” one participant noted. Another added, ”As stated before, we use our equipment until the end. We recycle it when we cannot find any use for it.” Reliability was also a key concern, with one respondent stating ”Hardware must be reliable. We use Life Cycle Management to guarantee reliable hardware for research and education.”

The interviewees also discussed the importance of reevaluating needs and repurposing equipment. ”Rethink if you need to upgrade, reuse what is still usable, repair when it is broken for reuse, or refurbish it for a second life elsewhere,” advised one expert. They also pointed out the environmental benefits, with comments such as ”Recycle and recover are the same for me.” In general, these insights highlight the careful consideration given to sustainability and efficiency in hardware management in data centers.

# 4 State of the art model

The section discusses the complexities of data center architecture, focusing on the integration of sustainability and circular economy principles into their operational and cost assessment models. Highlights the use of the ArchiMate modeling language to delineate the structure of data centers, which includes various technological, application, and business domains. The traditional Total Cost of Ownership (TCO) model, which focuses primarily on direct monetary expenses, is criticized for its inadequacy in addressing environmental impacts and sustainability. The section advocates for an expanded TCO model that incorporates sustainability metrics, such as energy efficiency and waste management, and introduces new cost categories such as *"Environmental Impact"* and *"Circular Economy Initiatives"*. This adaptation aims to provide a more comprehensive understanding of the holistic costs associated with data center operations, including the life-cycle impact of servers and the integration of component-level costs into maintenance strategies.

## 4.1 The Data Center Architecture

Data centers are complex businesses that are made up of complex business models, a wide and large IT landscape, and complex facilities to allow the working of the servers, while minimizing the effects on the environment. To model it in complete detail would make for a very extensive paper, well outside the scope of the current one. Therefore, some assumptions are made about the business model of the data center. Figure 16 shows an enterprise architecture model of the data center. The model is created using TOGAFs Archimate modeling language [3, 41].

> For readers unfamiliar with the *Archimate* modeling language, the ArchiMate specification offers tools that allow Enterprise Architects to clearly delineate, examine, and illustrate the connections between business domains. It is important to make a distinction between the following domains:
>
> - Technology, marked as green entities, represents physical assets to the business.
> - Application, marked as blue entities, represents software or software components in the enterprise.
> - Business, marked as yellow entities, represent business processes or functions within the enterprise.

Figure 16, describes the data center as a physical facility which is composed off of multiple server clusters as well as cooling equipment. Furthermore, there exist multiple

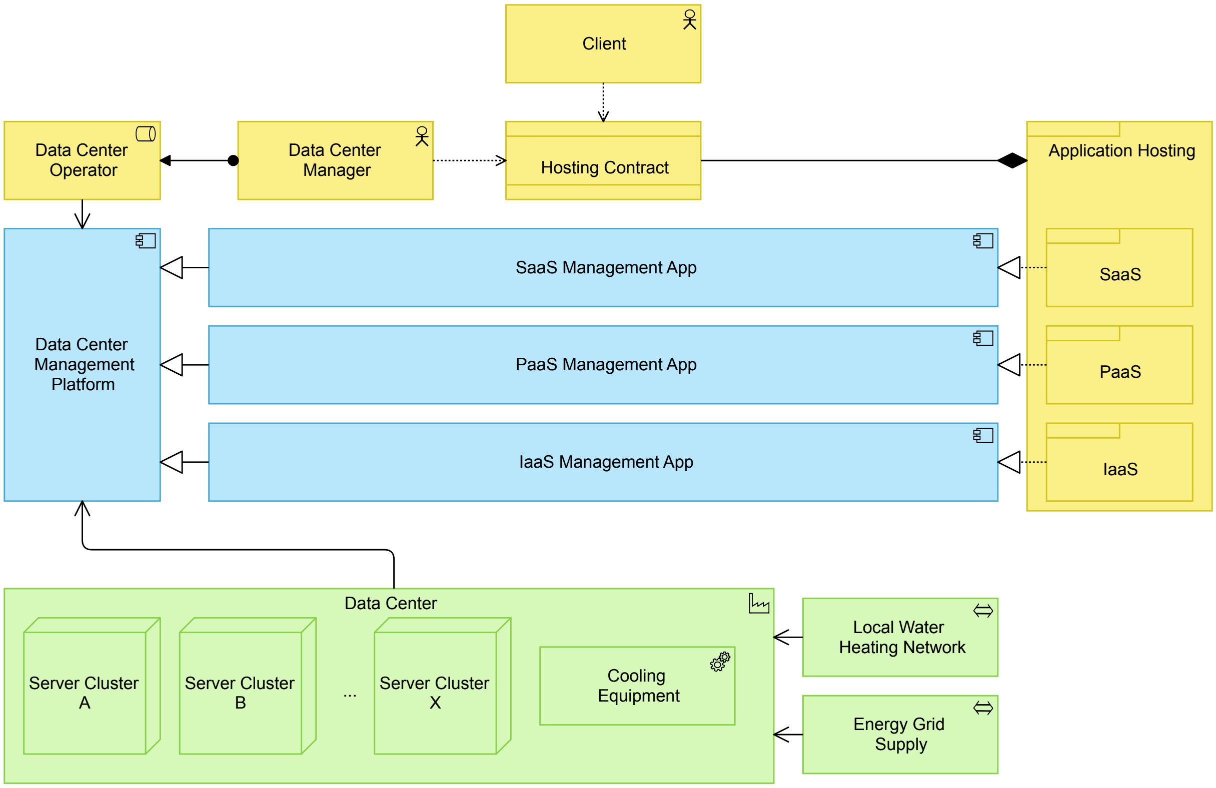



**Figure 16:** A business model for the data center

connections to the local water district as well as the electricity grid. As a general model, the data center serves a generic *Data Center Management Platform*, which encompasses all the management software required to manage the data center such as hypervisors for the servers, network management, cooling management as well as access in the data center.

The model also provides some specializations to the management platform, respectively the Software as a Service (SaaS),Platform as a Service (PaaS), and Infrsatructure as a Service (IaaS) applications. Those applications are intended to model hypervisors or other software capable of supporting related business cases. Lastly, linked to those application entities are business entities. The data center model is made up of a data center manager, who has the role of overseeing the entire well-working of the data center. There exists a specific role, taken by other employees namely *Data Center Operator* which operates the *Data Center Management Platform* or specific applications. Through the use of contractual agreements between the data center manager and a potential client, the business object of *"Application Hosting"* can be achieved, which in turn is linked to one or more of the management applications mentioned above.

Furthermore, it is important to consider the motivation of the *Data Center Manager* to pursue a change in the data center. Figure 17 illustrates the drivers, goals, and outcomes that the manager desires and should attend to in order to adopt a more sustainable way of working. This relates to concerns regarding the sustainable approach. Table 6 and Table 7 show the drivers, goals, and results, respectively.

During the interview, participants were also asked to answer questions formatted as a likert to identify their motivation for sustainability in data centers. Using the responses from the survey and combined with the trends identified in the literature review, a motivational view was created. Figure 17 shows how each of the drivers is related to the goals and further to the results.

| Driver | | Explaination |
|---|---|---|
| Improve the reliability of the data center | | As a data data center manager, the reliability of the hardware is of high importance, therefore planned hardware decomissionings take place, to replace old hardware which is growing a risk of failure |
| Become sustainable | | Due to pressure from competition, high energy prices and tighter regulations, data center managers need to find ways in which they can minimize the costs and grow revenue |
| Profitability | Costs | As any other business, profitability is affected negatively by the increase in costs, therefore it is a driver to decrease the costs as much as possible |
| | Revenue | Equally to costs, revenue positively impacts profitability and businesses strive to increase revenue |
| Reduce Electronic Waste (WEEE) | | Due to tighter regulations, data centers need to reduce their impact on the waste generation of electronic waste |
| Improve the scalability of the data center | | Analogue to reliability, regardless of the hardware chosen, it must ensure that the data center is capable to scale to needs. |

**Table 6:** Table displaying the drivers for sustainability

## 4.2 The model

According to the Total Cost of Ownership (TCO) model proposed by Hardy et al. [34], the evaluation of costs within data center operations covers five primary categories:

| Expected Outcome | Goal | Argument |
|---|---|---|
| Reduce decommissioning of hardware | Increase the life-span of server hardware | By offering older hardware as lower packages, data centers can increase the life span of existing hardware and reduce the costs associated with decommissioning |
| Reduce $CO_2$ Emissions Adhere to sustainability regulations | | By increasing the life-span of hardware, $CO_2$ emissions are reduced (there is no need to produce as much) and the business is more likely to adhere to sustainability regulations |
| | Be competitive on the market | Taxes imposed on the consumption of energy and generation of $CO_2$ contribute to extra costs for data centers. By minimizing those taxes, have the upper hand in becoming more profitable. Furthermore, by offering more services, it increases competitiveness on the market |
| Reduce the generation of WEEE | | Waste management costs money in most cases, therefore reducing waste directly increases the sustainability of the data center. |

**Table 7:** Goals and their intended outcomes

*Server*, *Power*, *Maintenance*, *Network*, and *Infrastructure*. However, with the growing emphasis on sustainability and the integration of circular economy principles into data center management, it becomes evident that the existing model presents constraints in accommodating these evolving considerations within its current framework. Expanding beyond traditional cost considerations, incorporating sustainability metrics and circular economy principles into the operational assessment of data centers is becoming imperative. This expanded focus covers not only the immediate financial expenses but also the environmental impacts, resource efficiency, and longevity of the infrastructure. When sustainability is explored, factors such as energy efficiency, renewable energy use, reduction of carbon footprints, and waste management become crucial determinants of operational costs.

> Table 8 displays all the parameters present in the proposed model. Notice the difference between notations, ***C*** parameters stand for costs, ***D*** notations stand for a duration of time, ***N*** duration stand for a quantity or a representative number, ***P*** represents parameters related to power figures, while ***O*** parameters represent possible opportunities.

Likewise, the circular economy concept, which emphasizes resource conservation, reuse,

| Mathematical | Parameter |
|---|---|
| $PUE$ | Power Usage efficiency |
| $A_{perRack}$ | Aria of one server rack |
| $C_{building_extra}$ | Extra building costs |
| $C_{building/sqmeter}$ | Costs of building per square meter |
| $C_{cooling}$ | Cost of cooling equipment |
| $C_{cooling/watt}$ | Cost of cooling |
| $C_{electricity/KWh}$ | Cost of electricity |
| $C_{employee}$ | Cost of salaried employees |
| $C_{downtime}$ | Cost of downtime |
| $C_{planning}$ | Cost of planning for sustainability |
| $C_{components}$ | Expected cost of replacement components |
| $C_{dataSanitization}$ | The cost of sanitizing data before decommission |
| $C_{recycling/Kg}$ | The cost of recycling |
| $C_{serverUnit}$ | The cost of one server |
| $C_{networkUnit}$ | The cost of one networking equipment |
| $P_{serverPower}$ | The power consumed by one server |
| $P_{networkPower}$ | The power consumed by the network equipment |
| $D_{server}$ | The expected life-time of the server |
| $D_{network}$ | The expected life-time of the network equipment |
| $D_{datacenter}$ | Life-time longevity of the data center |
| $D_{avgRepair}$ | Duration of an average repair |
| $D_{avgDowntime}$ | Duration of an average downtime event |
| $D_{planning}$ | Duration of sustainability planning |
| $N_{racks}$ | The number of server racks installed in the data center |
| $N_{TbToBeSanitized}$ | The amount of data that needs to be sanitized |
| $N_{serverUnits}$ | The number of server units installed. |
| $N_{networkUnits}$ | The number of network equipment installed |
| $N_{serverWeight}$ | The weight of one server. |
| $N_{networkWeight}$ | The weight of one network hardware |
| $O_{sellingHardware}$ | The opportunity of selling hardware |
| $O_{estimatedSellingPrice}$ | The estimated value of hardware |

**Table 8:** Parameters present in the final model

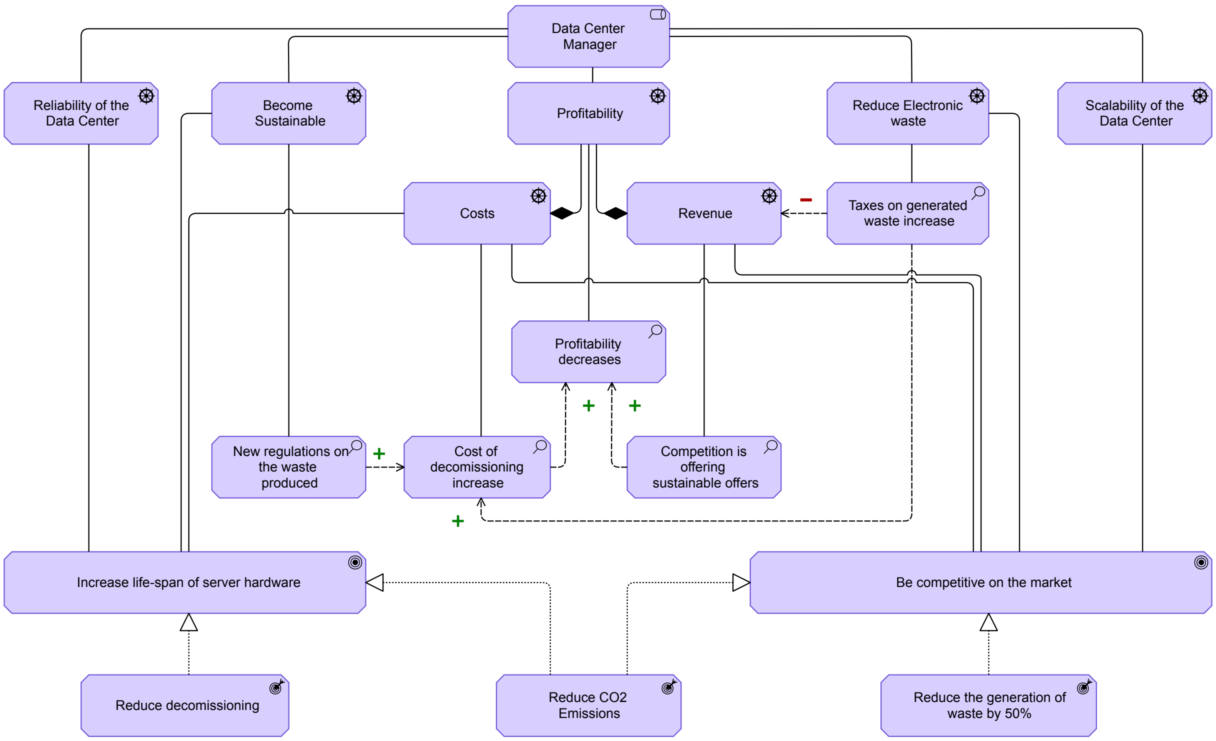



**Figure 17:** An Enterprise Architecture view of the motivational aspects behind sustainability in data centers

and recycling, introduces additional complexities to cost assessment. Incorporating principles such as product longevity, material recovery, and waste minimization into existing cost categories requires a more nuanced approach than what the conventional model offers. Therefore, while the Total Cost of Ownership (TCO) model provides a structured framework for assessing data center costs, it requires adaptation to accommodate the evolving priorities of sustainability and circular economy. This adaptation may involve the creation of new cost categories, such as ***"Environmental Impact"*** or ***"Circular Economy Initiatives"*** and the integration of relevant metrics and assessment methodologies to comprehensively assess the holistic costs associated with data center operations.

Integrating the cost of decommissioning into the ***Server Costs*** category offers a strategic approach to address sustainability considerations within data center operations. This expanded perspective enables users to account for the full life-cycle impact of servers, including their eventual decommissioning and disposal. Within this category, users can factor in additional costs associated with sustainability features and certifications that influence the server's environmental footprint.

Furthermore, the current approach to maintenance costs primarily focuses on human labor expenses, overlooking the importance of component-level costs. Although labor costs constitute a substantial portion of maintenance expenses, they do not encompass the full spectrum of resources involved in equipment maintenance. Incorporating component-level costs into the maintenance category provides a more comprehensive understanding of the financial implications associated with the maintenance of the data center infrastructure over its operational life.

By accounting for the costs of components such as spare parts, consumables, and specialized tools required for maintenance activities, users can better assess the total cost of ownership and make informed decisions about equipment maintenance strategies. This nuanced approach enables organizations to optimize maintenance practices, minimize downtime, and extend the longevity of data center assets while effectively controlling operational expenses.

# 5 Complete Total Cost of Ownership Overview

This section updates Hardy et al.'s Total Cost of Ownership (TCO) categories, emphasizing the importance of cold spares and refining maintenance costs calculations. Cold spares are crucial for immediate hardware replacement, minimizing downtime, and maintaining operational continuity. Despite upfront costs, they are long-term cost-effective by avoiding extended downtime and expedited shipping expenses. Furthermore cold spares, reduce IT staff stress, help in testing and upgrading, and improve inventory management while complying with regulatory standards and industry best practices. Warm and hot servers offer higher redundancy and faster recovery times, but incur higher operational costs; the use of refurbished hardware is suggested for sustainability. The current TCO model lacks comprehensive operational costs, such as hardware repairs, software licensing, and service agreements. A new formula is proposed to include these elements, which offers a complete financial picture. Service costs, covering power, cooling, security, and compliance certifications, are often passed on to customers through fees. Software licensing reflects recurring costs under SaaS models, calculated per unit over the contract duration. Finally, the power-cost formula is updated to consider multiple server configurations for a more accurate estimation. In general, these additions provide a more accurate and detailed financial overview of data center operations.

## 5.1 Importance of hot and cold spares

Cold spares play a crucial role in the operation and maintenance of data centers for several important reasons. First, they allow immediate replacement of faulty hardware components, such as servers, storage devices, or network equipment. This rapid replacement capability is essential for minimizing downtime, which is critical to maintaining the high availability and reliability that data centers are expected to provide. Minimizing downtime is not only about maintaining operational efficiency, but also about preventing significant financial losses and protecting the company's reputation. By having cold spares on hand, data centers can quickly restore full functionality, thereby ensuring they meet service level agreements (SLAs) and keep customers satisfied.

From a financial perspective, maintaining an inventory of cold spares, despite the upfront cost, can be more cost-effective in the long run. It avoids the extended downtime and higher expedited shipping costs that might occur if replacements had to be ordered after a failure. Thus, cold spares mitigate these risks and associated expenses. Operational continuity is another critical benefit. Data centers often support critical applications and services that must operate continuously. Cold spares provide an immediate backup option, ensuring that there are no interruptions in service. This is particularly vital in industries where uninterrupted service is paramount, such as financial services, healthcare and cloud computing [37].

In addition, having cold spares reduces the stress on IT staff. When a failure occurs, they can quickly replace defective components without the added pressure of prolonged downtime or customer dissatisfaction, leading to a more manageable and less stressful work environment [62]. Cold spares also serve a valuable purpose in testing and upgrading. IT staff can use them to test new configurations or updates before deploying them to live systems. This ensures compatibility and prevents interruptions in ongoing operations. In addition, cold spares contribute to better inventory management and planning. Data center managers can monitor spare usage and predict future needs, ensuring that stock levels are always adequate to handle potential failures [62, 28].

Lastly, from a compliance and risk management perspective, having cold spares can be part of a strategy to meet specific regulatory standards or industry best practices. It shows a proactive approach to handling hardware failures, which can be a requirement in certain compliance frameworks [37, 69, 42].

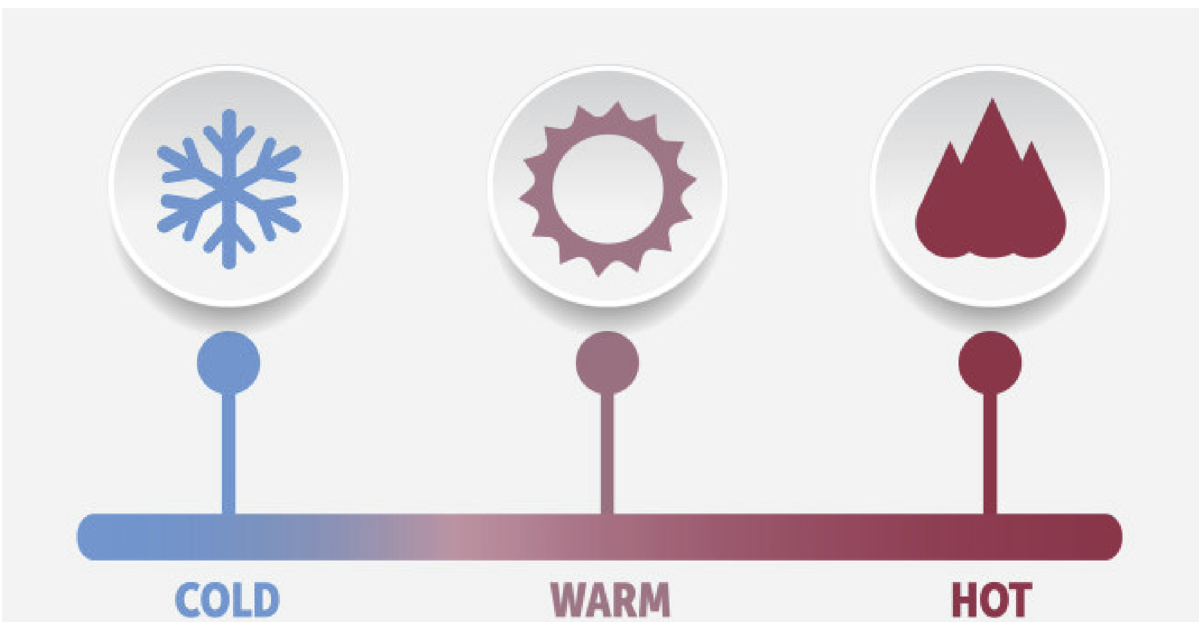



**Figure 18:** Different types of redundancy [39]

As shown in Figure 18, warm and hot servers offer more redundancy compared to cold ones. They are powered on and ready to take over in case of failure, but do not handle the primary workload under normal conditions. They are partially synchronized with the primary servers, often maintaining up-to-date copies of data, or are partially configured to take over tasks. This setup reduces downtime compared to cold servers, as the transition involves less delay. However, warm servers incur operational costs higher than cold servers because they consume power and resources to stay ready.

Hot servers are fully operational and are continuously synchronized with the primary servers. They either handle a part of the workload or stand by in a fully redundant setup to take over immediately without any noticeable delay in case of failure. This arrangement offers the fastest recovery time, ensuring high availability and reliability. However, it is the most expensive option because of the continuous consumption of power, cooling and resources, as well as the constant synchronization and maintenance required.

> While cold or warm servers provide lower maintenance costs, they provide low-to-no-revenue generations for the data center business. Furthermore, from a sustainability perspective, it is considered a waste to buy new, expensive hardware which is only intended to be used in critical situations. This is not to say that redundancy should not be taken into consideration, but rather that the servers which make up the warm and cold pools be refurbished hardware.
>
> In case of an outage where hot servers need to be replaced, work can be moved to warm or even cold servers, which offer less performance than their hot counterparts due to them being older generation servers or refurbished hardware. By making this switch, data centers can reuse older hardware, lower costs on acquiring new hardware components, and provide redundancy services to customers.

Most servers operate at about 5% to 15% capacity while using full power. Therefore, you are essentially using more energy to power less server utilization. Migrating data from your lightly used physical servers to virtual machines allows you to maximize server space and reduce the number of physical servers drawing power at any given time. Connect to interview answers

## 5.2 Maintenance Cost

Hardy et al. describe maintenance costs as the costs of repairing hardware, as well as the cost of employees. They define the maintenance cost as follows:

$$C_{maintenance} = \frac{N_{coldspares} \times C_{srvmodules}}{D_{srv} \times 12} + (N_{racks} \times C_{salaryperrack}) \tag{12}$$

Although its approach incorporates the salaries of employees, it falls short by not including a comprehensive range of other operational costs essential to the functioning of a data center. These omitted expenses include significant repair costs that arise from hardware failures or infrastructure maintenance. In addition, it overlooks the ongoing expenses associated with software licensing. Service costs for electric grid connections, which ensure a reliable power supply, are also not taken into account. Likewise, consistent internet connections are essential for continuous data transmission and network connectivity, but these costs are overlooked. Other contractual obligations, such as agreements with third-party service providers for various maintenance and support services, are also missing from their cost analysis. By not encompassing these additional operational costs, the approach presents an incomplete financial picture of data center operations potentially underestimating the true costs involved..

> A new improved formula for maintenance cost would be the following:
>
> $$C_{maintenance} = C_{Services} + C_{SoftwareLicensing} + C_{Employees} + C_{Repairs} \tag{13}$$
>
> With such a division, the updated maintenance cost allows for the distinction between services such as the electrical grid, internet connections, and other third party contracts; it defines a clear way in which to calculate software licenses and decouples the cost of repairs from the salaries of employees.

An important dimension in equation 13, is that of the decoupling between the cost of repairing and maintaining hardware and the salaries of employees. Although it is true that the cost of repairs grows linearly to that of the employee salary, it is not correct to assume that employees are paid per rack repaired or carry out only maintenance work. Furthermore, there can be discrepancies between the salaries of different employees, which makes the calculation more complex.

### 5.2.1 Service Costs

Service costs (represented by $C_{Serices}$ represent costs that the data center incurs and can represent, but are not limited to *power and cooling services*, *security services* or *compliance and certifications*. Data centers incur costs for electricity to power IT equipment and cooling systems to maintain optimal operating temperatures. These costs may be passed on to customers as part of the overall service fee or billed separately based on Power Usage Efficiency (PUE) or cooling capacity requirements.

In addition, data centers invest in physical security measures such as surveillance cameras, access controls, biometric scanners, and security guards to protect their facilities and customers' data. The cost of which is incurred by the data center; however, it is often included in packages to customers. In addition, data centers often undergo third-party audits and certifications to demonstrate compliance with industry standards and regulations such as SSAE 18 (SOC 1 and SOC 2), ISO 27001, HIPAA and GDPR [4, 42, 69]. Costs associated with obtaining and maintaining compliance certifications may be passed on to customers through higher service fees or surcharges.

Because of the nature of services not to generate direct revenue, service costs can be expanded by mathematical addition. Each new service can be defined as two variables, the complete cost of a contract, indicated as $C_{ServiceCost}$ and a part of the duration of the contract, indicated by $D_{ServiceDuration}$. The duration part is important to transpose the cost in daily, weekly, monthly, quarterly, or yearly costs. An example base formula is represented by:

$$C_{Services} = \frac{C_{ElectricalGridConnection}}{D_{ElectricalContract} \times 12} + \frac{C_{InternetConnection}}{D_{InternetContract} \times 12} \tag{14}$$

The equation defines two costs, one for the electrical grid and one for the internet connection. The duration of those contracts is considered to be in years; therefore, it is multiplied by 12 to get a monthly cost.

### 5.2.2 Software Licensing

While some software vendors offer a one-time license deal, in today's world, it is more and more common to find Software as a Service deals or different support schemes, which incur a recurring cost. These licenses can vary from software to software and, therefore, are defined by a contract and a term. While this is the same as the cost definition for services, it must be noted that while services are individual, or form individual acts of service, a license can be bound to only a number of components of the data center.

$$C_{SoftwareLicensing} = \sum_{n=1}^{k} \frac{N_{Units} \times C_{LicensePerNode}}{D_{LicenseDuration}} \tag{15}$$

Equation 15 presents an example of how to compute the licensing cost. It is formed as the sum of the total cost of the license divided by the duration. The total cost uses the variable $N_{Units}$ to define the number of units to which this license is applicable. A unit can be any object or component, as some licenses are bound to an entire server chassis, system, Central Processing Unit or software.

> Please note that $D_{LicenseDuration}$ represents the period for which the license is paid. This duration must therefore be converted into the unit of measurements in which the entire TCO is computed: months or years.

### 5.2.3 Repairing Costs

While every hardware enthusiast hates to admit it, hardware failures do occur, especially in systems that are used for extended periods of time. Therefore, one of the options that promotes the sustainable use of materials is to replace individual components when they fail rather than replacing the entire server infrastructure. With the abundance of components on the market, it is impossible to arrive at an exact number; rather an estimate for the cost is made with grounded ideas and realistic costs in mind. The cost of repairing a server ($C_{Repair}$) involves several key components: hardware component costs, labor costs, downtime costs, shipping costs, and maintenance contract costs. The hardware component costs ($C_{Components}$) encompass the price of new parts needed or the expense of refurbishing existing ones. Labor costs are calculated based on the time IT personnel spend diagnosing, replacing, and testing the components, multiplied by their hourly rate. Downtime costs account for the financial impact of the server being non-operational, which includes lost revenue, penalties for failing to meet service level agreements (SLAs), and reduced productivity.

$$C_{repair} = C_{components} + (D_{labour} \times C_{labour}) + (D_{downtime} \times C_{downtime}) \tag{16}$$

## 5.3 The cost of power

$$C_{power} = PUE \times \frac{(C_{elecperKWh} \times 30 \times 24)}{1000} \times (SPUE \times P_{totalsrv} + P_{totalnetwork}) \tag{17}$$

The cost of power is computed in the initial TCO model as the efficiency of the data center PUE multiplied by the cost of electricity per month multiplied by the efficiency of the server. Considering that a data center may have more than one server configuration, we propose equation 18 which considers $\boldsymbol{k}$ possible server configurations. Because the power consumed by the servers is calculated in the International System reference [55], watts (W), the cost $C_{elecperKWh}$ must be converted from KWh to Wh.

$$C_{power} = PUE \times C_{elecperKWh} \times \sum_{i=1}^{k} (SPUE_i \times P_{i\_totalsrv} + P_{i\_totalnetwork}) \tag{18}$$

# 6 Right Hardware for the job

This section aims to provide the reader with a selection of best practices for deciding which hardware to acquire. It follows the example of EcoUp [75], as well as the well-known scaling options.

The section discusses strategies for acquiring suitable hardware in the context of sustainability and efficiency in information and communication technology (ICT). It emphasizes the importance of aligning hardware procurement with sustainability policies and goals. Firstly, it suggests conducting a thorough requirement elicitation to understand the hardware needs and their connection to sustainability objectives. Considerations include scaling options such as vertical and horizontal scaling, particularly pertinent in cloud computing environments accommodating diverse tenant requirements. Moreover, it advocates for selecting hardware designed for easy disassembly, enabling repair and updates without causing damage. This entails providing clear instructions for non-destructive repair or replacement of components and ensuring availability for authorized third parties for an extended period. Furthermore, the section discusses the significance of firmware in hardware selection, especially in data center settings. Firmware updates can improve power efficiency, security, and functionality, thus prolonging the useful life of the hardware. In addition, it briefly touches on the implications of software licensing and elasticity in distributed systems, suggesting potential avenues for optimizing resource allocation.

## 6.1 Requirements elicitation

An essential first step to make ICT procurement more sustainable is to determine what is really needed and how it is related to sustainability policies and goals. Important considerations should be taken with respect to hardware requirements, and hardware should be acquired to meet these needs [75]. The scaling of the data center should be taken into consideration, where-as some data centers can only scale vertically, others run applications which scale horizontally, especially in the context of the cloud where different tenants have specific requirements (see picture Figure 19). In other words, not all applications are built to take advantage of large systems, which a data center operator can take advantage of when specifying the requirements of the data center [59]. Moreover, by directly eliciting such requirements, the data center business implements the first R of the 9Rs[29, 47], respectively, refusal. Their actions directly refuse the purchase of extra hardware that has the possibility to not be used at its fullest.

Large data centers tend to have more than 50 thousand machines [8, 12]. With such large installations, it is essential that data centers use central infrastructure management software. Therefore, one of the important requirements that new hardware must meet is that it work with the existing infrastructure management setup [45]. Regu-

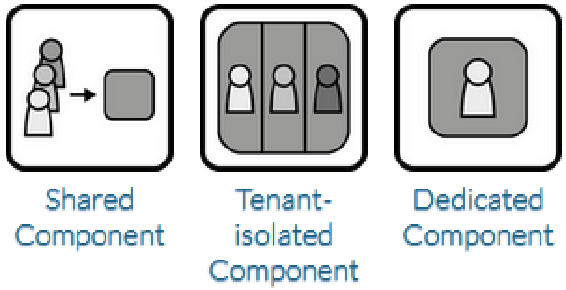



**Figure 19:** Cloud tenancy types [25]

latory bodies can also issue Approved Vendor Lists (AVL) AVLs are lists that specify the vendors which are approved for wide-use. This is important to prevent malicious actions.[76, 9, 17].

## 6.2 Disassembly oriented hardware

Critical Raw Materials (CRM) refer to materials essential to the economy and whose supply may be at risk. These materials are crucial for various industries, including manufacturing, technology, and energy production. CRMs are materials that are essential for the functioning of modern economies, but are subject to high supply risk. This risk can arise from geopolitical factors, scarcity, or lack of sustainable extraction methods. CRMs are indispensable in numerous sectors, including automotive, aerospace, electronics, renewable energy, and pharmaceuticals [22]. Identifying CRMs involves assessing their economic importance and supply risk. The European Commission has established criteria for identifying CRMs, which include economic importance, supply risk, and the ability of the EU to influence supply conditions. The list is periodically updated on the basis of evolving market dynamics and geopolitical factors. The European Union (EU) has a well-defined list of CRMs, which undergo regular review based on their economic importance and supply risk [23, 24].

> Although it is important to mention the relevance of Critical Raw Materials, it needs to be acknowledged that data centers are not always in a position to influence the materials used by hardware vendors. Although some large data centers might have the financial backing to influence the design of specific hardware and advocate for the usage of fewer Critical Raw Materials, many more use pre-defined configurations from various vendors.

In the IT sector, several Critical Raw Materials (CRM) are essential for the manufacturing of computers, servers, and data centers. Some of the most commonly used CRMs in this sector include:

1. Rare Earth Elements (REE) such as neodymium, dysprosium, and praseodymium are crucial for manufacturing magnets used in Hard Disk (HDD), Optical Disc Drive (ODD), and electric motors in cooling fans and power supplies.

2. Germanium is used in the production of optical fibers and infrared optics used in data transmission and optical communications equipment.

3. Platinum Group Metals such as platinum, palladium, and rhodium are used in various electronic components and catalysts found in IT equipment, including servers and networking devices

4. Cobalt is used in the production of lithium-ion batteries that power many devices, such as laptops, smartphones, and tablets. These batteries are also used in uninterruptible power supplies (UPS) for data centers or the motherboard of servers and computers.

5. Tantalum is used in capacitors, which are essential components in electronic devices to store and regulate electrical energy. Capacitors are widely used in computers, servers, and other IT equipment.

6. Gallium is used in the production of gallium arsenide (GaAs) semiconductors, which are used in high-speed electronic devices such as transistors and diodes. GaAs-based components are used in high-performance computing applications and telecommunications equipment.

> Data centers have the option to avoid discarding those materials as waste by choosing to reintegrate them into the material stream through refurbishment or recycling. This approach conserves valuable resources such as Critical Raw Materials (CRM). By adopting these practices, data centers can contribute to a more sustainable and circular economy, leading to advantages for both the environment and their financial performance.

These CRMs play a vital role in the functionality and performance of IT equipment, and their sustainable sourcing and recycling are essential to ensure the resilience of the IT supply chain. In addition, efforts to reduce the dependency on these materials through material substitution, design optimization, and recycling initiatives are becoming increasingly important in the IT sector [21]. Therefore, data centers are advised to pay attention and consider the design of the server machine to ensure a design that allows repair and update [17]. Requirements must provide clear instructions to enable a non-destructive repair or replacement of Data Storage Devices (HDD, SSD), Random Access Memory (RAM), Central Processing Unit (CPU), Expansion Cards (Graphical Processing Unit), Power Supply Unit (PSU). A list of current recycling statistics can be found in

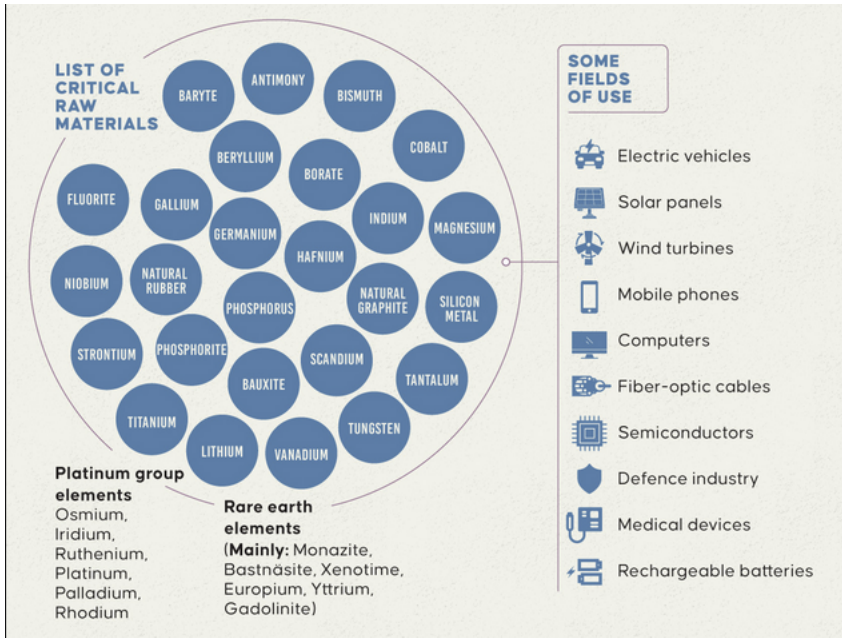



**Figure 20:** Critical Raw Materials and their fields of use [2]

Figure 21.

According to literature, one of the best practices when acquiring new hardware is to present instructions which are available to authorized third parties, including brokers, spare parts repairers, spare parts providers, recyclers, and maintenance providers via registration on the manufacturer's website. These instructions must be available for a minimum of 8 years after the server product is placed on the market [17, 76, 27] and guide relevant parties to safe dismantling of the hardware for later reuse or recycling.

## 6.3 Costs of Software

### 6.3.1 Firmware

Unfortunately, IT solutions consist not only of hardware but also of software. Firmware is one of the major factors in hardware selection, especially in the context of data centers [15]. Firmware consists of small, low-level code fragments meant to coordinate the communication between a server's components, which over time may get improvements or even fixes for certain issues reported or to improve security. Furthermore, firmware updates may include new or improved functionality, meaning that the old server hardware receives an improvement in terms of power efficiency or tool set. These are important aspects when purchasing new hardware, as they dictate the usage period.

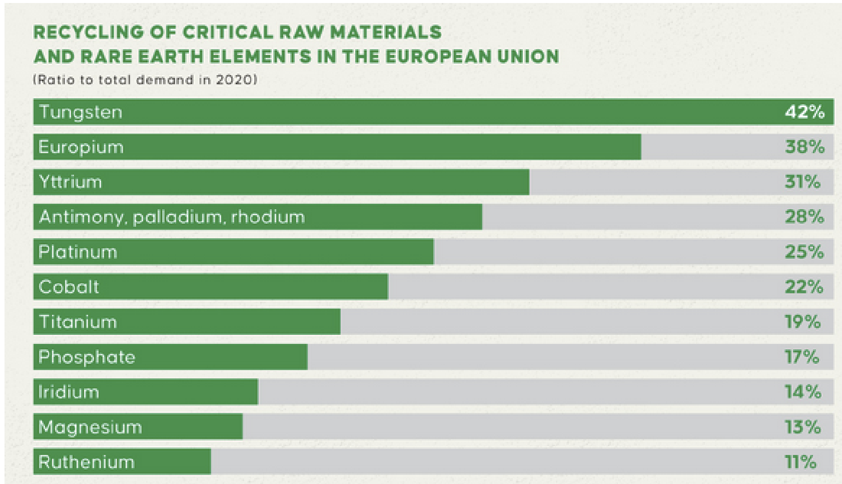



**Figure 21:** Percentage of CRM volume recycled in the European Union [2]

Analogously to this are the update cycles of smartphones. Studies show that phones that receive extended software updates from the manufacturer tend to have a higher market margin [76]. This is a beneficial factor for the user, as the most up-to-date software ensures security and bug fixes, as well as the possibility of new features. In the same context, hardware in data centers can be used for extended periods of time if the firmware and software support allow it. Frequently, vendors may impose an additional fee for extended software support, considering it a distinct service within the realm of business-to-business transactions.

### 6.3.2 Software licensing / Elasticity

As shown in subsubsection 3.3.1, while Hardy et al. argued for the important distinction between hot and cold systems, current practices ignore or do not acknowledge an important distinction. Although current practices do not show a distinction between the two and suggest that there should not be one, it is important to acknowledge that cold spares do offer data center businesses the elasticity required to perform critical work. Elasticity in distributed systems refers to the system's capacity to independently adapt resource allocation to meet varying workload requirements, a fundamental characteristic that sets cloud computing apart from earlier models such as grid computing. It involves dynamically adjusting resources according to demand, with the aim of enhancing resource distribution and user satisfaction.

Key obstacles include delays in resource allocation, intricate monitoring of resource allocation changes, addressing various stakeholder needs, and overseeing multiple levels of control strategies. Control-theoretic approaches, such as predictive control, present viable options to improve resource distribution in cloud settings. Therefore, the elasticity of the infrastructure must be considered as a requirement for any hardware that is currently running or that is considered. We distinguish here between two cases:

A Application servers which host additional unused resources such as spare RAM, CPU cores, external cards or storage. In such cases, the server can elastically scale and occupy resources for a given time, making sure to free them once they are no longer needed.

B Hot spare machines. In critical applications, hot spare machines are machines that are on and can do additional work, usually used in the context of load balancing. Such machines have the benefit of providing redundancy so that if one of the other server machines fails, they are ready to replace the work.

C Cold spare machines. In noncritical applications that can experience downtime, cold spare machines can be utilized. In such cases, the hardware requirements need to provide the minimum level of requirements for the application to run.

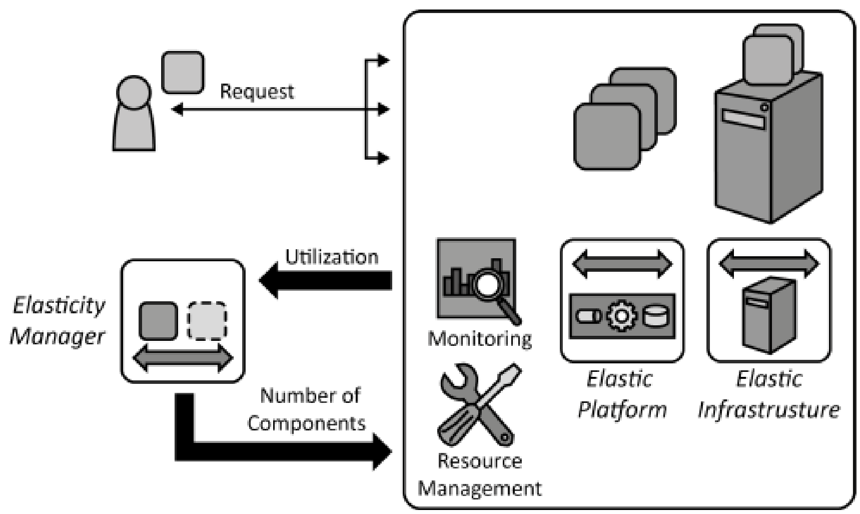



**Figure 22:** Elastic Management of Application requests [25]

Software licensing can also play a role as a limiting factor for hardware. Although many software vendors offer solutions for all hardware platforms, some only offer specific ones. Examples can range from x86 and x64 CPU variants, to ARM based SoC to newly introduced RISC V CPU architectures.

## 6.4 Sustainable Development Goals

By applying the requirements and recommendations expressed in this section, companies actively support the United Nations Sustainable Development Goals (SDG). The recommendations emphasize the selection of hardware that not only meets performance


**Figure 23:** Sustainable Goals as proposed by the United Nations [56]

requirements, but also aligns with various Sustainable Development Goals (SDG).

Firstly, it addresses **SDG 7: Affordable and Clean Energy** and **SDG 12: Responsible Consumption and Production** by recommending hardware that improves power efficiency, thus reducing energy consumption and promoting the use of clean energy in data centers. This includes advocating for firmware that supports energy efficiency and hardware designed for easy disassembly and encouraging repairability and updates without damage, thus minimizing waste and promoting the reuse of materials. Secondly, it aligns with **SDG 9: Industry, Innovation and Infrastructure** by promoting the use of modular, repairable, and recyclable hardware, fostering the development of resilient and sustainable industrial processes.

Finally, the emphasis on making repair and replacement instructions available to authorized third parties supports **SDG 17: Partnerships for the Goals**, encouraging collaboration between manufacturers, service providers, and data centers to achieve sustainability goals through shared knowledge and resources. These integrated approaches ensure that the hardware selection process in data centers not only meets operational needs but also contributes to broader global sustainability objectives. In general, the entire process supports **SDG 13: Climate Action** by proposing a solution aimed at reducing the carbon footprint of data centers.

# 7 Effects of sustainability

The following section describes the principles of circular economy with regard to data centers, examining their feasibility and associated costs. It begins by introducing the dynamic nature of data centers and the importance of circular sustainability in reducing waste and improving efficiency throughout their lifecycle. While some circular economy concepts like Refuse, Reduce, Reuse, Remanufacture, and Repurpose aren't directly applicable to data centers due to their operational requirements, others like Repair, Refurbish, and Recycle are discussed as feasible strategies. Reuse, for example, involves considering options such as purchasing second-hand equipment or reusing existing hardware, focusing on potential cost savings and environmental benefits. Repair extends product lifetimes by fixing faulty components, while refurbishment involves modifying products to restore or enhance performance, often within factory settings. Recycling plays a role in salvaging usable components from decommissioned equipment, with non-operating parts sent to recyclers for treatment. In general, the text emphasizes the importance of integrating circular economy principles into data center operations to minimize waste and environmental impact while optimizing resource utilization.

## 7.1 Why choose circular sustainability

There might be a temptation to view data centers and IT facilities as unchanging, where the installations remain in place for extended periods before being decommissioned. However, in reality, most facilities experience continuous activity: new assets are introduced and set up, while current ones are updated, reorganized, repurposed, relocated to different areas, or substituted.

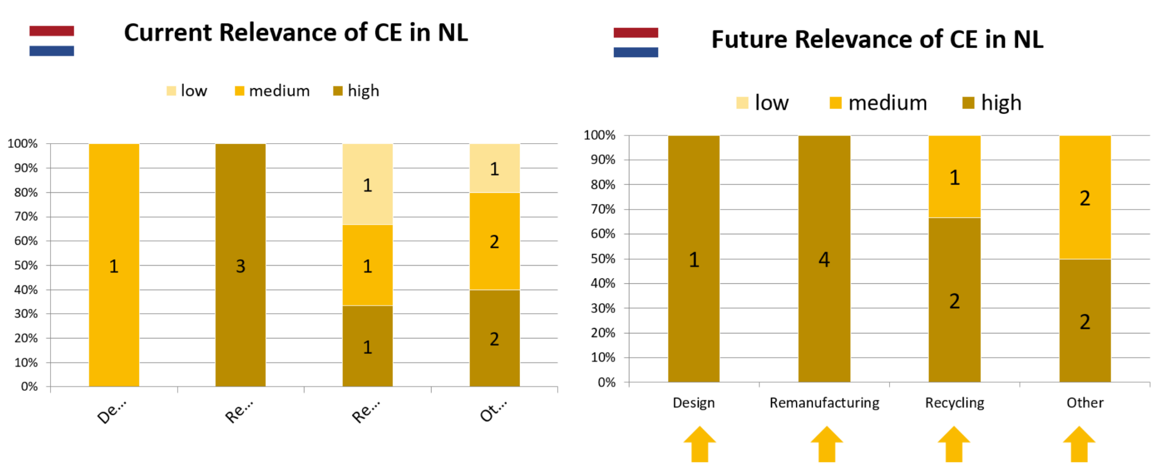



**Figure 24:** Importance of Circular Economy

Circular sustainability represents a comprehensive framework that prioritizes waste reduction and efficiency optimization throughout product life cycles [29]. This approach is of significant importance to mitigate the environmental repercussions of consumption and production processes. By reducing resource depletion and minimizing pollution and greenhouse gas emissions, circular sustainability stands as a cornerstone for sustainable development. In addition, its emphasis on innovation and collaboration fosters novel business models and economic opportunities.

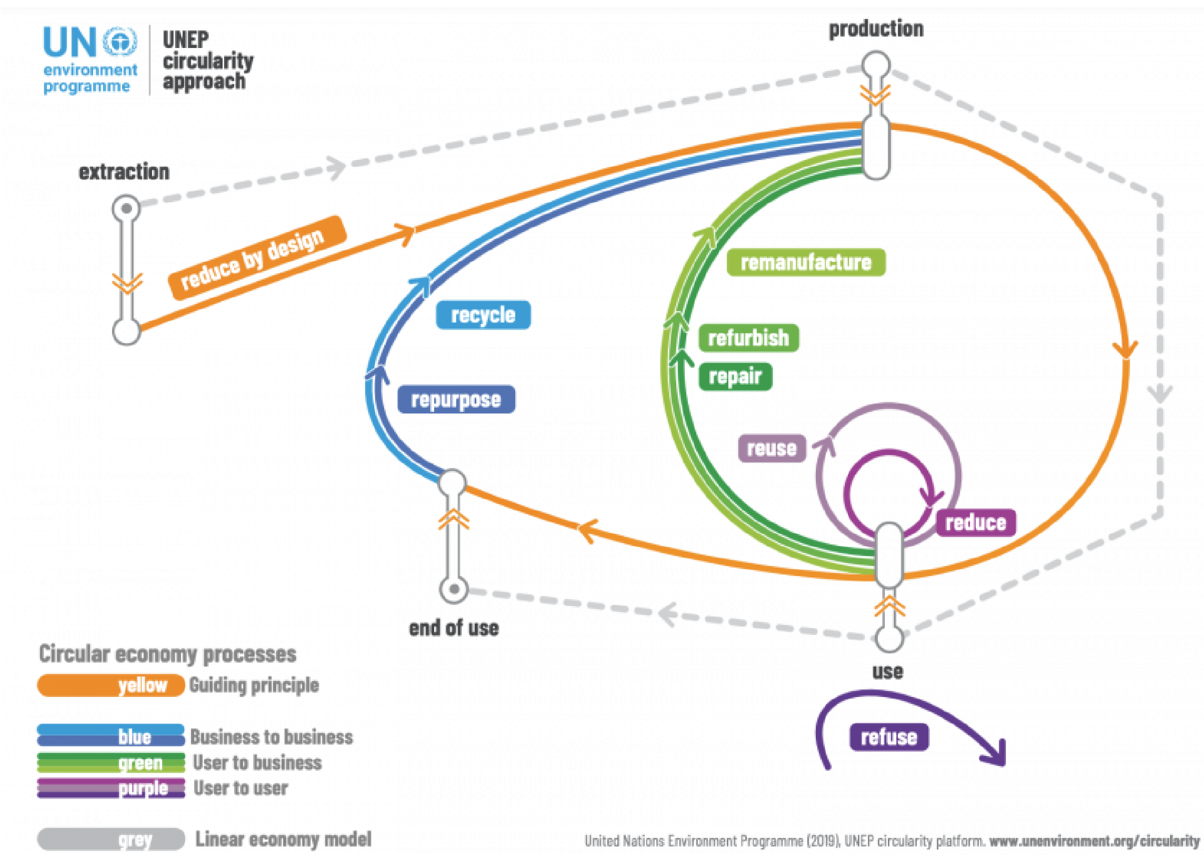



**Figure 25:** The 9 RE steps towards a circular economy [20]

Given that each of the circular strategies acts at a different level, a discussion is made on the feasibility and costs involved with each of the strategies in the context of data centers.

## 7.2 The non-applicable

While academic studies, as well as the United Nations Environment Programme (UNEP), argue for the importance and distinction of the 9Rs [47], present in Figure 25. However, when considering the review of the literature in section 2 and other related research on the topic, such as the EcoUp model or the refurbishment strategies [16, 29], it becomes clear that not all aspects can be applied to the *data center* industry. The following subsections express concerns and reasoning for the exclusion of the following concepts in the circular economy of data centers: ***Refuse, Reduce, Reuse, Remanufacture, Repurpose***. The UNEP categorizes some of the concepts of the circular economy in the interaction between actors. It differentiates between users, who are generally considered day-to-day end users of a product, and businesses, which have the capacity to generate additional value for a product which is considered at its end of life.

> Not all 9Rs are relevant in practice for the data center industry. From one perspective, data centers act as users of products and can **repair** and **refurbish** their own servers, something the United Nations Environment Programme considers to be an action present only between users and businesses. Meanwhile, data centers are businesses, meaning that they can better interact with other businesses to facilitate recycling of server machines.

### 7.2.1 User To User

The United Nations Environment Programme describes the user-to-user interaction category to illustrate the various behavioral decisions a user can take when buying or disposing of a product. Since data centers operate within a business context, all concepts in this category are either inapplicable or implicitly covered by other categories.

> User-to-User interactions are sparse in the data center industry. Such concepts operate on the underlying decisions of the user and not of the business and, therefore, are unrealistic in the data center industry.

***Refusing*** is a user's choice to buy or use less, saying no. It implies changing to more sustainable lifestyles, for example, rejecting packaging, shopping bags, or other products or services that are considered unnecessary [20]. Refuse can also be applied to a specific element of a product, such as refusing the use of hazardous substances in the design of a product. By refusing to buy or consume a specific product or service, users send a strong signal to the market, helping economies transition to more circular models. However, in the case of data centers, these strategies are not viable. The data center operators must be able to quickly adapt to current situations and events. Furthermore, since the business model of data centers is centered around the hosted infrastructure and the capabilities of the said infrastructure, it is clear that refusing to upgrade or replace server hardware in data centers is not a viable business strategy. At the same time, the method proposed in section 6 of acquiring hardware, more specifically the data-driven approach of EcoUp [75], directly promotes the concept of refusing to purchase hardware and prioritizes the purchase of relevant hardware for the business, in contrast to the vertical and horizontal scaling alternatives.

***Reduction*** involves people rethinking how they can meet their needs while minimizing their impact on the planet and those in their vicinity. It involves users opting to use items and services for an extended period of time and buying less frequently. Implementing the reduction concept is free and has a significant potential to prolong the value of a product or service over time [20]. It aims at reducing the impact on material depletion; however, in the context of data centers, it reduces overlaps with the concept of reuse. This is because while the data center operator decides to reduce material depletion, they, in fact, continue to reuse the same hardware. In the data center industry, the concept of reduction is achieved by performing maintenance, repairing, and refurbishing tasks on the existing hardware.

> The ideas of **refusal** and **reduction** are philosophical and not practical for data centers. However, they contribute to more detailed and practical concepts like renovation, repair, and recycling.

### 7.2.2 User To Business

Circular interactions can exist between users and businesses at the end of life of a product. These interactions occur in the form of repair, refurbishment, and remanufacturing of products. A user might contact other companies, including the original seller or independent firms, to dispose of their product, sometimes in the form of economic compensation [1, 70, 63].

Original Equipment Manufacturer (OEM) are transitioning to remanufacturing, indicating an increase in new capital allocations towards remanufacturing. This is a notable worldwide trend due to the growing acceptance of circularity and sustainability principles, driven by environmental laws and the pursuit of profit [35].

**Remanufacturing** contributes to sustainability by converting previously used goods back to functional condition, thus reducing waste creation in terms of both resources and energy. It completes the supply chain cycle, thus advocating the idea of a circular economy [32]. However, it has been shown to be difficult to remanufacture Waste of Electrical and Electronic Equipment (WEEE) products at a low raw material level. Furthermore, the remanufacturing of WEEE products requires extensive OEM planning [72, 67].

> Although remanufacturing takes place in other industries, the data center industry generates WEEE which makes it difficult to remanufacture at a raw-material level. In addition, the initial product vendors must offer the option of remanufacturing.

### 7.2.3 Business to Business

The last category offered by United Nations Environment Programme in its 9R strategies (see Figure 9 is Business-to-Business. The UNEP defines this category as a long loop, in which a product or component loses its original function through processes such as recycling and repurposing.

The process of **repurposing**, involves reusing discarded goods or components adapted for another function, the material has a different life cycle (i.e. plastics used in handbags). Converting old or discarded materials into something useful allows one to return them into the economy retaining some of its value, if not all of its value. In the context of data centers, this is often seen as a way to reduce heat by repurposing it as hot water for close neighborhoods [54, 46]. However, when it comes to WEEE, the product cannot be repurposed as easily.

> Data center server-grade components can indeed be sold on the second-hand market and repurposed by individuals in home-lab scenarios. Many tech enthusiasts and hobbyists use these components to build powerful and cost-effective home labs. These setups often use refurbished servers from brands such as Dell, HP, and Supermicro, which are designed to run 24/7 and handle substantial workloads efficiently. This practice is popular among those looking to experiment with virtualization, cloud computing, and home automation.
> Although this repurposing has become quite common, there have not been significant breakthroughs beyond these existing applications. The primary uses remain in enhancing home networks, running complex simulations, and hosting small-scale personal or development environment.

## 7.3 Reduce by design

As an overall principle, applied in the early stages of design, reduction by design leads to the design of products and services that use less materials per unit of production, and/or during their use. **Reduce by design** therefore influences all stages of the life cycle of a product or a service: less raw material is extracted, production has been designed to use fewer materials, consumption patterns, and end-of-life of such products and services are influenced by the design in order to lead to less impact and less waste.

$$C_{sustainability} = C_{planningPerMonth} \times D_{DurationOfPlanning} \tag{19}$$

Where $C_{planningPerMonth}$ is the cost of hiring employees to design and propose a system capable of meeting the requirements in section 6. $D_{DurationOfPlanning}$ is the number of months required to carry out the design. It is calculated in months because it is thought that it would take a person a month to make a complete proposal of the required hardware.

## 7.4 Reuse

Reuse refers to the continuity of usage of a product, object, or substance that is not waste, for the same purpose for which it was conceived, without the need for repair or refurbishment, according to the Basel Convention. Reuse and resell imply a user choice to hand over to another user, most frequently without intermediary and with no modification of the product or service. It applies to the use of second-hand products or products that are reused after cleaning.

> In the data center industry, the reuse of hardware does not incur additional costs in acquisition. Maintains the same costs for power consumption, cooling, and maintenance. Therefore, reusability as a concept involves a cost of 0 on the Total Cost of Ownership and can be *left out / ignored*.

However, when considering the industry of data centers and enterprise IT settings, the only viable options are **repairing** existing hardware, or, **refurbishing** exiting hardware, both of which imply the concept of reusing hardware. There does not emerge a clear definition of re-usability in the practice world of data centers; therefore, we propose the following definition:

> Reusability is translated as the choice of the data center operators to either *not* acquire new hardware *not* or to consider old, second-hand or refurbished hardware.

### 7.4.1 Repair

Repair refers to the repair of a specified fault in an object that is a waste or a product and / or replacement of defective components to make the waste or product a fully functional product to be used for its original intended purpose. Repair extends the useful life of the product, for example, by replacing broken parts or removing defects. A user sends its product for repair to a business intermediary, either from the retailer or directly from repair shops. The product returns to its original user or to a new one in working order.

When applied as a concept to the data center industry, repair can be seen as a means of replacing components of existing server machines in such a way that their initial capabilities have been restored. This process therefore comes at the cost of acquiring the new component at the time of breaking.

$$C_{repair} = C_{components} + (D_{labour} \times C_{labour}) + (D_{downtime} \times C_{downtime}) \tag{20}$$

From a business perspective, data center operators must determine if the cost of repairing the server hardware is a justified cost considering the required capabilities of the server machine. The alternative is to refurbish, recycle, or sell as waste.

### 7.4.2 Refurbish

According to the Cambridge Dictionary, refurbishment is *to repair and clean the equipment so that its condition is like new*. Furthermore, literature defines refurbishment as the modification of an object that is a waste or a product to increase or restore performance and/ or functionality or to meet applicable technical standards or regulatory requirements, with the result of making a fully functional product to be used for a purpose that is at least the one that was originally intended. Restoration of functionality, but not value, allows for a partial new service life for the product. Refurbishment that takes place within maintenance or intermediate maintenance operations.

> For the data center industry, refurbishment can be seen as two equally possible opportunities. On the one hand, data centers can acquire refurbished hardware in favor of new cutting-edge options usually at a reduced cost [60].

As the possibility of re-usability increasingly influences purchasing decisions for products which have previously been used, it introduces the possibility to sell hardware on the second-hand market, therefore resulting in a negative cost, otherwise known as profit. The profit can be considered as a maximum between deciding to re-use the hardware for the same data center, resulting into no extra costs for decommissioning or the profit made by selling the hardware on the second-hand market.

$$C_{Sustainability} = max(O_{Selling} \times O_{EstimatedSellingPrice}, 0) \tag{21}$$

Profit is calculated as the *opportunity* $O_{Selling}$ to sell a certain hardware artifact multiplied by the estimated selling price $O_{EstimatedSellingPrice}$. The opportunity is calculated as the probability of selling the hardware.

Modeling the opportunity and estimated selling price is required to represent the type of hardware sold. Proprietary hardware intrinsically has a lower opportunity to be sold compared to mainstream, standardized hardware. Furthermore, considering *opportunity* as a probability, the model can understand multiple possible sales prices at different opportunities.

## 7.5 Recycle

Recycling is the process of converting waste materials into new materials and objects. It involves collecting, sorting, processing, and remanufacturing materials such as paper, glass, metal, plastics, and electronics to produce new products. It is an essential component of sustainable waste management and environmental conservation efforts. However, in the data center industry, energy is considered the most recycled material. The current literature focuses on researching the improvements that can be made to data centers in order to reduce power consumption or even better to recycle energy through, for example, water cooling servers and reusing the heat generated to heat up neighboring neighborhoods.

> Recycling is an encompassing word for all of the sections above, representing the last step for products to be re-used entirely either in the same form or in a new one.

However, the purpose of this paper is to consider the recycling options for server machines that are decommissioned from data centers. Once removed, the DC equipment is tested to determine whether it still functions. Products can be reused directly as *"new"* (see subsection 7.4 or can be dismantled, and up-to-date parts are *"harvested"* and sold to spare parts providers or reused internally. They can be used to build equipment of equal or lower technological profile [16]. The complete disassembly, destruction and filtering of the product into reusable materials is in most cases done by specialized businesses, often referred to as ***recyclers***.

> Non-operating components can be repaired or refurbished (see subsubsection 7.4.1 and subsubsection 7.4.2) if it is required and/or profitable. Unusable parts are sent to recyclers, where they are treated to recover metals, incinerated with or without energy recovery, or landfilled.

Some electronic components from data centers are frequently reused; these include Hard Disk (HDD) and memory cards where reuse rates are >40%. Meanwhile Central Processing Unit (CPU) (5.2%), motherboards (2.7%) and Power Supply Unit (PSU) are reused less frequently due to the fast development of new technologies, which often optimize energy efficiency and increase storage capacity [16]. After each reuse cycle, HDDs and memory cards must go through the data sanitation phase. Data sanitization is a process through which decommissioned hardware which previously stored data is destroyed in order not to leave traces of data behind. Depending on the kind and level of data santitzation, different costs can be incurred.

$$\begin{aligned} C_{Sustainability} = {} & (C_{DataSanitizationPerGB} \times N_{GBofDataToDestroy}) \\ & + (C_{RecyclerPerKg} \times N_{KilogramsOfHardware}) \end{aligned} \tag{22}$$

# 8 Model Enhancements

The text highlights several ongoing research gaps that hinder the complete implementation of circular principles in data centers. A significant gap identified is the lack of detailed costing for the decommissioning process, crucial to assessing the financial viability of circular strategies. Current research focuses primarily on energy improvements and cooling efficiency, neglecting the detailed costs associated with maintenance. Additionally, there is a lack of clear strategies aligning hardware capability with actual requirements, leading to inefficiencies and contributing to the underexplored issue of e-waste generation. Addressing these gaps through targeted research on decommissioning costs, maintenance expenses, optimal hardware sizing, and e-waste management is crucial to advancing sustainable practices in the data center industry.

The proposed enhancements include introducing comprehensive requirements for acquiring new hardware to align with the principles of circular sustainability and customer needs. An emphasis is placed on expanding maintenance costs to prioritize repairability and refurbishment, thereby extending the useful life of the equipment. The research also investigates the environmental and economic impacts of sustainability initiatives, particularly focusing on computing hardware. A novel method is proposed to quantify ongoing decommissioning costs through various circular strategies, offering an adaptable framework for mathematical refinement. In addition, the section outlines methods and best practices for extending service life by incorporating the principles of resource reduction, recycling, and responsible disposal.

Furthermore, the model is enriched by various additions, particularly focusing on enhancing the motivation of data center actors and organizational changes necessary for acquiring and implementing new capabilities. It suggests industrial symbiosis as a means of efficient recycling, repair, or refurbishment through collaboration and resource exchange between industries. Organizational processes for capacity realization are also discussed, emphasizing the need for infrastructure maintenance teams to promote modularity, repairability, and recycling in server design.

The section concludes with the proposal of a complete Total Cost of Ownership (TCO) model, addressing all identified research gaps. The expanded TCO includes costs for sustainability, acknowledging the growing demand for hardware designed for decommissioning, and estimating expenses associated with adopting various sustainability paths such as recycling, refurbishing, or repairing existing hardware.

## 8.1 Current Research Gaps

Despite significant advances in the field of circular economy within the data center industry, several research gaps persist that hinder the complete implementation of circular principles. One primary gap is the lack of detailed costing for the decommissioning process, which is essential for assessing the financial viability of circular strategies.

| Aspect | Research Gap | Improvements |
|---|---|---|
| Total Cost of Ownership | Cost of decommissioning | Proposed a method to quantify the ongoing costs of decommissioning hardware through multiple circular strategies. Designed as an umbrella term, which can be further refined mathematically. |
| | Research Focuses on improvements in cooling of Data Centers | Performed research on the environmental and economical impact of sustainability on Data Centers while focusing on the impact of the computing hardware used. |
| | Maintenance costs are shallow | The maintenance costs have been expanded towards repair-ability and refurbishment |
| Acquisition Process | Interest is on energy use | Proposed a list of requirements for acquiring new hardware, oriented towards a circular sustainability and adjusting to the needs of customers |
| | No clear strategy for a system fit to the needs (unless the needs are on the high-end) | |
| Circular Economy | Research focuses on reusability of heat exchanged in data centers | Proposed calculation methods and practices for elongating the lifespan of servers through Re- duce / cylce / fuse |
| | No discussion on the waste generated by servers. | |

**Table 9:** Identified research gaps

Current research focuses primarily on energy improvements and cooling efficiency, leaving the costs associated with maintenance insufficiently detailed. In addition, there is no clear strategy to align hardware capability with actual requirements, often resulting in oversized and underutilized equipment. This mismatch not only leads to inefficiencies but also contributes to the underexplored issue of e-waste generated by data centers. Addressing these gaps through targeted research on decommissioning costs, detailed maintenance expenses, optimal hardware sizing, and e-waste management will be crucial for advancing a sustainable circular economy in the data center industry.

This paper proposes several improvements to advance the circular economy within the data center industry. First, it introduces a comprehensive list of requirements for acquiring new hardware, ensuring alignment with circular sustainability principles and customer needs. In addition, maintenance costs are expanded to emphasize repairability and refurbishment, facilitating a longer useful life of the equipment. The research also investigates the environmental and economic impacts of sustainability initiatives, with a specific focus on the computing hardware used in data centers. A novel method is proposed to quantify ongoing decommissioning costs through various circular strategies, designed as an adaptable framework for further mathematical refinement. Finally, the paper outlines the methods and best practices for extending service life, incorporating the principles of reducing, recycling, and discarding resources. These improvements collectively aim to improve the sustainability and efficiency of data center operations.

## 8.2 Model Additions

### 8.2.1 Capabilities

The first expansion affects the motivation of the actors. Following the ArchiMate model presented in Figure 17 (see section 4). Figure 26 displays an updated motivation flow chart that shows not only the intended results and goals of a data center manager to invest in sustainability or to consider the implications of sustainability in their daily operations. The figure links important goals and outcomes to direct actions such as *"Explore Sustainable Options"* or *"Buy scaled hardware for the needs"*, both of which are specializations of the much greater action of improving sustainability. Furthermore, all actions are then linked to existing or additional capabilities, which the data center might already have or need to acquire. The capabilities follow the five sustainable ways through which a data center can improve sustainability presented in section 7.

Furthermore, data centers need to perform organizational changes in order to be able to acquire, implement, and make use of the new capabilities. Figure 27 shows an example of organizational and inter-organizational processes and interactions that can lead to the acquisition of such capabilities.

In the case of recycling, it is unlikely that a data center business will obtain the knowledge, infrastructure, and processes to perform internal recycling of decommissioned hardware. Instead, it is more economically efficient to interact with an established Waste of Electrical and Electronic Equipment (WEEE) recycling entity to carry out recycling work. Recylcing, Repairing or Refurbishing are all acts which can be done through industrial symbiosis, which is a concept that promotes collaboration and resource exchange between industries to optimize resource use, reduce waste, and enhance environmental sustainability. In industrial symbiosis, one company's waste or by-products become another company's raw materials or inputs, creating a mutually beneficial rela-

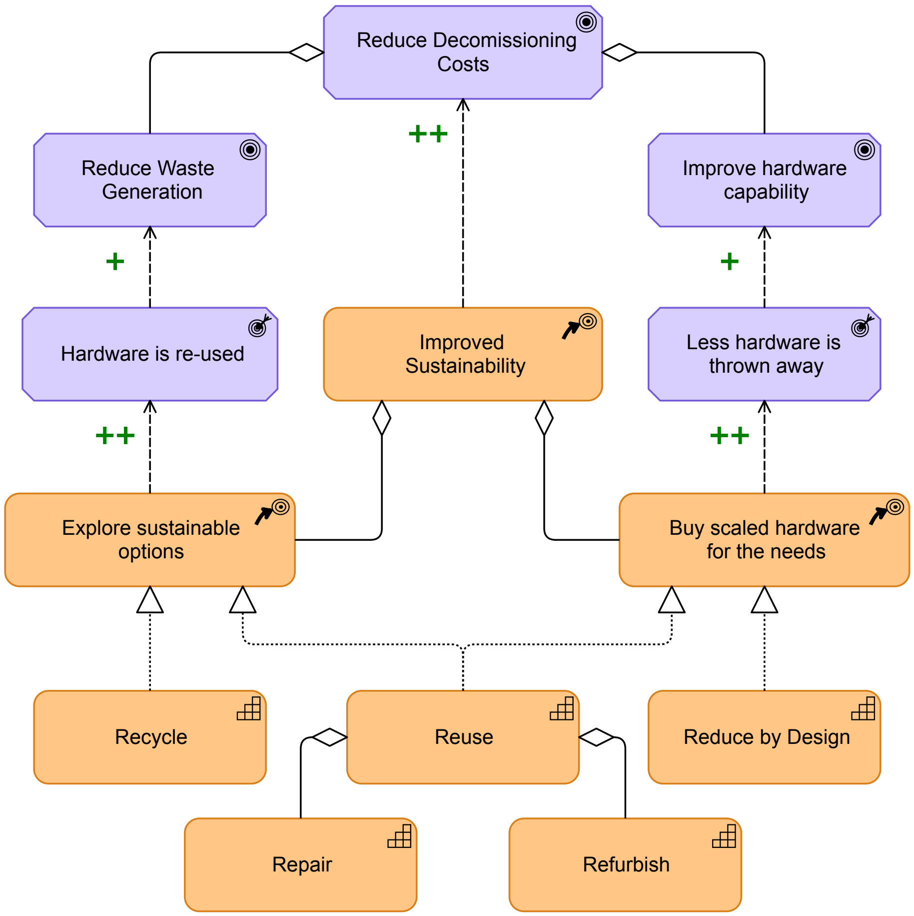



**Figure 26:** Mapping of the most important drivers onto required capabilities for a sustainable transformation

tionship.

Furthermore, the Infrastructure Maintenance Team is responsible for the maintenance of the data center equipment. Subtasks may include, but are not limited to, *Decomissioning, Acquisition, Consolidation, or Repair*. As you can imagine, each one of those subprocesses then promotes a certain capability. An interesting outcome of decommissioning is that the design of servers, server components and server mounting is changed so that by design it allows modularity, repairability, and recycling.

For example, when opting for more expensive servers that prioritize modularity and upgradeability, users are investing in infrastructure that facilitates easier component replacement and refurbishment, thus reducing the overall environmental impact of server decommissioning. These servers may incorporate design features that enable efficient disassembly, promote component reuse, and facilitate responsible recycling practices. Consequently, while the upfront cost of such servers may be higher, their long-term sus-

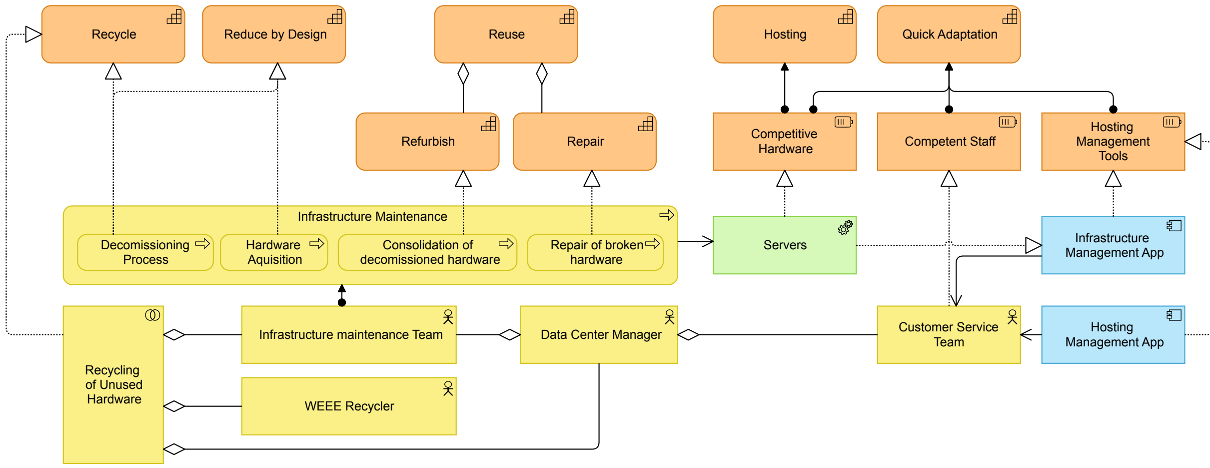



**Figure 27:** Example process for the realization of each capability introduced

tainability benefits contribute to mitigating decommissioning expenses and minimizing environmental harm.

### 8.2.2 Complete TCO Model

Considering all the arguments, critiques and additions made throughout the paper, a complete new TCO has been proposed. The proposed Total Cost of Ownership aims to answer all research gaps present in Table 9.

In addition, servers certified for reduced Critical Raw Materials (CRM) usage present another avenue for cost consideration within the ***Server Costs*** category. By selecting servers manufactured with materials that minimize the use of critical resources or eliminate hazardous substances, users can mitigate the environmental impact associated with server decommissioning. These certifications affirm compliance with stringent environmental standards and endorse sustainable manufacturing practices, aligning with the broader objectives of reducing resource consumption and promoting the principles of circular economy.

Figure 28 displays the technology Archimate model of the complete TCO (see Figure 31). In order to compute the TCO, a central object needs to be define, thus the *Data Center* facility emerges. Every other part of the model eventually flows down into the *Data Center* object. Mathematically, this means the costs are added between all relations. Further more, at implementation time, it can be decided if certain components need be left out, in turn, deleting all related elements on the specific branch. On the same level, we find the Electrical Grid, as well as a communication path as expressed in subsubsection 5.2.1 together with their related contracts. Those services then allow

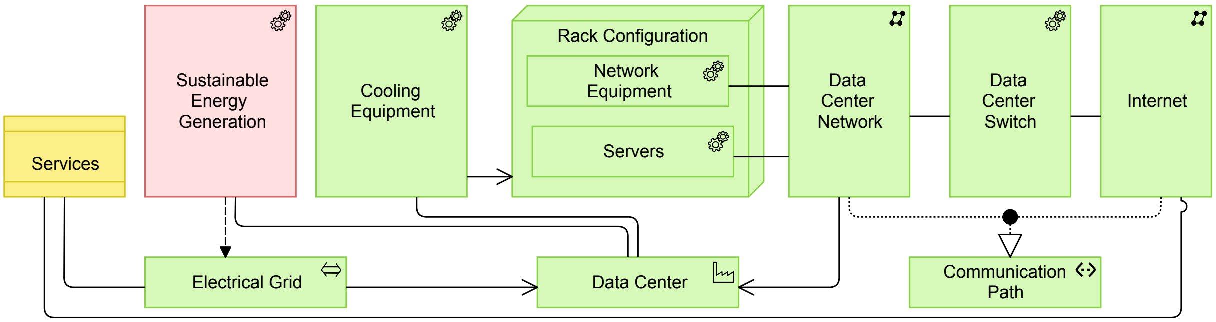



**Figure 28:** Technology part of complete TCO Model

for the interconnection of the data center to the outside world and are crucial for its existence.

The data center also requires internal infrastructure such as *Cooling, Networking* and *Computing* equipment. The network infrastructure consists of the Internet and the internal network created within the data center between multiple computing nodes and switches. Both the cooling and the networking equipment are directly connected to the rack configuration, more specifically the servers that are designed to be modular and easy to disassemble (see subsubsection 7.4.1 and subsection 6.2).

> You can notice the equipment named *"Sustainable Energy Generation"* added to the TCO. Although this paper does not explicitly talk or suggest changes in regard to power generation in data centers, it acknowledges the need for greener energy in data centers and an improvement towards generating the required electrical energy within the limits of the data center. Examples in this case can be: Solar arrays, wind turbines, and geothermal energy. [43]

In addition to server, network, and cooling equipment, there exists a management interface, which allows stakeholders to control and manage each equipment individually. Moreover, due to the increased performance of servers, it is unlikely that the equipment is running as bear bones but rather equipped with a hypervisor capable of supporting multiple virtual machines. Here, the model makes a distinction between ***Free and Open-Source Software (FOSS)***, and proprietary Operating Systems. The distinction is necessary so as not to omit the licensing costs associated with proprietary software.

> The Archimate model, as presented in Figure 29 and Figure 31, showcases only one instance of each object. However, it must be noted that there can be multiple variations to the configuration of servers, the number of hypervisors used, or the number of Virtual Machines. In addition, data centers may have multiple connections both to the electricity grid

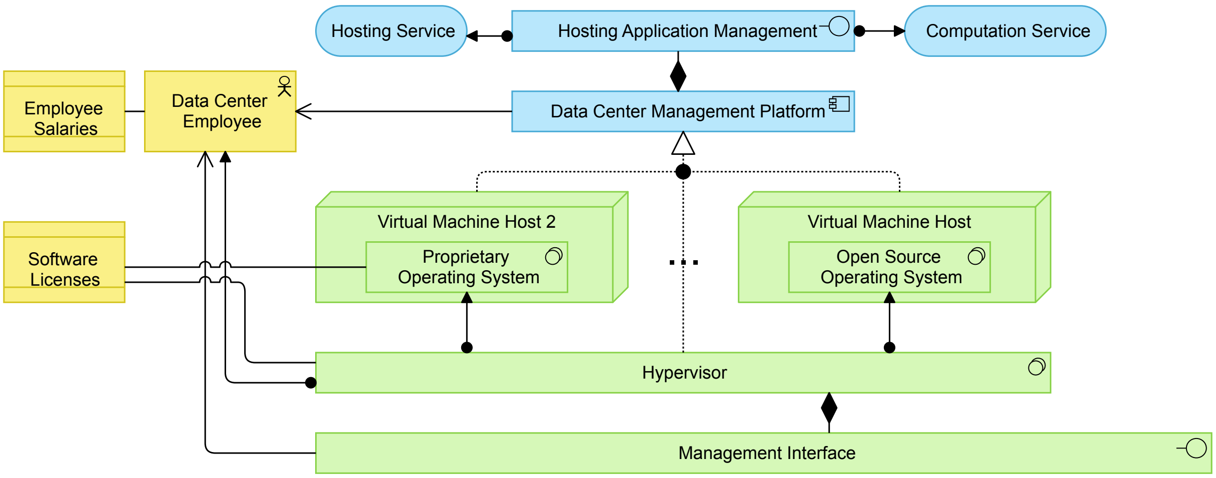



**Figure 29:** Application part of complete TCO Model

> and to the Internet to allow redundancy. So, feel free to add as many instances as you see fit.

All of the technology layers (marked green in the model) serve the ultimate revenue drivers, such as services. For illustration purposes, the model proposes a distinction between hosting services and computing services.

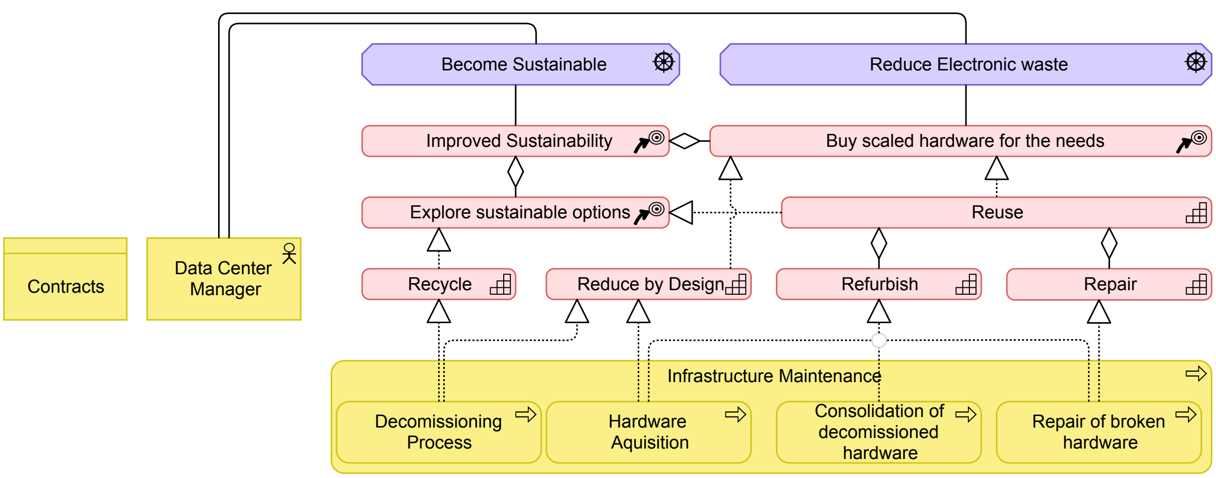



**Figure 30:** Business part of complete TCO Model

Figure 30 illustrates the business actors, processes and motivation to invest into sustainable technologies in the data center environment. While data centers are complete organizations, often with multiple departments, the model chooses to define two actors and stakeholders. The *Data Center Manager* which is a person or a board of managers which take decisions and actions in the best interest of the data center. *The Data Center Employee* is a generic employee representing System Administrators or Data Center

Technicians which take responsibility in maintaining the internal (and external) infrastructure of the data center.

Employees are bound to the organization by their employment contract, which defines their salary and responsibilities. In addition, employees must carry out the general maintenance process. The maintenance process includes repairing and ensuring the well functioning of the data center equipment. As an addition in the scope of sustainability, the paper adds three new sub-processes: *Decommissioning*, *Hardware Acquisition*, *Consolidation of decommissioned hardware*. All of which are linked to respective capabilities that the data center can ensure.

The last aspect of the business perspective is related to the motivation of the actors to make such changes in the data center. In this case, we identify the need to become sustainable and reduce electronic waste as the primary drivers of change for the Data Center Manager. As explained in subsection 4.1, related to the drivers are the action points for **exploring sustainable options** and to **purchasing scaled hardware**, both of which are specializations of the more general aspect of *improving sustainability*. The aspects of reusing, recycling, refurbishing, and repairing, defined in section 7, are then the capabilities that link the business processes to the action objectives of the managers.

Lastly, Figure 31 combines Figure 28, Figure 29 and Figure 30 together, resulting in the complete Archimate model of the Total Cost of Ownership.
Following the example from Hardy's 2013 Total Cost of Ownership (TCO) [34], the new Total Cost of Ownership can also be expressed mathematically, by the following formula:

$$\begin{aligned} TCO = {} & C_{Server} + C_{Network} + C_{Infrastructure} \\ & + C_{Power} + C_{Maintenance} + C_{Sustainability} \end{aligned} \tag{23}$$

Where the cost of the server ($C_{Server}$) has been expanded to acknowledge the growing demand to invest in hardware designed for decommissioning, with servers tending to cost more upfront than normal servers. In addition, the cost of maintenance ($C_{Maintenance}$)has been expanded to include estimates for hardware replacement as well as software licenses, employee salaries, and services such as connection to the electrical grid or the network.

The new addition made is that of a cost for sustainability ($C_{Sustainability}$) which aims to encompass the costs associated with adopting various sustainability paths such as Recycling, Refurbishing or Repairing of existing hardware.

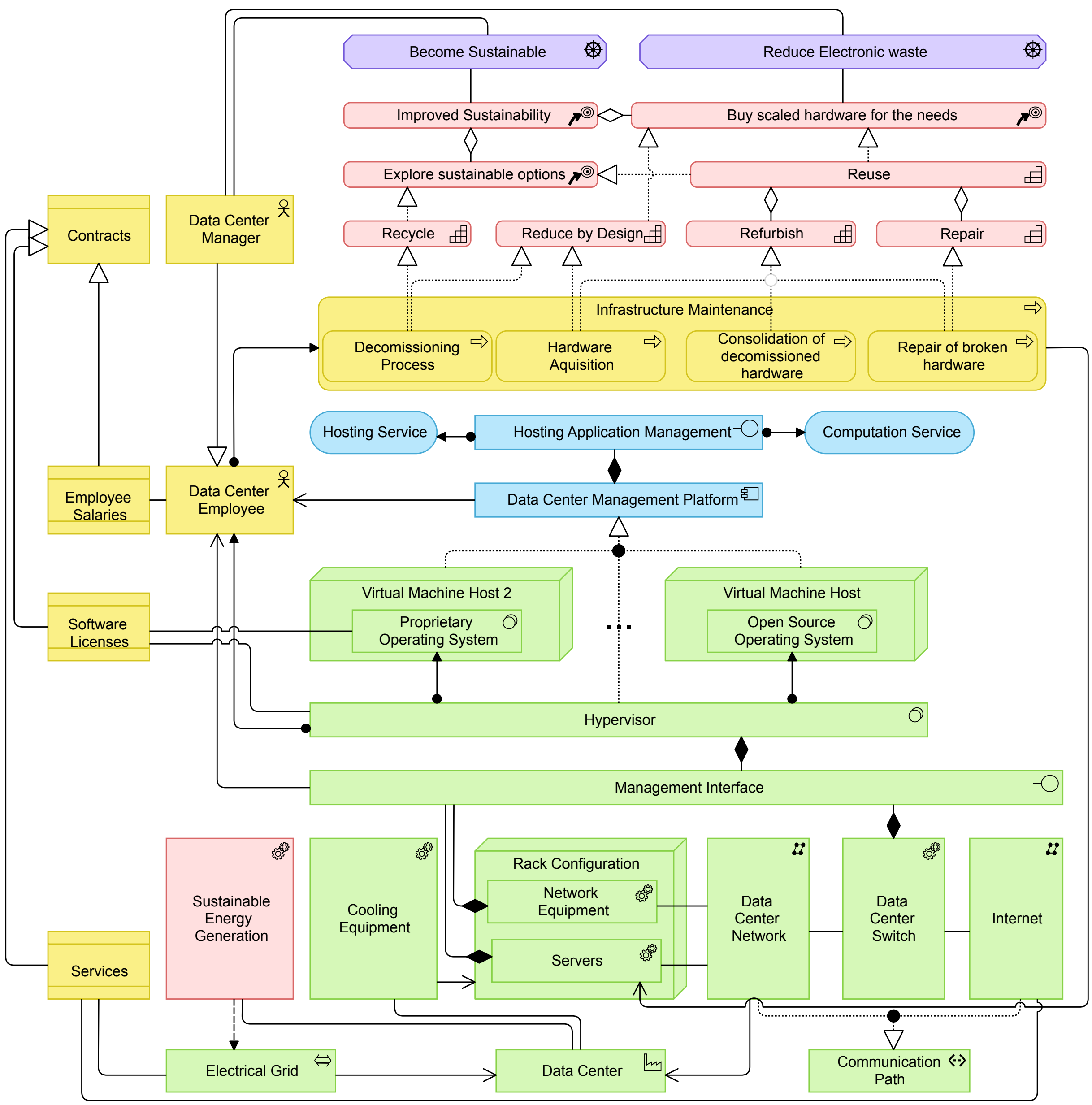



**Figure 31:** Complete top-down TCO model

# 9 Case-Studies

In this section, we apply our developed Total Cost of Ownership (TCO) mathematical model including sustainable considerations, to a real-world scenario. The purpose of this case study is to illustrate the practical application and benefits of incorporating sustainability into daily operations. By examining a specific industry and company, our aim is to demonstrate how our model can provide a more comprehensive understanding of the true costs associated with their ownership, including environmental impacts, beyond traditional financial metrics.

The selected case study focuses on a fictive entity within the data center industry which is known for its substantial environmental footprint and complex supply chain dynamics. This industry presents an ideal case for evaluating the efficacy of our TCO model due to its high relevance to sustainability challenges and the potential for significant cost savings and impact reduction through improved practices.

We will begin by outlining the specific characteristics and operational context of the company, followed by a detailed application of the TCO model. Key parameters, assumptions, and sustainability metrics will be highlighted, illustrating how each component influences the overall cost and sustainability profile. The results will then be analyzed to obtain insight into cost savings, efficiency improvements, and sustainability enhancements achievable through informed decision making.

Ultimately, this case study aims to validate the applicability and robustness of the model, providing a blueprint for other organizations seeking to integrate sustainability into their TCO assessments. Through this practical demonstration, we emphasize the critical importance of a holistic approach to cost analysis that encompasses economic, environmental, and social dimensions, paving the way for more sustainable business practices.

## 9.1 Model Validation

$$TCO = C_{server} + C_{network} + C_{infrastructure} + C_{power} + C_{maintenance} + C_{sustainability} \tag{24}$$

Model validation is defined within regulatory guidance as the set of processes and activities intended to verify that models are performing as expected, in line with their design objectives and business uses. It also identifies *potential limitations and assumptions, and assesses their possible impact*. With that in mind, the Total Cost of Ownership (TCO) model (see section 8 and Equation 24) needs to define the validation steps. Being a financial model, each variable is subject to the following set of rules:

$$C_{\text{server}} \geq 0$$
$$C_{\text{network}} \geq 0$$
$$C_{\text{infrastructure}} \geq 0$$

$$C_{\text{power}} \geq 0$$
$$C_{\text{maintenance}} \geq 0$$
$$C_{\text{sustainability}} \geq -C_{\text{incentives}}$$

The following bullet point aims at providing an explanation of each component and the reasoning of its rule:

- **Cost Server ($C_{\text{server}}$):** This includes the purchase, lease, or depreciation costs of servers. These costs are always non-negative as they represent actual expenditures incurred in acquiring or maintaining server hardware.

- **Cost Network ($C_{\text{network}}$):** This includes expenses related to network equipment, bandwidth, and connectivity services. These costs are inherently non-negative since they involve real outlays for network infrastructure and services.

- **Cost Infrastructure ($C_{\text{infrastructure}}$):** This covers the physical facilities and utilities required to support servers and network equipment, such as data center space, cooling, and security. These costs are also non-negative as they represent essential operational expenditures.

- **Cost Power ($C_{\text{power}}$):** This includes the electricity required to operate servers, network equipment, and cooling systems. These costs cannot be negative as they are based on the consumption of electrical power and associated utility bills.

- **Cost Maintenance ($C_{\text{maintenance}}$):** This involves regular upkeep, repairs, and support services for servers, network equipment, and infrastructure. These costs are always non-negative because they represent necessary expenditures to ensure continued operation and reliability.

- **Cost Sustainability ($C_{\text{sustainability}}$):** This is unique in that it can encompass both expenses and potential revenue streams. It includes costs related to implementing sustainable practices, such as energy-efficient technologies or waste reduction measures. However, it can also include revenues from sustainability incentives, rebates, or credits (e.g., renewable energy credits). As such, the sustainability cost can be negative, reflecting a net gain if sustainability incentives exceed expenditures.

## 9.2 Model Evaluation

> Prices used in the model evaluation section are taken as averages from major providers such as Dell, HP, SuperMicro, Ubiquity, Netgear by accessing their web stores and picking the most popular product on offer.

> Furthermore, ongoing costs are calculated with a cumulative 2% interest rate,dar with the intention of simulating real-world scenarios.

A small company is evaluating its Total Cost of Ownership (TCO) over a 5-year period. The company management decides to rent rack space at a third party location. The company decides that their requirements are for seven warm hypervisor machines. Furthermore, the company considers 10% of the number of nodes as cold spare nodes, therefore, one cold spare machine. The company decides that they will rent-out a complete rack and that for their use case there is a need for three switches in order to offer the required redundancy. Table 10 shows the average prices used for all calculations in the following scenarios.

| Variable | Case 1 | Case 2 & 3 |
|---|---|---|
| $C_{Infrastructure}$ | \$12,000 | \$15,000 |
| $C_{Server}$ | \$10,000 | \$12,000 |
| $C_{NetEquipment}$ | \$1,500 | \$1,750 |
| $C_{ElectricityPerKW}$ | \$0.29 | \$0.20 |
| $K_{ServerConfig}$ | 1 | 1 |
| $N_{Server}$ | 8 | 6 |
| $N_{ColdSpare}$ | 1 | 1 |
| $N_{NetEquipment}$ | 3 | 2 |
| $P_{ServerEquipment}$ | 2000 W | 1500 W |
| $P_{NetEquipment}$ | 300 W | 250 W |
| $PUE_{DataCenter}$ | 1.6% | 1.56% |
| $SPUE$ | 1.5% | 1.48% |

| Variable | Case 1 | Case 2 & 3 |
|---|---|---|
| $C_{SalaryPerYear}$ | \$42,000 | \$42,000 |
| $C_{GridConnection}$ | \$0 | \$0 |
| $C_{InternetConnection}$ | \$1,200 | \$1,200 |
| $C_{SoftwareLicensing}$ | \$47,000 | \$47,000 |
| $C_{Repairs}$ | 0 | TBD |
| $N_{Employees}$ | 1 | 1 |
| $C_{SanitizationPerGB}$ | \$0 | \$0.007 |
| $C_{RecyclerPerKg}$ | \$0 | \$5 |
| $N_{DataToDestroy}$ | 0 GB | 50 TB |
| $N_{HardwareWeight}$ | 0 Kg | 350 Kg |
| $C_{DesignPlanning}$ | \$0 | \$1,500 |
| $O_{Selling}$ | 40% | 60% |
| $O_{EstSellingPrice}$ | \$5,000 | \$7,500 |

**Table 10:** Values of variables used for calculations

### 9.2.1 Case 1: Ignoring Sustainability

The first scenario shows a case where the company decides to ignore sustainability measures in their small data center. Furthermore, this scenario does not account for repairs made to the hardware after acquisition. Below are the given and estimated costs for each category, excluding sustainability costs, all costs are brought up to the determined 5-year period. The total costs for the first year of running the data center are presented in Table 11. The table also includes a yearly increase in the costs of 2% due to inflation. The static costs are composed of the server costs, the network costs, and the Oracle license.

| Cost | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | **Total** |
|---|---|---|---|---|---|---|
| Static Costs | $137,000 | $0 | $0 | $0 | $0 | **$137,000** |
| Operational Costs | $98,132 | $100,094 | $102,096 | $104,138 | $106,221 | **$512,681** |
| Total | **$235,132** | **$100,094** | **$102,096** | **$104,138** | **$106,221** | ***$649,681*** |

**Table 11:** Running costs for scenario 1

| Cost | Sub-costs | Costs per Year | Costs per Period | Explanation |
|---|---|---|---|---|
| Server ($C_{Server}$) | Initial purchase | $80,000 | $80,000 | The initial overall cost of the server hardware. According to Super-Micro's website, the average price per unit is $5,000 depending on the configuration. The Power Usage Efficiency (PUE) of chosen servers is 1.5 |
| Networking ($C_{Network}$) | Initial Setup | $10,000 | $10,000 | The initial investment into networking equipment. Considered the average price for 3 Ubiquity Enterprise switches with added leeway for installation costs and miscellaneous. |
| Infrastructure ($C_{Infrastructure}$) | Rent | $12,000 | $62,470 | Because the company decides to rent rack space, the costs include the rent and the security of the data center, considering that most data rack lenders include basic security services in their package. |
| Power ($C_{Power}$) | Cost of electricity | $42,932 | $223,417 | The cost of electricity for one year of running the servers 24/7. The cost per month evaluates to $3,000, meaning it is ~$36,000 per year. |
| Maintenance ($C_{Maintenance}$) | Grid Connection | $0 | $0 | There is no cost of connecting to the grid as this is done through the 3rd party vendor |
| | Network Connection | $1,200 | $6,245 | A dedicated Internet connection with an internal VLAN with an average price of $100 per month |
| | Salary costs | $42,000 | $218,648 | The salary of a full-time system administrator |
| | Licensing Costs | $47,000 | $47,000 | Cost of running one Oracle Enterprise Edition Database |
| TOTAL | | $235,132 | $647,780 | |

**Table 12:** Scenario 1: Example of Costing for a 5 year period with no sustainability measures taken

### 9.2.2 Case 2: Building a Sustainable Data Center

The small business decides to approach a rack rental data center capable of offering sustainable options. One of the options offered by the third-party data center is the green energy provided by the installation of solar panels. This in turn affects the price of energy from ***$0.29 per KWh*** to ***$0.2 per KWh***. With such a measure, the business tries to profit not only from the cost savings from reduced electricity consumption but also from potential renewable energy credits that could offset expenditures, such as tax reliefs.

| Cost | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | **Total** |
|---|---|---|---|---|---|---|
| Static Costs | $126,500 | $0 | $0 | $0 | $0 | **$126,500** |
| Operational Costs | $83,087 | $84,748 | $86,443 | $88,172 | $89,936 | **$432,386** |
| Total | **$209,587** | **$84,748** | **$86,443** | **$88,172** | **$89.936** | ***$558,886*** |

**Table 13:** Running costs for scenario 2

Furthermore, the small company decides to invest in hardware that has been designed for disassembly. This requires the company to go through a planning phase that costs the company $1 500, while the server costs increase by 20%, resulting in the cost of a server being as $12 000. Moreover, from the analysis, the company decides it's business processes only require six servers and one warm spare. Furthermore, as a result of the planning, the company decides to also invest in more power efficient switching gear, reducing power consumption from 300W to 250W, however, the new switching gear now costs $250 more. Due to the reduced number of servers, they argue internally that they can achieve a good enough level of redundancy with only two switches instead of three. Furthermore, improvements to the hardware also result in slightly increased Power Usage Efficiency (PUE) and Server Power Usage Efficiency (SPUE) ratings. Table 13 shows the running costs for the second scenario.

| Cost | Sub-costs | Costs per Year | Costs per Period | Explanation |
|---|---|---|---|---|
| Server ($C_{Server}$) | Initial purchase | $70,000 | $70,000 | The initial purchase cost of the server hardware. |
| | Initial Planning | $1,500 | $1,500 | Research and market analysis for hardware oriented to disassembly |
| Networking ($C_{Network}$) | Initial Setup | $8,000 | $8,000 | The initial investment in networking equipment. |
| Infrastructure ($C_{Infrastructure}$) | Rent | $15,000 | $78,087 | Because the company decides to rent out rack space, the costs include the rent and the security of the data center, considering that most data rack lenders include basic security services in their package. |
| Power ($C_{Power}$) | Cost of electricity | $24,887 | $129,511 | The cost of electricity for one year of running the server 24/7. The cost per month evalutes to $1,388, meaning it is ~$16,655 per year. |
| Maintenance ($C_{Maintenance}$) | Grid Connection | $0 | $0 | There is no cost of connecting to the grid as this is done through the 3rd party vendor |
| | Network Connection | $1,200 | $6,245 | A dedicated Internet connection with an internal Virtual Local Area Network (VLAN) with an average price of $100 per month |
| | Salary costs | $42.000 | $218,648 | The salary of a full-time system administrator |
| | Licensing Costs | $47,000 | $47,000 | Cost of running one Oracle Enterprise Edition Database |
| TOTAL | | $209,587 | $558,991 | |

**Table 14:** Scenario 2: Example of Costing for a 5 year period with adequate computing power

### 9.2.3 Case 3: A Circular Economy Decommissioning Strategy

In the last scenario, the small company, which has been operating for four years, finds that its servers are beginning to show signs of aging. Recognizing the need to maintain efficient and cost-effective operations, the company decides to adopt more sustainable practices. This decision is driven by the dual objectives of reducing operational costs and improving the profitability of their services. In addition, they understand that their customer base values sustainability, and they see an opportunity to strengthen their market position by highlighting their commitment to environmentally friendly practices. A table of the running costs can be view

| Cost | Year 1 | Year 2 | Year 3 | Year 4 | Year 5 | **Total** |
|---|---|---|---|---|---|---|
| Static Costs | $126,500 | $0 | $0 | $51,000 | $0 | **$177,500** |
| Operational Costs | $83,087 | $84,748 | $86,443 | $57,475 | $84,625 | **$396,378** |
| Total | **$209,587** | **$84,748** | **$88,172** | **$108,475** | **$84,625** | ***$578,878*** |

**Table 15:** Running costs for scenario 3

To achieve these goals, the company formulates three distinct strategies to manage their aging hardware:

I ***Use Refurbished Hardware*** — Instead of investing in brand-new servers, the company opts to purchase high-quality refurbished equipment. This hardware is sourced from certified resellers who guarantee its reliability and performance. By choosing refurbished servers, the company can significantly lower its capital expenditure while still maintaining the necessary technological standards.

II ***Re-sell old hardware Hardware*** — As part of their strategy to manage outdated equipment, the company decides to sell any surplus hardware to third-party resellers. This approach not only helps in recouping some of the initial investment but also prevents functional equipment from ending up in landfills. By re-selling, the company contributes to the circular economy and minimizes waste.

III ***Recycle Existing Hardware*** — For hardware that can no longer be used or resold, the company partners with a certified recycling facility. This ensures that valuable materials are recovered and repurposed, reducing the environmental impact associated with electronic waste. By engaging in responsible recycling, the company adheres to environmental regulations and demonstrates its commitment to sustainability.

Each of these options is carefully evaluated to determine the most effective and sustainable approach. By comparing the costs and benefits of each strategy, the company aims to make informed decisions that align with their financial goals and sustainability commitments. Through these efforts, they hope not only to reduce operational costs but also to enhance their reputation as a responsible and forward-thinking business, ultimately leading to greater customer satisfaction and loyalty.

Furthermore, for this case, the following variables are introduced:

| Variable | Case 1 |
|---|---|
| $C_{GarbageCollection}$ | \$200 |
| $C_{Server}$ | \$8,500 |
| $P_{ServerEquipment}$ | 1400 W |

**Table 16:** Extra variables required for Scenario 3

> All of the above can be combined to accentuate the sustainability effect of the company's actions. In a real scenario, it is very likely that the reseller for refurbished hardware would also accept the old equipment as a reimbursement.
>
> Additionally, the company could also sell the hardware on the second-hand market itself.

When considering selling. The company decides to sell all six server machines. They estimated that each machine could be sold for a price of \$7,500, which would result in a net gain of \$45,000. However, second-hand sales do not always go to plan and there is the possibility that the company cannot sell all the hardware. However, they believe with 60% certainty that they can sell all the hardware. Including this certainty, it results in a *guaranteed* gain of \$27,000.

In the calculation shown in Table 17, the company is assumed to have sold 60% of the hardware, which means that they have sold 4 machines. The other 2 machines have been sent for recycling. Nothing is thrown away as garbage; however, the costs are included for illustration purposes.

| Cost | Sub-costs | Costs per Year | Costs per Period | Explanation |
|---|---|---|---|---|
| Server ($C_{Server}$) | Initial purchase | $70,000 | $70,000 | The initial purchase cost of the server hardware. |
| | Initial Planning | $1,500 | $1,500 | Research and market analysis for hardware oriented to disassembly |
| | Refurbished Hardware | $51,000 | $51,000 | Purchase of refurbished hardware. |
| Networking ($C_{Network}$) | Initial Setup | $8,000 | $8,000 | The initial investment in networking equipment. |
| Infrastructure ($C_{Infrastructure}$) | Rent | $15,000 | $78,087 | Renting costs to third party provider. |
| Power ($C_{Power}$) | Cost of electricity | $24,887 $22,863 | 125,436 | The first year costs $24,887 and accumulates 4 years of *interest*. Meanwhile the 5th year is calculated for the new hardware |
| Maintenance ($C_{Maintenance}$) | Network Connection | $1,200 | $6,245 | A dedicated Internet connection with an internal Virtual Local Area Network (VLAN) with an average price of $100 per month |
| | Salary costs | $42.000 | $218,648 | The salary of a full-time system administrator |
| | Licensing Costs | $47,000 | $47,000 | Cost of running one Oracle Enterprise Edition Database |
| | Data Sanitization | $350 | $350 | The cost to sanitize data before decommissioning |
| Sustainability ($C_{Sustainability}$) | Re-sell Value | $-27,000 | $-27,000 | The value at which the company negative,to sell the 4 year old hardware. |
| | Recycling value | $-500 | $-500 | The value that the company would be reimbursed for in the case of hardware recycling. |
| | Throw to garbage | $70,000 | $70,000 | The cost of throwing the servers as garbage. Resulting in no sustainable gains. |
| TOTAL | | $209,587 | $578,766 | |

**Table 17:** Scenario 2: Example of Costing for a 5-year period with hardware change during the fourth year

## 9.3 Key Takeaways

Cases one through three show the numbers and calculations of building a sustainable data center. However, numbers alone do not tell the complete story. The differences between the first and second cases illustrate that a conscious business investing in sustainable hardware can significantly outperform those not aiming for sustainability. For example, the total cost of ownership for Case 1 was $649,681, while for Case 2, the total cost was only $558,886. The stark difference arises from the nearly $15,000 discrepancy in operating costs between a sustainable data center and its less eco-friendly counterpart. This highlights the financial benefits of sustainable investments, such as reduced energy consumption and lower long-term maintenance costs, which contribute to overall savings.

Furthermore, when compared to the $578,878 required to completely build and replace hardware in Case 3, both previous scenarios seem relatively modest. There is a $20,000 difference between Case 2 and Case 3, with the second case being the less expensive option. However, it is essential to note that in the third case, the small business manages to completely refurbish its hardware catalog within the five-year study period. Additionally, Case 3 is the only scenario that incorporates decommissioning techniques, which, as mentioned previously, should be considered as additional costs. This case demonstrates the potential for long-term sustainability through regular updates and maintenance, ensuring that the infrastructure remains up-to-date and efficient while managing the lifecycle of the hardware responsibly.

# 10 Conclusion

In this study, we addressed the complexities of modern data center operations through a comprehensive exploration of sustainability practices and the provision of a circular economy model within this sector. To this end, our primary objective was to develop and expand an existing Total Cost of Ownership (TCO) model to incorporate the costs associated with hardware decommissioning. This model provides a holistic, more accurate view of the financial impacts of data center life-cycle, from acquisition, through operation, to end-of-life management.

We defined our model following a 4-step approach, guided by 4 research subquestions. In the following paragraphs, we revisit these questions and the answers our research has provided. Through these research questions, we offer, to the best of our knowledge, the most comprehensive and current version of Hardy et al's Total Cost of Ownership (TCO) model. This study conceptualizes the broad issue of sustainability in data centers and explores how related concepts can be utilized in detail. Consequently, this work is valuable to practitioners and serves as a foundation for future research.

**RQ1: What are the key challenges and benefits of implementing circular economy principles in data center operations?** We provided a first overview of these challenges and benefits by conducting a literature review (section 2). We found that while there is a growing recognition of the importance of sustainability, significant gaps remain in the practical implementation of the principles of circular economy. Our study further highlighted the difficulties of integrating the principles of a circular economy, such as the need for better decommissioning processes and the financial implications of sustainable practices, into operational data center models. However, as the benefits include potential cost savings and improved environmental performance, addressing these challenges can lead to sustainable practices that are beneficial to data center operators. In addition, we expanded on previous research by highlighting current research areas in the data center industry. Thus, we identified gaps in the current literature, such as the lack of financial costing with respect to sustainability. Furthermore, interviews (section 3) conducted with industry stakeholders confirmed the gaps identified in the literature study.

**RQ2: How do maintenance costs, including service costs, software licensing, and repair costs, contribute to the overall TCO?** Section 4 and section 5 both answer our second research question by identifying the most comprehensive TCO model and the gaps in this model. We expanded it further to address maintenance costs, service costs, and software licensing. We proposed these expansions, providing argumentation for each of the additions from the new model. Analysis of maintenance costs revealed that including repairability and refurbishment in the TCO model can significantly extend the life of hardware and reduce overall costs. This aligns with the findings of the interviews in which the stakeholders emphasized the importance of efficient maintenance practices.

**RQ3: What are good practices for hardware acquisition that align with sustainability goals?** Section 6 answers this third research question. We identified good practices for hardware acquisition, highlighting the need for thorough requirement elicitation and alignment with sustainability objectives. This is crucial to ensure that new hardware investments are both cost-effective and environmentally friendly. We also provided links and connections to the existing Sustainable Development Goals (SDG) where investing in sustainable hardware practices can have significant impact.

**RQ4: How do different sustainability practices (reduce, reuse, recycle) affect the operational costs and efficiency of data centers?** We answered RQ4 by section 7 where the impact of various sustainability practices on operational efficiency was evaluated through case studies (section 9). It was found that practices such as recycling and refurbishing not only reduce environmental impact but also offer economic benefits by lowering operational costs.

In summary, following this 4-stages approach, our work addresses the main gaps we identified in the relationship between the data center industry and sustainability actions. We do so by providing an updated and working model[1] that can ease scenario analysis and thus can provide more data for an informed decision towards adopting sustainable practices in data center operation. Thus, we ultimately answered our main research question - *"What Total Cost of Ownership model can be employed to evaluate sustainable growth strategies within data center environments?"* - by providing a better TCO model to help organizations implement circular economy practices for datac enters.

## 10.1 Contributions and Challenges

This research provides significant contributions to the field of sustainable data center (DC) management by addressing several critical challenges and proposing comprehensive solutions:

- We define a circular economy model in the context of data centers.
- We provide a practical implementation of our model that can be used for scenario analysis.
- We provide a comprehensive questionnaire for collecting data center sustainability practices, and present the state-of-practice in the industry based on 8 responses to this questionnaire.

[1]`https://gitfront.io/r/gorgonea/wmvfe3LSugWE/Sustainable-TCO/`

- We propose an improved, operational, quantitative TCO model that includes sustainability metrics and can be used to assess the impact of the proposed circular economy.
- We propose a scenario-based validation method and present three scenarios aligned with three realistic case-studies.

In the following paragraphs, we highlight the relevance and main challenges surrounding these contributions.
Our main contributions are the definition of a circular economy in the context of data centers and the definition of our novel data center Total Cost of Ownership (TCO) model that includes these circular economy provisions towards sustainability. The main challenge we encountered in integrating sustainability into data center operations was the ambiguity in definitions and the gap between theoretical concepts and practical implementations. For example, the 9R model's *refusal* is a clear concept, but it cannot always be achieved in practice.

On a more practical note, our improved TCO model required aligning sustainability metrics and models alongside traditional operational metrics and financial costs. This alignment required various types of normalization, which in turn imposed the refinement of existing approaches/tools and the transformation of metrics present in the literature. For example, making sure that every value is correctly normalized over the same duration of time, such as one month and not one year, was an important aspect in defining a workable, quantitative TCO model. Further, to aid in understanding the complex relationships between costs, an *Archimate* [41] model has been employed, showcasing the entirety of the TCO model in the form of business, application, and technology layers.

We also contribute a scenario-based validation approach for our TCO model and propose three scenarios, inspired by three different case studies. These case studies demonstrate the model's effectiveness in real-world scenarios, but they have not been straightforward to create: finding feasible and realistic data for an accurate estimation of their parameters is based on integrated information from literature, vendors and integrator price lists, various service companies offers, and real-estate costs.

Furthermore, we proposed a comprehensive form, inspired by our literature study, to conduct a campaign of interviews with stakeholders in the data center industry. By conducting these interviews with industry stakeholders and analyzing current practices, our study sheds light on the state of sustainability practices in real data centers. This step was necessary because it was unknown how real practices in a data center are aligned with those presented in academic papers. The designing of the questionnaire and the definition of the right questions for the interviews to keep them short, yet informative was an additional challenge.

## 10.2 Practical Implications

This research contributes significantly to the academic discourse on sustainable data center management by presenting an advanced framework to evaluate and improve sustainability. Our improved Total Cost of Ownership (TCO) model, integrating environmental and social costs alongside traditional financial metrics, emphasizes the critical role of sustainability in data center operations, addressing the entire life cycle of data center assets, from acquisition to end-of-life management.

The new TCO model is comprehensive and incorporates the costs associated with infrastructure, servers, networking, power, maintenance, and sustainability initiatives. Provides a more accurate and holistic view of the financial impacts of data center operations, considering factors such as energy-efficient technologies, waste reduction measures, and potential revenues from sustainability incentives. This nuanced approach allows practitioners to balance technological advancement with environmental responsibility, ensuring that sustainable practices do not compromise profitability.

Furthermore, the model highlights the importance of modularity and upgradeability in server design, which facilitates easier component replacement and refurbishment, thus reducing the environmental impacts associated with server decommissioning. By selecting servers with certifications for reduced Critical Raw Materials (CRM) usage, data centers can mitigate their environmental footprint and align with the principles of circular economy, promoting resource conservation and waste minimization.

In practical terms, we applied this enhanced TCO model to real-world scenarios through three detailed case studies. These case studies illustrate the financial benefits of sustainable investments, such as reduced energy consumption and lower long-term maintenance costs, which contribute to overall savings. The findings demonstrate how integrating sustainability into TCO assessments can lead to significant cost savings, efficiency improvements, and enhanced sustainability profiles for data centers.

In summary, this study offers a robust framework for industry professionals looking to improve sustainability while maintaining profitability. It paves the way for more sustainable business practices by providing actionable insights and practical solutions to integrate environmental and social considerations into data center management.

## 10.3 Limitations

Despite the comprehensive nature of this study, we identify three main limitations.

The first limitation is inherent to the model itself: the Archimate model [41], by its very nature, cannot be entirely complete. Although it provides a broad perspective on the interconnections of elements within the business, application, and technology landscape of sustainable data centers, it fails to capture the specific nuances and solutions offered by individual vendors. This can lead to gaps in understanding the full scope of sustainable practices in different technological ecosystems. A second limitation is linked to the study's reliance on interviews conducted within a specific geographic region, coupled with the limited number of interviews, which poses a challenge to the generalizability of the findings. However, while the insights gathered may not fully represent the diverse practices and challenges faced by data centers in other regions or contexts, the approach of collecting such information - the questionnaire and interviews - is portable across these regions and contexts.

Third, and final: our work represents a new state-of-the-art in modeling a circular economy in the data center field. However, we acknowledge a critical limitation in assessing the timeliness of the findings. Throughout the study, a multitude of new papers and developments emerged in the realms of sustainability and the data center industry, possibly exceeding the insights presented here [18, 11, 10, 5]. Although our study aims to offer up-to-date and pertinent information, rapid advancements in this dynamic field require additional, periodic checks to determine whether the assumptions made and the state-of-practice changes (including due to the partial adoption of this model) cause some contributions and/or conclusions to become outdated.

## 10.4 Future Work

Future research should aim to validate the enhanced Total Cost of Ownership (TCO) model across diverse contexts and geographic regions to ensure its broader applicability and robustness. This includes testing the model in various types of data centers, from small-scale operations to large-scale cloud facilities, to understand how different environments impact the model's effectiveness and accuracy.

Furthermore, further exploration of additional sustainability strategies is essential to improve the circular economy within data centers. This could involve investigating innovative approaches to energy efficiency, such as advanced cooling techniques, the use of renewable energy sources, and the integration of AI-driven energy management systems. Research should also further investigate modular and upgradeable server designs that facilitate easier component replacement and refurbishment, thereby extending the lifecycle of data center hardware and reducing electronic waste.

Expanding the model to include more Sustainable Development Goals (SDG) is another critical area for future work. Identifying and implementing strategies that align with the SDGs, such as affordable and clean energy (SDG 7), industry innovation and infrastructure (SDG 9), and responsible consumption and production (SDG 12), will be vital in promoting sustainable data center practices.

Moreover, it is important to consider **SDG 11: Sustainable Cities and Communities**. Data centers, often located in urban areas, can significantly impact local environments and communities. Future research should explore how sustainable data center practices can contribute to making cities and human settlements more inclusive, safe, resilient, and sustainable. This includes examining the integration of data centers with local energy grids, utilizing waste heat for district heating, and ensuring minimal levels of noise and pollution. In addition, **SDG 6: Clean Water and Sanitation** should be taken into account. Data centers require substantial cooling, often involving significant water usage. Future work should investigate water-efficient cooling technologies and strategies to minimize water consumption, such as using recycled water for cooling purposes. Ensuring that data centers do not adversely affect local water resources and contribute to clean water initiatives is crucial to holistic sustainability.

Furthermore, more empirical testing of the framework with larger studies is necessary to enable its practical applications. Longitudinal studies tracking the financial and environmental impacts of sustainable practices over extended periods can provide a deeper understanding of the benefits and challenges of implementing the enhanced TCO model.

Finally, fostering collaboration between academia, industry practitioners, and policy makers will be crucial in driving forward sustainable data center development. By sharing knowledge, resources, and best practices, stakeholders can work together to overcome barriers to sustainability and achieve significant advancements in the field. Further research into how such a multidisciplinary, multi stake-holder framework for joint development should be designed and implemented is essential.

By addressing these areas, future research can significantly contribute to the ongoing development and refinement of sustainable practices in data center management, ensuring that technological advancements are balanced with environmental responsibility.

# References

[1] James D Abbey, Margaret G Meloy, Joseph Blackburn, and V Daniel R Guide Jr. Consumer markets for remanufactured and refurbished products. *California Management Review*, 57(4):26–42, 2015. pages 58

[2] Anadolu Ajansi. Critical raw materials and rare earth elements could reduce carbon emissions, 2023. pages 51, 52

[3] Adina Aldea, Maria-Eugenia Iacob, Jos van Hillegersberg, Dick Quartel, Lianne Bodenstaff, and Henry Franken. Modelling strategy with archimate. In *Proceedings of the 30th Annual ACM Symposium on Applied Computing*, pages 1211–1218, 2015. pages 35

[4] American Institute of Certified Public Accountants (AICPA). Statement on Standards for Attestation Engagements (SSAE) No. 18: Attestation Standards: Clarification and Recodification. Technical report, American Institute of Certified Public Accountants (AICPA), 2017. pages 45

[5] Taha Amir and James Henry. Securing sustainable data stores: Cybersecurity strategies for environmental responsibility. Technical report, EasyChair, 2024. pages 87

[6] Deborah Andrews, Elizabeth J. Newton, Naeem Adibi, Julie Chenadec, and Katrin Bienge. A circular economy for the data centre industry: Using design methods to address the challenge of whole system sustainability in a unique industrial sector. *Sustainability*, 13(11), 2021. pages 25

[7] Rabih Bashroush, Nour Rteil, Rich Kenny, and Astrid Wynne. Optimizing server refresh cycles: The case for circular economy with an aging moore's law. *IEEE Transactions on Sustainable Computing*, 7(1):189–200, 2022. pages 25, 29

[8] Ken Baudry. *DATA CENTER SITE SEARCH AND SELECTION*, chapter 22, pages 367–380. John Wiley & Sons, Ltd, 2021. pages 48

[9] Fiona Brocklehurst. International review of energy efficiency in data centres. *IEA EBC Building Energy Codes Working Group*, 2021. pages 49

[10] Rajkumar Buyya, Shashikant Ilager, and Patricia Arroba. Energy-efficiency and sustainability in new generation cloud computing: A vision and directions for integrated management of data centre resources and workloads. *Software: Practice and Experience*, 54(1):24–38, 2024. pages 87

[11] Onyinyechukwu Chidolue, Peter Efosa Ohenhen, Aniekan Akpan Umoh, Bright Ngozichukwu, Adetomilola Victoria Fafure, and Kenneth Ifeanyi Ibekwe. Green

data centers: Sustainable practices for energy-efficient it infrastructure. *Engineering Science & Technology Journal*, 5(1):99–114, 2024. pages 87

[12] RackSolutions .com. How many servers does a data center have?, Sep 2020. pages 48

[13] Yan Cui, Charles Ingalz, Tianyi Gao, and Ali Heydari. Total cost of ownership model for data center technology evaluation. In *2017 16th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITherm)*, pages 936–942, 2017. pages 19

[14] Andrew R. Curtis, S. Keshav, and Alejandro Lopez-Ortiz. Legup: Using heterogeneity to reduce the cost of data center network upgrades. In *Proceedings of the 6th International COnference*, Co-NEXT '10, New York, NY, USA, 2010. Association for Computing Machinery. pages 21

[15] Rajarshi Das, Jeffrey O Kephart, Charles Lefurgy, Gerald Tesauro, David W Levine, and Hoi Chan. Autonomic multi-agent management of power and performance in data centers. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems: industrial track*, pages 107–114, 2008. pages 51

[16] Andrews Deborah. Cedaci, 2023. pages 29, 57, 62

[17] Nicholas Dodd, Felice Alfieri, L Maya-Drysdale, J Viegand, S Flucker, R Tozer, B Whitehead, A Wu, and F Brocklehurst. Development of the eu green public procurement (gpp) criteria for data centres server rooms and cloud services. *Retrieved from JRC Science for Policy report: https://ec.europa.eu/environment/gpp/pdf/jrc118558_2020_0605_data_centres_technical_report_jrc_clean_with_id. pdf*, 2020. pages 49, 50, 51

[18] Lieven Eeckhout. Focal: A first-order carbon model to assess processor sustainability. *NotAvailable*, 2024. pages 87

[19] Lisa M. Ellram. *Total Cost of Ownership*, pages 659–671. Gabler Verlag, Wiesbaden, 2002. pages 19

[20] UN EP. Understanding circularity, Aug 2023. pages 29, 56, 57, 58

[21] Union European. Circular Economy Action Plan. `https://ec.europa.eu/environment/circular-economy/index_en.htm`, 2023. Accessed: April 11, 2024. pages 50

[22] Union European. Critical Raw Materials. `https://ec.europa.eu/growth/sectors/raw-materials/specific-interest/critical_en`, 2023. Accessed: April 11, 2024. pages 49

[23] Union European. List of Critical Raw Materials. https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52017DC0490, 2023. Accessed: May 11, 2024. pages 49

[24] Union European. Raw Materials Initiative. https://ec.europa.eu/growth/sectors/raw-materials/policy-strategy_en, 2023. Accessed: May 15, 2024. pages 49

[25] Christoph Fehling, Frank Leymann, Ralph Retter, Walter Schupeck, and Peter Arbitter. *Cloud computing patterns: fundamentals to design, build, and manage cloud applications*, volume 545. Springer, 2014. pages 49, 53

[26] Michael Ferdman, Almutaz Adileh, Onur Kocberber, Stavros Volos, Mohammad Alisafaee, Djordje Jevdjic, Cansu Kaynak, Adrian Daniel Popescu, Anastasia Ailamaki, and Babak Falsafi. Clearing the clouds: A study of emerging scale-out workloads on modern hardware. In *Proceedings of the Seventeenth International Conference on Architectural Support for Programming Languages and Operating Systems*, ASPLOS XVII, page 37–48, New York, NY, USA, 2012. Association for Computing Machinery. pages 22

[27] Leylane Ferreira, Patricia Takako Endo, Daniel Rosendo, Guto Leoni Santos, Demis Gomes, André Luis Cavalcanti Moreira, Glauco Estácio Goncalves, Judith Kelner, Djamel Sadok, Amardeep Mehta, and Mattias Wildeman. Standardization efforts for traditional data center infrastructure management: The big picture. *IEEE Engineering Management Review*, 48(1):92–103, 2020. pages 51

[28] Dave Fredricks. Data center challenges: The importance of power, cooling and connectivity, 2021. pages 43

[29] Martin Geissdoerfer, Paulo Savaget, Nancy M.P. Bocken, and Erik Jan Hultink. The circular economy – a new sustainability paradigm? *Journal of Cleaner Production*, 143:757–768, 2017. pages 48, 56, 57

[30] Albert Greenberg, James Hamilton, David A. Maltz, and Parveen Patel. The cost of a cloud: Research problems in data center networks. *SIGCOMM Comput. Commun. Rev.*, 39(1):68–73, dec 2009. pages 9, 10

[31] Albert Greenberg, Parantap Lahiri, David A. Maltz, Parveen Patel, and Sudipta Sengupta. Towards a next generation data center architecture: Scalability and commoditization. In *Proceedings of the ACM Workshop on Programmable Routers for Extensible Services of Tomorrow*, PRESTO '08, page 57–62, New York, NY, USA, 2008. Association for Computing Machinery. pages 22

[32] Hasith Gunasekara, Janaka Gamage, and Himan Punchihewa. Remanufacture for sustainability: A review of the barriers and the solutions to promote remanufacturing. In *2018 International Conference on Production and Operations Management Society (POMS)*, pages 1–7, 2018. pages 58

[33] László Gyarmati, András Gulyás, Balázs Sonkoly, Tuan A. Trinh, and Gergely Biczók. Free-scaling your data center. *Computer Networks*, 57(8):1758–1773, 2013. pages 22

[34] Damien Hardy, Marios Kleanthous, Isidoros Sideris, Ali G. Saidi, Emre Ozer, and Yiannakis Sazeides. An analytical framework for estimating tco and exploring data center design space. In *2013 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, pages 54–63, 2013. pages 13, 18, 19, 27, 29, 37, 70

[35] Janaka R. Gamage Hasith N.W. Gunasekara and Himan K.G. Punchihewa. Remanufacture for sustainability: a comprehensive business model for automotive parts remanufacturing. *International Journal of Sustainable Engineering*, 14(6):1386–1395, 2021. pages 58

[36] C van Hoorn. Improving on the open-loop reverse supply chain for data centre servers-the environmental and financial aspects of data centre server disposal. Master's thesis, University of Utrecht, 2016. pages 25

[37] Uptime Institute. Tier standard: Topology, 2021. pages 42, 43

[38] International Monetary Fund. How pandemic accelerated digital transformation in advanced economies. *NotAvaliable*, 2023. Accessed: 2024-03-27. pages 9, 10

[39] Automation IT. Cold, warm, and hot redundancy: Determining how much you need, 2024. pages 43

[40] Editor JetBrains. The state of developer ecosystem 2023, Dec 2023. pages 10

[41] Andrew Josey, Marc Lankhorst, Iver Band, Henk Jonkers, and Dick Quartel. An introduction to the archimate® 3.0 specification. *White Paper from The Open Group*, 2016. pages 35, 85, 87

[42] HIPAA Journal. Hipaa compliance checklist, 2021. pages 43, 45

[43] Moses Jeremiah Barasa Kabeyi. Geothermal electricity generation, challenges, opportunities and recommendations. *International Journal of Advances in Scientific Research and Engineering (ijasre)*, 5(8):53–95, 2019. pages 68

[44] Zain Kamran. Creating a scalable saas platform: Architectural strategies, scalability, and technology stack..., Dec 2023. pages 21, 22

[45] Krishna Kant. Data center evolution: A tutorial on state of the art, issues, and challenges. *Computer Networks*, 53(17):2939–2965, 2009. Virtualized Data Centers. pages 48

[46] Ilhan Keskin and Gurkan Soykan. Distribution grid electrical performance and emission analysis of combined cooling, heating and power (cchp)-photovoltaic (pv)-based data center and residential customers. *Journal of Cleaner Production*, 414:137448, 2023. pages 59

[47] Julian Kirchherr, Denise Reike, and Marko Hekkert. Conceptualizing the circular economy: An analysis of 114 definitions. *Resources, Conservation and Recycling*, 127:221–232, 2017. pages 27, 48, 56

[48] Adithya Kumar, Iyswarya Narayanan, Timothy Zhu, and Anand Sivasubramaniam. The fast and the frugal: Tail latency aware provisioning for coping with load variations. In *Proceedings of The Web Conference 2020*, WWW '20, page 314–326, New York, NY, USA, 2020. Association for Computing Machinery. pages 21

[49] Chunlin Li, Jianhang Tang, and Youlong Luo. Elastic edge cloud resource management based on horizontal and vertical scaling. *The Journal of Supercomputing*, 76:7707–7732, 2020. pages 21

[50] Chien-Yu Liu, Meng-Ru Shie, Yi-Fang Lee, Yu-Chun Lin, and Kuan-Chou Lai. Vertical/horizontal resource scaling mechanism for federated clouds. In *2014 International Conference on Information Science & Applications (ICISA)*, pages 1–4, 2014. pages 21

[51] Benedikt Martens, Marc Walterbusch, and Frank Teuteberg. Costing of cloud computing services: A total cost of ownership approach. In *2012 45th Hawaii International Conference on System Sciences*, pages 1563–1572, 2012. pages 19

[52] McKinsey & Company. Covid-19 digital transformation & technology. *NotAvailable*, 2021. Accessed: 2024-03-27. pages 9, 10

[53] Victor Millnert and Johan Eker. Holoscale: horizontal and vertical scaling of cloud resources. In *2020 IEEE/ACM 13th International Conference on Utility and Cloud Computing (UCC)*, pages 196–205, 2020. pages 22, 23

[54] Adreon Raymond Murphy and Alan S. Fung. Techno-economic study of an energy sharing network comprised of a data centre and multi-unit residential buildings for cold climate. *Energy and Buildings*, 186:261–275, 2019. pages 59

[55] David B Newell, Eite Tiesinga, et al. The international system of units (si). *NIST Special Publication*, 330:1–138, 2019. pages 47

[56] government .nl. The netherlands and sustainable goals, Sep 2015. pages 54

[57] Laura Talens Peiro, Fulvio Ardente, et al. Environmental footprint and material efficiency support for product policy. *Analysis of material efficiency requirements of enterprise servers*, 2015. pages 25

[58] Yichuan Qi, Dan Feng, Binbing Hou, Fang Wang, Jianxi Chen, and Yun Liu. A simulation-based study on the reliability of data center upgrades. In *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, pages 401–408, 2020. pages 21, 27

[59] Sazrina Ramli and Dian Indrayani Jambari. Capacity planning for green data center sustainability. *Int. J. Adv. Sci. Eng. Inf. Technol*, 8(4):1372–1380, 2018. pages 48

[60] J. Redžepagić, V. Dakić, and Z. Morić. Establishing the guidelines for using refurbished hardware in creating new data centers to lower the amount of e-waste. In *2023 46th MIPRO ICT and Electronics Convention (MIPRO)*, pages 486–490, 2023. pages 61

[61] Kevin Reilly and Manuel Silva Paulus. How scrappers cash in on gold from your old computer, Jan 2024. pages 11

[62] E Riso. Cold spares vs hot spares: Understanding redundancy in data centers, 2021. pages 43

[63] Chad Robertson. Integrating unorganised waste reclaimers into formal recycling systems: the positive role of key brokers. *NotAvaliable*, 2023. pages 58

[64] Fabiana Rossi, Matteo Nardelli, and Valeria Cardellini. Horizontal and vertical scaling of container-based applications using reinforcement learning. In *2019 IEEE 12th International Conference on Cloud Computing (CLOUD)*, pages 329–338, 2019. pages 22

[65] Simon Spinner, Samuel Kounev, Xiaoyun Zhu, Lei Lu, Mustafa Uysal, Anne Holler, and Rean Griffith. Runtime vertical scaling of virtualized applications via online model estimation. In *2014 IEEE Eighth International Conference on Self-Adaptive and Self-Organizing Systems*, pages 157–166, 2014. pages 22

[66] Fa-Qiang Sun, Gui-Hai Yan, Xin He, Hua-Wei Li, and Yin-He Han. Cpicker: Leveraging performance-equivalent configurations to improve data center energy efficiency. *Journal of Computer Science and Technology*, 33:131–144, 2018. pages 21

[67] Evandro Leonardo Silva Teixeira, Benny Tjahjono, Macarena Beltran, and Jorge Julião. Demystifying the digital transition of remanufacturing: A systematic review of literature. *Computers in Industry*, 134:103567, 2022. pages 58

[68] UNCTAD. How covid-19 triggered the digital and e-commerce turning point. *NotAvaliable*, 2021. Accessed: 2024-03-27. pages 9, 10

[69] European Union. General data protection regulation (gdpr), 2021. pages 43, 45

[70] Takashi UNNO. Recycling of personal computer vendors. *Journal of MMIJ*, 123(12):823–827, 2007. pages 58

[71] Aitor Villar-Martínez, Luis Rodríguez-Gil, Ignacio Angulo, Pablo Orduña, Javier García-Zubía, and Diego López-De-Ipiña. Improving the scalability and replicability of embedded systems remote laboratories through a cost-effective architecture. *IEEE Access*, 7:164164–164185, 2019. pages 23

[72] Johan Vogt Duberg, Gustav Johansson, Erik Sundin, and Jelena Kurilova-Palisaitiene. Prerequisite factors for original equipment manufacturer remanufacturing. *Journal of Cleaner Production*, 270:122309, 2020. pages 58

[73] Jane Webster and Richard T Watson. Analyzing the past to prepare for the future: Writing a literature review. *MIS quarterly*, pages xiii–xxiii, 2002. pages 16

[74] World Economic Forum. How can companies have a successful digital transformation? *NotAvaliable*, 2023. Accessed: 2024-03-27. pages 9, 10

[75] Guihai Yan, Jun Ma, Yinhe Han, and Xiaowei Li. Ecoup: Towards economical datacenter upgrading. *IEEE Transactions on Parallel and Distributed Systems*, 27(7):1968–1981, 2016. pages 13, 21, 23, 27, 32, 48, 57

[76] Zhongtao Yin. Creating an e2e process to manage the avl firmware at the datacenter. *Theseus Finland*, 2023. pages 49, 51, 52

# Annex

## A Interview Questions

# Server Sustainability and Circularity in Data Centers

**15 Jun 2024**

Statistical data shows that e-waste has been growing steadily for a number of years, combined with the necessity for newer hardware. However through sustainable decomissioning techniques, the costs associated with e-waste can be lowered and considered upon the initial investment.

* Required

## Data about you

We ask you to share some data about you and your way of working in order to conduct proper quantification

1

Email address *

2

What is your current company? *

3

Current position at your company? *

4

What is the size of the infrastructure you consider when answering? *

- ◯ LESS than 10 machines
- ◯ BETWEEN 10 and 100 machines
- ◯ BETWEEN 100 and 500 machines
- ◯ MORE than 500 machines
- ◯ Other

5

Will you want to receive and/or discuss results of this study? *

- ◯ No
- ◯ Yes, only receive a written copy via email
- ◯ Yes, be part of a live discussion on the findings

## Hardware Acquisition Costs

TCO stands for Total Cost of Ownership. It is a financial estimate that helps businesses and individuals assess the complete costs associated with owning, operating, and maintaining a particular asset or undertaking a project over its entire lifecycle. TCO analysis takes into account not only the initial purchase price but also all other costs that may arise throughout the asset's lifespan.

In the context of Data Centers we consider the following aspects in the TCO:

1. **Number of servers:** The number of servers that are often acquired and their distribution
2. **The cost of a server:** The impact the cost of the server has on the acquisition process
3. **The cost of power:** Consider if the new server hardware incresease or decreases the power consumption of the data center.
4. **Maintenance cost:** Consider if the new server hardware has an effect on the costs related to maintenance, please consider also long-term provisions.

Please indicate how much you identify with the following statements

6

When acquiring new machines, I always differentiate between HOT and COLD spares *

◯ Yes

◯ No

◯ Other

7

Please explain your answer

8

There always exists a budgeting scheme I need to follow. *

- I agree
- I disagree
- Prefer not to respond

9

The power efficiency of the hardware is an important factor for me *

- Yes
- No
- Other

10

Please explain your answer

11

I prefer acquiring servers that fit my datacenter's application usages
*

◯ Yes

◯ No

◯ Prefer not to respond

12

If you can respond, what is the criteria you use to when choosing hardware

13

The energy and cooling efficiency of the server is an important aspect of mine *

◯ Yes

◯ No

◯ Prefer not to respond

◯ Other

14

Please explain your answer

15

Do you consider sustainability when acquiring new hardware? *

- ◯ Yes
- ◯ No
- ◯ Prefer not to mention

16

Please select the aspects that you consider when acquiring new hardware *

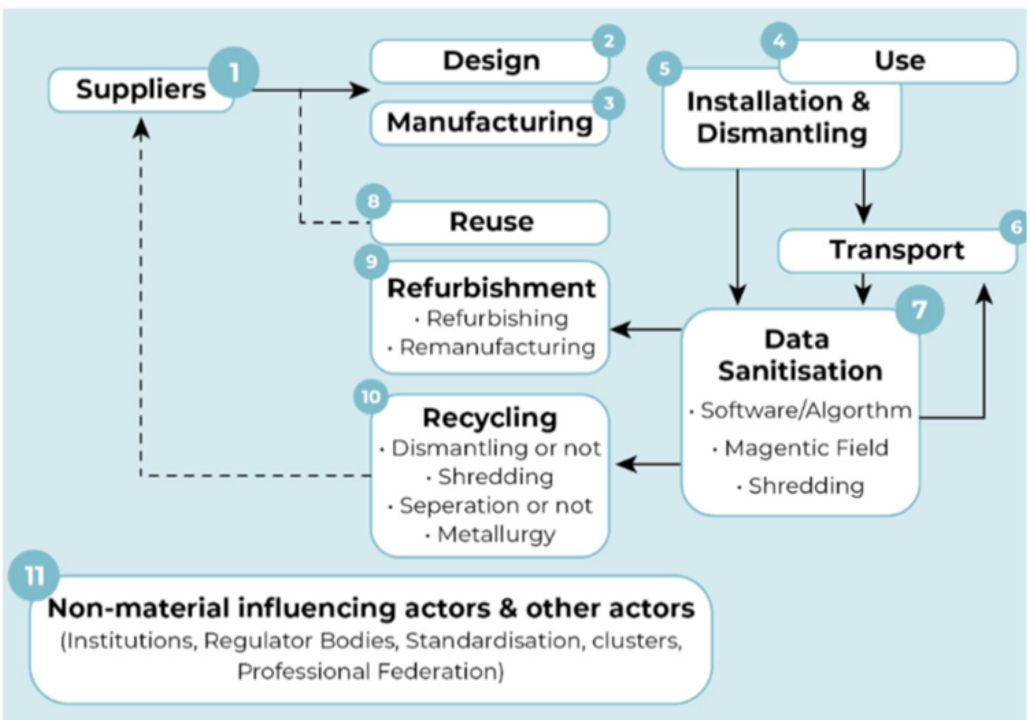
- [ ] Re-usability of the server
- [ ] Value of the server on the second-hand market
- [ ] Reduced energy consumption
- [ ] Reduced number of materials used
- [ ] Drop-in replacement
- [ ] Easy upgradeability
- [ ] None of the above

## Hardware decomissioning costs

This section explores various cost factors involved in the decommissioning process, including but not limited to:

1. **Disposal Expenses**: Expenses related to responsibly disposing of outdated or non-functional hardware in compliance with environmental regulations.
2. **Physical Removal and Logistics**: Costs related to physically removing servers from racks, transporting them to designated disposal or recycling facilities, and managing logistics throughout the decommissioning process.
3. **Equipment Salvage and Recycling**: Costs associated with salvaging reusable components or recycling materials from decommissioned hardware, which may include metals, plastics, and other recyclable materials.

Please indicate how much you identify with the following statements



17

How often do you remove old / broken hardware from your harware inventory ? *

- ◯ Weekly
- ◯ Monthly
- ◯ Seasonal (4 times a year)
- ◯ Yearly
- ◯ Never

18

How do you dispose of decomissioned hardware? *

- ☐ Dismantle the hardware
- ☐ Refurbish components for other servers
- ☐ Throw to garbage
- ☐ Return to manufacturer
- ☐ Recycle hardware
- ☐ Other

19

Can you describe your decomissioning process? What steps do you take w.r.t. sustainability?

20

Are there costs associated with the disposing method ? *

- ◯ Yes
- ◯ No
- ◯ Prefer not to say

21

What could you do more in terms of sustainability when decomissioning servers *

- [ ] Sepparate materials
- [ ] Sell hardware on the second hand market
- [ ] Re-use older hardware
- [ ] Re-thinking hardware requirements
- [ ] Planning and budgeting decomissioning
- [ ] None of the above

22

If you selected none of the above, are there other methods you use?

23

Hardware vendors should make use of standardised components in machines *

- ○ Yes
- ○ No
- ○ Prefer not to say
- ○ Other

24

Please explain the answer of "Other" above

## Circular Sustainability Strategies

The "*10 Rs of Sustainability*" represent a set of principles aimed at promoting sustainable practices and reducing environmental impact.

1. **Refuse**: Involves saying ***no*** to unnecessary or unsustainable products, services, or practices that contribute to waste generation or environmental degradation.
2. **Rethink**: Encourages individuals and organizations to reconsider their consumption patterns, habits, and choices to minimize waste and environmental impact.
3. **Reduce**: Focuses on minimizing consumption and resource usage wherever possible, such as using less energy, water, and materials.
4. **Reuse**: Promotes the reuse of products, materials, and resources to extend their lifespan and reduce the need for new production.
5. **Repair**: Encourages repairing items rather than discarding them when they break or become damaged, thus extending their useful life and reducing waste.
6. **Refurbish**: Encourages exchanging / replenishing only a few components of the entire system. I.e dead hard drive.
7. **Repurpose**: Involves finding alternative uses for items or materials that may no longer serve their original purpose, thus reducing waste and resource consumption.
8. **Recycle**: Emphasizes the recycling of materials and products to divert waste from landfills and conserve natural resources. The emphasis here is to recover working components and re-use those, whils the unusable ones are recovered.
9. **Recover**: Recover prime material from products. I.e aluminium & steel from server chasis.
10. **Remanufacture**: Re-manufacture the hardware using already existing material.

***Industrial symbiosis*** is a concept in which different industries and businesses collaborate to use each other's by-products, waste, energy, and other resources in a mutually beneficial way. In data centers the main resource to be exchanged is heat and energy however recent studies show that hardware can also be repurposed for other industires.

25

Considering what you've known so far, please arrange the following items in order from the most viable to the least. *

- Rethinking
- Reuse
- Repair
- Refurbish
- Recycle
- Recover

26

Please explain why you made this choice *

27

Have you heard the term of Industrial Symbiosis in the context of data centers? *

- ◯ Yes
- ◯ No
- ◯ Don't know

28

Please select an apropriate "feeling" with regards to each of the statements below. *

| | Completely disagree | Disagree | Neutral |
|---|---|---|---|
| Sustainability can only be achieved in datacenters through better cooling | ○ | ○ | ○ |
| Servers need to be changed anyway therefore we do not care how what we do with them | ○ | ○ | ○ |
| Decomissioning servers costs are important to our infrastructure | ○ | ○ | ○ |
| Decomissioning strategies should be known at the time of purchase | ○ | ○ | ○ |
| As a manager I am willing to invest time in decreasing the waste generated by my data center | ○ | ○ | ○ |
| The goal of Sustainability should overpower the costs it entails | ○ | ○ | ○ |



Laws and regulations

| | | | |
|---|---|---|---|
| should be changes so that more resources can be re-used from servers. | ◯ | ◯ | ◯ |
| | ◯ | ◯ | ◯ |

29

Please indicate your agreement with regards to the following statements *

| | Completely Disagree | Disagree | Neutral |
|---|---|---|---|
| Reusing hardware in datacenters is a crucial step towards a more sustainable IT infrastructure by reducing electronic waste and conserving natural resources. | ◯ | ◯ | ◯ |
| While reusing hardware has some environmental benefits, the performance and security risks associated with outdated equipment outweigh them. | ◯ | ◯ | ◯ |
| Reusing hardware offers mixed environmental impacts. | ◯ | ◯ | ◯ |

30

Please indicate your agreement with regards to the following statements *

| | Completely Disagree | Disagree | Neutral |
|---|---|---|---|
| Repairing hardware in datacenters extends its lifespan, reducing e-waste and resource use. | ○ | ○ | ○ |
| Repairing hardware can be expensive and time-consuming, making it less economical than replacing it with newer, more efficient equipment. | ○ | ○ | ○ |
| Hardware repair impact varies by complexity and lifespan extension. | ○ | ○ | ○ |

31

Please indicate your agreement with regards to the following statements *

| | Completely Disagree | Disagree | Neutral |
|---|---|---|---|
| Recycling hardware in datacenters minimizes landfill waste and recovers valuable materials, promoting a circular and sustainable IT economy. | ○ | ○ | ○ |
| Recycling hardware can be complex and costly, and the recovered materials might not always be readily usable in new equipment, reducing its overall environmental benefit. | ○ | ○ | ○ |
| The environmental impact of recycling hardware in datacenters hinges on factors like recycling efficiency, material recovery rates, and the energy used in the process. | ○ | ○ | ○ |

32

Please indicate your agreement with regards to the following statements *

| | Completely Disagree | Disagree | Neutral |
|---|---|---|---|
| Recovering materials from decommissioned datacenter hardware allows us to "decommission, not deplete," closing the loop on our resource usage and fostering a truly circular IT ecosystem within the data-driven world. | ◯ | ◯ | ◯ |
| Recovering materials from hardware can be energy-intensive and involve hazardous materials, potentially creating new environmental issues. | ◯ | ◯ | ◯ |
| Recovering materials from hardware offers potential environmental benefits, but also presents challenges related to energy consumption, recycling efficiency, and responsible hazardous material handling. Careful | ◯ | ◯ | ◯ |

consideration of these factors is crucial for an accurate assessment of its overall impact.

33

Please indicate your agreement with regards to the following statements *

| | Completely Disagree | Disagree | Neutral |
|---|---|---|---|
| Rethinking laws to incentivize hardware reuse, refurbishment, and responsible recycling in datacenters can unlock a sustainable IT future by promoting a circular economy and minimizing environmental impact. | ○ | ○ | ○ |
| Stricter regulations and law changes could stifle innovation and increase costs for datacenter operators, potentially hindering technological advancements and economic growth. | ○ | ○ | ○ |
| The role of law in promoting sustainable hardware practices in datacenters | | | |

| | | | | |
|---|---|---|---|---|
| needs careful consideration. Balancing environmental goals with economic realities and technological progress requires a nuanced approach that involves stakeholders from various sectors. | ○ | ○ | ○ | |

## Thank you!

Thank you for taking part into this research. We are very happy to share the results with you once the research is complete. If you shared any information that is confidential, please contact us at the email address:
**a.gorgan@student.utwente.nl**