Degree project

UNIVERSITY OF SKÖVDE

Environmental impact of producing and operating server infrastructure: A systematic literature review

Bachelor Degree Project in Information Technology

First Cycle 22,5 credits

Spring term 2021

Student: Andoni François Retegui Schiettekatte

Supervisor: Thomas Fischer

Examiner: Jianguo Ding

Acknowledgments

To my family and friends in Arrasate and my family in Xenia, for their unconditional love and support.

Thanks to Beñat Clemente, for the unstoppable help.

Thank you to the teaching staff from Högskolan I Skövde and Mondragon Unibertsitatea.

Abstract

Technologies that use cloud computing have increased considerably in the last decades. Computer services that we had in local devices have evolved to cloud services, resulting in applications such as online shopping stores, streaming platforms, or databases of large companies. In order to store and manage all the information that these types of services create, data centers were built and have been increasing in number and size in the last years. A big part of these databases is formed by server infrastructures, which are the devices that are responsible for storing and managing the data. Considerable research has been carried out to analyze the quantity of energy that servers consume and reduce it. What has been less considered is the environmental impact that servers have, not when they are operational but in the production phase (material gathering, transportation, assembly, etc.). A *Systematic Literature Review* (SLR) has been carried out to answer the research question that has been defined. The environmental impact of different stages is presented confirming the idea that other life cycle phases also have a big importance in the total environmental impact. Methods to reduce that impact are also explained.

Keywords: server infrastructure, environmental impact, life cycle assessment, production, energy consumption.

Table of Contents

Table of figures

Table of tables

List of abbreviations

CED	Cumulated Energy Demand
CMD	Cumulated Material Demand
CPU	Central Processing Unit
EOL	End Of Life
GHG	Greenhouse gases
GW	Gigawatt
IT	Information Technology
ICT	Information and Communication Technology
I/O	Input / Output
LCA	Life Cycle Assessment
LP	Low Power
Mt	Metric Tonne
PC	Personal Computer
TWh	Terawatt per hour
PCB	Printed Circuit Board
US	United States

1 Introduction

Cloud computing has been growing considerably in the last decades and has now become one of the main parts of existing computer services in the world. This is mostly due to how digital consumption has evolved in the latest years, where everything can be done anywhere with any type of device. This can partly be attributed to the growth in demand for devices such as smartphones, which has helped the expansion of this type of service. This can be seen in the increase of the global data traffic by 30 between 2005-2015.

The data that is used to offer these services has to be stored and that is where data centers enter the game. To see the extent of the energy consumption of this type of buildings, if they were treated as a country, they would have been the fifth-biggest electricity consumer worldwide in 2011 (Cook & Horn, 2011).

This can be also seen in countries like Germany, where 12TWh (Terawatt hour) were consumed by the professional data centers in 2014 and this number was expected to grow to 14TWh by 2020 (Peñaherrera et al., 2019). When looking for data about the global power consumption, in 2007, the annual average power consumption of data centers was 216 TWh (Lannoo et al., 2013).

(Lannoo et al., 2013) shows that the quantity of energy that data centers consume grew considerably in the next years, as data centers were estimated to consume 268 TWh in 2012. (Malmodin & Lundén, 2018) also states that the total consumption of data centers worldwide was 240 TWh in 2015. (Malmodin & Lundén, 2018) calculates that the carbon footprint of data centers and enterprise networks was 160 Mt CO_2e (Metric Tonne of CO_2 equivalent) in 2016. (Li et al., 2013) states that worldwide server market size was supposed to multiplicate by three for 2020, consuming 1000 TWh annually.

Data centers are infrastructures built by IT companies to manage and store big quantities of information. They are divided into three subsystems: IT infrastructure, power infrastructure, and cooling infrastructure. The one that will be analyzed in this thesis will be the IT infrastructure, which contains servers, networking equipment, and storage devices (Marwah et al., 2010). More specifically, the servers will be the devices that this report will analyze.

The servers were expected to account for 52% of the data center carbon footprint of the data centers worldwide for 2020, emitting around 136 Mt of CO_2e. For the same date, these devices were expected to account for 9% of the total ICT (Information and Communication Technology) footprint (Montoya et al., 2007). Other studies show that the number of servers was going to increase by a factor of 10 and the information managed by data centers by 14 by 2020 (Gantz et al., 2012).

The services that are offered through cloud computing usually have to be available at every moment and inefficient management of these devices can lead to significant electricity consumption leading to economic and technical consequences for the company. However, there is increasing concern about the environmental impact that this type of infrastructure has (van Heddeghem et al., 2012). The problem can be seen when we take into account that most of the country's energy source is still dependent on burning fossil fuels (Ren et al., 2012).

In response to these problems, several scientific works have been done regarding the reduction of energy consumption, including algorithms and strategies to efficiently manage servers. Nevertheless, these researches are only focused on the energy efficiency in the operational stage of the devices, not considering other processes that also have a big impact on the environment.

The environmental impact that this type of hardware piece has on the rest of its life cycle has not been addressed in the same way. Apart from the stage when the server infrastructure is operating, the device passes from numerous phases that also generate environmental impact such as resource gathering, manufacturing, transport, assembly, etc. In these phases, the environmental impact is big enough to consider.

(Chang et al., 2012) shows that the environmental impact related to production can reach 20% of the total impact of the device and can improve 30-40% with design changes. Due to the increasing market of the ICT (Information and Communication Technology) sector, the environmental impact of these devices will continue increasing in the short or medium term, and significant environmental improvements could be done by the means of material efficiency (Talens Peiró et al., 2020).

To answer this problem, two main actions should be taken. On the one hand, the usage of energy sources that contribute to big GHG (Greenhouse Gas) emissions should be lowered. However, this is not a completely feasible solution, as there are a lot of setbacks in the use of renewable sources, such as limited transmission capacity and losses, and intermittence. Because of this, the energy efficiency of each one of the phases in the life cycle of a server should be addressed so that this can be reduced to the maximum.

2 Background

Climate change concern is increasing over the years as the effects that it has and will have in the future are very worrying. Therefore, there is a general effort in the ICT industry to optimize the devices so that it becomes more efficient regarding energy consumption.

Cloud computing is a group of applications, hardware, and software delivered to the end user with the use of the internet. It is one of the most popular technologies because of its new way of offering services by managing data and services in the cloud. It has continuously been increasing in the last decade and because of that, IT companies have started to invest more in storage capacity and data center infrastructure.

The data center is a building which's purpose is to store and manage data usually from the internet. This type of infrastructure is completed with several devices, including cooling systems, cables, routers, switches, firewalls, servers...

2.1 Server

A server is an ICT device that offers various services such as storing information or executing transactions for the client, which is the device connected to it on the other side. There are different server types, mail servers, web servers, or enterprise servers for example. In this review, enterprise servers will be the ones that will be under discussion. These are devices used for business-to-business communication usually located in data centers (Talens Peiró et al., 2020).

This hardware piece contains a lot of different components inside of it that can vary depending on the model, but usually, they have components like hard drives, memory, processors, power source units, memory, and PCB boards (Hannemann et al., 2010).

This specific type of server is responsible for storing and managing big amounts of data. It must be available all the time, as any user who is using the service that is hosted in the server can try to access it at any moment. That model of use is what makes this type of device's environmental impact interesting and important to study.

2.2 Environmental impact

An environmental impact is a change to the environment resulting from an activity, product, or living being. In the last centuries, the increase of the annual emissions of CO_2 to the atmosphere has had a big environmental impact. These emissions affect the planet in a lot of different ways. For instance, it makes the average temperature of the planet grow, having consequences like the melting of the poles and increasing the level of the water. It also affects the pH concentration of the ocean's surface, increasing acidity and making it more difficult for marine wildlife to live (Lindsey, 2020).

The environmental impact of this type of device is something that must be assessed to lower it as it could be a very big harm in the near future. For that, a transition to greener energy would be the perfect solution, and even if changes are being done in that direction, it is unrealistic to think that the energy production industry will change fast.

However, decrementing the environmental impact by lowering the energy consumption of the whole life cycle of these devices is a plausible solution. Even if it would not solve the problem of using polluting energy sources, it would decrease the amount of energy needed and consequently, the pollution created by it.

Before implementing any change, the server's energy consumption should be addressed in its entire life cycle. By doing this, the effect that the device has on the environment can be identified and decisions can be taken knowing the real consumption of it. For that, several metrics can be used which measure different assets, even if all of them have the common objective of quantifying environmental impact. The following metrics can be used:

- **Carbon Footprint:**

It is a metric used to measure the total amount of CO_2 that is caused by an action or is piled up over the life phases of a product (Wiedmann & Minx, 2008). Even if there are more gases in greenhouse gases, there are much fewer data available compared to carbon. In this study, different variants of carbon footprint indicators will be used.

  - kg of CO_2: Quantity of CO_2 emitted by the process.
  - $CO_2e(q)$: Measures an amount of greenhouse gas whose impact has been standardized in units of CO_2, following the global warming potential of it (U.S. EPA, 2014).
  - Mt: Metric tonne is a unit that is equivalent to 1000kg of mass.

- **Exergy:**

This metric is based on the second law of thermodynamics, which states that all processes are irreversible and there is always a loss of energy. Exergy is defined as the energy of an irreversible process, in which there is a loss of energy (Jørgensen & Svirezhev, 2004). Several studies have argued about how the destruction of exergy represents the irreversibility of the process and environmental sustainability (Chang et al., 2010). This metric can be represented using units like Joules or watts per hour. These units can be also presented with their variants when speaking about big numbers (Mega, Giga, Tera…).

- **CED (Cumulated Energy Demand):**

This measure estimates the energy consumption that has taken place during the life cycle of a product direct and indirectly. This calculation includes the energy use during all the phases of the life cycle (Huijbregts et al., 2005). As it is an energy consumption estimator, it can be represented with the same units as exergy.

3 Problem definition and motivation

First, the aim of the project will be discussed. Second, the motivation to answer the research question will be explained. Finally, the limitations that apply to the project will be described.

3.1 Aim

The objective of the project is to show the actual environmental impact that server hardware has. For this, an answer will be given to the next question:

What is the environmental impact of producing and operating server infrastructure?

This question will be answered globally taking all phases of the server life cycle into account, which are resource gathering, transport, assembly, operation, and end of life.

3.2 Motivation

Cloud computing is one of the fields that is growing and will heavily grow in the near future, that is why the industry is urged to find a solution towards the high energy consumption of server infrastructure devices throughout their whole lifetime (Hannemann et al., 2010).

This research is motivated by the fact that the information regarding this field is limited, making it difficult to find information about server environmental impact when investigating something related to it. This project wants to serve as a reference of information for researchers or anyone that want to start an investigation or broaden their knowledge in this field.

3.3 Limitations

As the topic of the research is one that has brought a lot of controversy to the field of computer science. There are always interests in not publishing data that can tarnish the reputation of certain companies, so these brands work on hiding the impact that they have on the environment from their customers, using techniques like greenwashing (Hazel & Brittany, 2020). This makes the action of gathering literature about this topic a difficult task and limits the results that can be collected.

The method that will be explained in the next chapter has also been limited, as developing this project with direct access to servers would have been a really good way of improving the project results as new data and information would emerge. This was mainly because of two reasons.

On the one hand, considering the actual pandemic, having to go to a company would have risen the risk of getting or spreading the virus. Also, as mentioned before, there are interests in not showing this part of the business, so that consumers are unaware of how this type of hardware piece affects the environment during its lifecycle, making it difficult to find a company that would accept to have an intern doing this study in their facilities.

4 Methodology

This section will describe the methodology that has been used to carry out the research, the stages that applying this method had and the results achieved from this methodology.

The method that has been used is the Systematic Literature Review, which is a unique method which objective is to review the state of the research in one specific subject or field. The name derives from the fact that it follows a systematic and rigorous standard to collect information (Okoli & Schabram, 2012).

This type of work must be a much-cited paper that researchers will use as a resource for getting the first insights into the topic. It is responsible for simultaneously summarizing the existing evidence, spot missing parts that should be investigated, and provide a general idea for positioning research objectives (Okoli & Schabram, 2012).

Even if it is exhaustive work, conducting an SLR is a tremendous service for the community and can be a work of great importance in the field that is being studied (Petticrew & Roberts, 2006).

For using this method, numerous steps must be taken. First, a goal or the purpose of the review has to be defined. For that, the reviewer studies the field in which the study is going to be based and achieves a good insight into it. After that, the protocol that is going to be followed will be documented giving a detailed description of it. Next, the search is done, and inclusion and exclusion screening will be applied, based on criteria defined by the reviewer. To go on, interesting information from each study will be systematically extracted and will be combined with other studies ' information. Finally, the method will be reported in sufficient detail to be able to reproduce it (Okoli & Schabram, 2012).

After having recollected the literature, another step should be taken to broaden the number of studies gathered and make sure that no study is left outside of the scope. For doing so, the backward snowballing technique has been used, which will be explained later.

4.1 Databases

Numerous databases of quality must be used to get a decent amount of information for the research. Novice researchers usually fall in the narrowness of search and gather only partial information of the field that they are studying. This usually happens in the computer science field due to the wide dispersion of resources among many databases and research should take time to expand the information source and conduct the research using multiple literature vendors (Levy and J. Ellis, 2006).

Because of using several databases, as the quantity of available documents is wider, duplication must be taken into account, as there is more chance of finding the same document repeatedly in different databases.

The following databases were considered to use as an information source for the research.

- ProQuest
- Scopus
- EBSCO
- IEEEXplore
- ACM Digital Library

(Norris, Oppenheim, & Rowland, 2008).

- ABI/INFORM
- JSTOR
- Science Direct
- Wilson Web

(Levy and J. Ellis, 2006)

- Citeseer
- DBLP
- Academic Search Elite
- LibSearch

From all these databases, only a few were selected for proceeding with the literature review as the other ones were not suitable for the review. On the one hand, ProQuest and ABI/INFORM were discarded, as it is not possible to make an institutional login with the university account and has a low number of results.

On the other hand, Citeseer, LibSearch, and Academic Search Elite were found not to have enough information or were focused on other fields instead of computer science.

Regarding DBLP, this is a different database, as it does not contain articles in it, but it has very useful information about them, such as references, citations, author information, a direct link to other databases where the article can be found, etc. This page may seem useless as it does not have the article itself available, but it is very useful in order to apply some practices that will later be explained, such as forward and backward search.

Other discarded databases are JSTOR and EBSCO, as they do not have a lot of information and the interface is not intuitive. With this being said, the following databases have been chosen as the ones that are going to be used for the research.

- Scopus
- IEEEXplore
- ACM Digital Library
- DBLP
- Springer Link
- Science Direct

4.2 Search terms

Taking the previously presented information as a reference, a study of the field was made to gather the most correct or meaningful words to carry the literature search. During this search, one of the conclusions that were made is that there is a lot more information about the energy consumption and environmental impact of computer hardware when they are operating (mainly in data centers) than the environmental impact that this type of device has in their entire life cycle. That is why the following words were discarded, as the articles that contain these words mainly focus on the operational phase:

- Datacenter design
- Datacenter
- Datacenter metrics
- Energy Consumption

And the followings strings were selected for making the literature search:

- server* AND (“environmental impact” OR pollution OR “carbon footprint” OR “environmental sustainability”)
- server* AND (“life cycle assessment” OR LCA OR “Lifecycle environmental impact” OR “lifecycle impact”)

4.3 Selection criteria

In this section inclusion and exclusion criteria that will be used to filter the literature will be discussed. This is the filter that the literature that has been found has to pass in order to be used in the final review.

4.3.1 Inclusion criteria

The following inclusion criteria were established:

- Written in English
- Published between 1/01/2005 – 1/01/2021
- Meaningful for the goal of the research
- Journals and scientific conferences
- Peer reviewed

Starting with the date that was selected as a filter, this date was selected because of two main reasons. On the one hand, technology progresses at such a fast pace that referring to old scientific papers for research in the actual days is pointless, as even if there is good literature, it will be outdated regarding the technology that is available and in use nowadays. On the other hand, it is true that the field that has been chosen for the report has not a big amount of information and that is why the date has not been limited in a very strict way, so that valuable information is not lost.

The third condition is that it has to be meaningful for the goal of the research. This means that the literature that has been analyzed need to have some data that can help to answer the previously defined research question direct or indirectly.

Regarding the requirement that it needs to be peer-reviewed, this means that other experts from the field have reviewed it before publishing it. This is made so that readers are sure that they can use the paper with complete determination and protect their work from being harmed by low-quality information (Davison et al., 2005).

4.3.2 Exclusion criteria

Regarding the exclusion criteria (explained above), the following ones were defined:

- Does not allow institutional login with the University of Skövde account.
- Must pay to get access to the work.
- Duplicated

4.4 Backward and Forward Snowballing Technique

To broaden the obtained literature, firstly backward snowballing can be done. This is the process of going backward to find more information in the field of study. For doing so, three main techniques can be used: backward reference search, backward author search, and previously mentioned keywords (Levy & Ellis, 2006).

Firstly, backward reference research consists of reviewing the documents that have been referenced in an article or paper that have been found with the literature search. Using this technique, new useful literature can be found that otherwise would not be identified, as it was not in the scopes of the string search that was previously done.

This technique is also mentioned by other researchers with the name of backward snowballing, in which the researcher reviews the articles that are in the references part (Wohlin, 2014). After doing so, the researcher must apply the basic criteria that were applied to previous papers and only look into the ones that pass these filters.

This can also be used as a way of finding the origins and other versions of the topic that is being investigated. This can be made on a lot of levels, repeating the process of reviewing the references as much as the researcher wants and giving the process a big depth.

Regarding backward author search, this consists of looking for other works that an author whose article was interesting has made to gather more information about the topic that is under investigation. Thirdly, looking into the keywords that the literature contains to find additional information about the topic would be the previously mentioned keywords method.

On the other hand, there is also another very useful technique which is called forward search. Regarding this type of search, there are two types of them: forward reference and forward author search (Levy and J. Ellis, 2006). The first one consists of looking for the literature where that article has been cited to find new trends or methods that have not been mentioned in that paper.

On the other hand, forward author search consists of seeking the new literature that the researcher has released since then to find, as in the previous technique, new findings in the field.

4.5 Validity Threats

(Wohlin et al., 2012) explains that validity threats are different factors that can influence the integrity of the results and objectivity of the study that has been carried. Scientific researches should always have a big resistance to validity threats. There are four different aspects of the validity that have to be taken into account, which will be explained in the next chapters.

4.5.1 Construct validity

To start, this validity measures to which extent the used protocol represents the principal idea of the researcher. In this case, the used methodology and the defined protocol go clearly in hand with the objective of the research. This can be proved seeing that the results that have been obtained at the end of the study have been optimal and have answered the initial question.

4.5.2 Internal validity

This type of validity is investigated when causal relations are important in the study. More specifically it is the risk that a non-identified factor affects the value that is going to be investigated without the knowledge of the researcher.

In order to avoid any bias by the researcher, all the protocol that has been followed in the investigation has been predefined and followed very carefully. The big number of studies that have been selected due to the defined protocol has helped to ensure internal validity, as all the events that affected the investigated factor were taken into consideration.

4.5.3 External validity

This validity type refers to the feasibility to amplify the finding to other cases and to what extent people that are outside of the case would be interested in it.

In this particular case, the methodology and investigation that has been done are of big interest to other cases, as the same assessments could be made for another type of hardware device. Apart from that, the results that have been obtained from the review are interesting for any researcher that wants to make a study regarding servers, as the environmental impact of it is something that must be taken into account prior to any study.

4.5.4 Reliability

This threat is about in what measure the investigation is dependent on the researcher that does it. In the best of cases, if research is later repeated with the same methodology, results should be identical.

In this particular case, the investigation that has been carried is totally reliable. This is mainly due to the strict and precise protocol that has been defined and used during the investigation. Every step that has been taken has been explained in deep detail and always backed with scientific evidence.

4.6 Methodology in practice

After having rigorously defined the methodology, this part will describe how this has been applied and which are the results that the reviewer has gathered.

On the one hand, databases' grammatical instructions were analyzed as search strings needed to be adapted to each one of them. Various things had to be considered.

First, booleans are terms that can be used to facilitate the search. Most of the databases use AND, OR, and NOT for manipulating the intended results. However, some of the databases also add other booleans such as NEAR and SAME which makes even easier the action of finding specific research. Second, special characters must be used to make the distinction between words and phrases, use it as a wildcard for a set of characters or specify terms that must appear next to each other.

After learning how each database grammar works, search strings for each database were created. After searching with the defined strings, the following steps were taken in order to only choose the documents that were useful for the review:

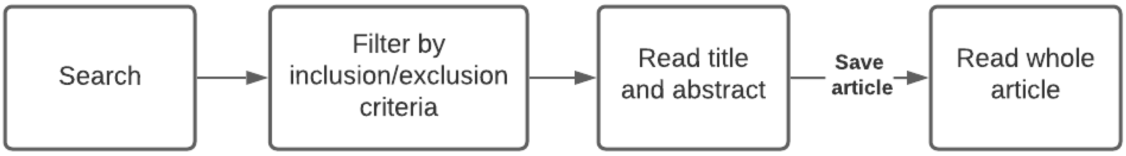

Figure 1 Articles processing method (Author's own)

The search was made following those steps in each database separately. Firstly, the search was made using the search strings adapted to each database. After that, inclusion and exclusion criteria were applied to narrow the results to the real scope. For that, filters offered by the host databases were utilized, some of them had the option of applying these filters manually, whereas in others the filters needed to be added to the search string. Some assets such as publication date, language, or publication type were filtered in this way. Other inclusion and exclusion criteria could not be applied mechanically and reading the documents was required. Therefore, other criteria were applied after saving the document.

- IEEExplore

Table 1 IEEExplore results

String Nº	Result Nº	Result after criteria[1]	Selected papers after criteria[2]
1	44	42	4
2	18	12	2

- ACM Library

Table 2 ACM Library results

String Nº	Result Nº	Result after criteria[1]	Selected papers after criteria[2]
1	52	47	4
2	8	4	0

- Science Direct

Table 3 Science Direct results

String Nº	Result Nº	Result after criteria[1]	Selected papers after criteria[2]
1	47	39	3
2	2	2	0

- Springer Link

Table 4 Springer Link results

String Nº	Result Nº	Result after criteria[1]	Selected papers after criteria[2]
1	39	36	1
2	7	5	0

- Scopus

Table 5 Scopus results

String Nº	Result Nº	Result after criteria[1]	Selected papers after criteria[2]
1	137	111	4
2	39	35	1

[1] These criteria were applied before saving the document (Date, language, and publication type)

[2] Criteria were applied after saving the document. (Peer reviewed, duplication and scope)

After having narrowed the selection of documents, each one of the studies was read to make sure that it was valuable for the review. To be efficient time-wise, as the reviewer could spend time reading documents that were not in the scope of the study, the studies were read in the following way.

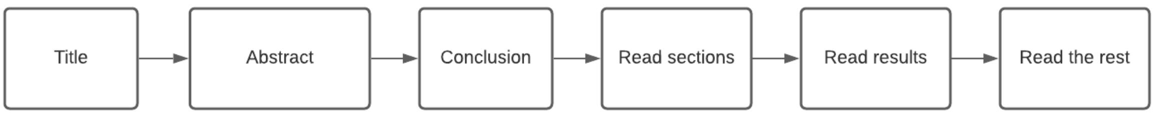

Figure 2 Steps for processing a study (Author's own)

This way of processing the study allowed a fast detection of false positives or research that did not meet the inclusion requirements. While doing so, the interesting or most valuable parts were noted, to facilitate the work in the gathering of information. With this done, the relevance of all the articles was confirmed.

Afterward, to facilitate the work of gathering information, a content table was created. In this table, each of the articles was marked manually depending on the information that they had concerning the research question.

Authors	Energy consumption	Use time env. Impact	Material	Transport	Assembly	EOL	Energy sources	Refresh cycle
F. Peñaherrera, J. Hobohm and K. Szczepaniak			x		x	x		
Jichuan Chang Justin Meza	x	x	x	x	x	x	x	
Christopher R. Hannemann, Van P. Carey	x	x	x	x	x	x		
Markus StutZ, Scott O'Connell, John Pflueger		x		x	x	x		x
Jichuan Chang, Justin Meza,	x	x	x	x	x	x		
Willem Vereecken1 , Ward Vanheddeghem								x
John Panneerselvam, Lu Liu	x	x						

Figure 3 Content table (Author's own)

After having read the articles, some of them were discarded as even if in the first insight it looked like it, it did not have valuable information for the study. Finally, the information was extracted from the articles and was synthesized for the review document.

After, in order to gather more information and complete the project, backward snowballing was done. For doing so, the references of the articles were read and the same methodology as in the first search was applied.

First, the exclusion and inclusion criteria were applied to each research. Second, the title was read and the usability of the work in relation to the aim of the project was considered. After that, as done with the first search, a table of content was made to facilitate the task of gathering information from the articles.

5 Results

In this part of the article, the results that have been achieved using the above-mentioned methodology will be presented. A brief introduction will be given, and afterward, the environmental impact of different life stages of servers will be discussed. After that, different ways of optimizing or reducing environmental impact will be assessed.

5.1 Environmental impact

The environmental impact of server infrastructure is undeniable at this stage. Five years ago, experts were expecting the energy consumption of data centers to grow to 14 TWh per year for 2020 and 16.4 TWh for 2025 (Hintemann, 2015). With the present energy production model, server energy consumption would have considerable negative effects on the environment, considering that still, 72% of the electricity consumed globally comes from non-renewable energies. Considering that the data centers are usually located in southwest Asia and the U.S., the consequences are even bigger, as the non-renewable electricity generation rises to around 75% (IEA, 2020).

Servers go through a lot of phases during their life cycle where the environmental impact is also present, but it has never been as investigated as much as the use phase. In the following lines, the negative impact of material gathering, transport, manufacturing, use phase, and end of life will be presented.

Each researcher divides the server life cycle into different stages, usually containing material gathering, transportation, manufacturing, use, and end of life. Other researchers such as (Chang et al., 2010) prefer to assess the impact that servers have through an architectural approach, stating that the model that most researchers use is ineffective.

Instead, they create three different aggrupation's: embedded, infrastructure, and operational impact. Inside of these three groups, they divide the components in an architectural approach instead of by process, separating the components that a server contains such as PCB, processors, and chassis.

Most of the studies use a metric called exergy to quantify the environmental impact, which has been previously explained. This measure has been used in numerous studies and has been proven that measures to reduce exergy go in hand with other measures that have environmental objectives (Shah & Meckler, 2009).

Using exergy for calculating the impact that these devices have on the environment is also more useful than other environmental indicators, such as carbon footprint, energy use, pollution, or impact on human health. This is because it is often difficult to relate one of these indicators with a design change, as one change can reduce toxic emissions but increase the impact on human health.

Another advantage of using exergy for measuring is the fact that there is a lot more publicly available information than with other environmental indicators.

The uncertainty of the environmental impact assessment was also studied, as various factors can heavily alter the calculations. First, the electricity emission varies depending on the country or geographical zone in which is assessed, as the energy sources vary and can have a different carbon footprint. Second, the distances of a particular geographical zone can be bigger than another, making the impact of shipping the servers higher (Weber, 2012).

LCA calculates the GHGs created in each phase of the product's life cycle (Pandey & Agrawal, 2011). For assessing the environmental impact of servers, each one of the stages of its life cycle will be analyzed. These stages are Material gathering, Manufacturing, Transport, Operational, and End of life. As it has been previously stated, apart from knowing the total environmental impact of the servers, the objective of this study is to determine the impact that the other phases have in comparison to the operational one, as the other phases are usually overlooked in scientific studies.

5.1.1 Material

Material selection is one of the crucial decisions that ICT device designers must take into account in order to reduce the carbon footprint that devices have mainly in the first and last stages of their life cycle.

In the production of a server, raw materials such as steel, copper, silicon, aluminum, polyethylene, polystyrene, and acrylonitrile butadiene styrene are used. However, the carbon footprint of the extraction and utilization of these materials can vary depending on the territory in which the study takes place.

(Hannemann et al., 2010) states that the material mass of the server mainly comes from the chassis and outer part, which is made of steel, accounting for almost 60% of the mass. 20% of the mass is due to the aluminum that the hard disk and power supply contains. Finally, copper and plastic are also materials that are a considerable part of the mass of the server.

Taking this into account, if we look to the exergy consumption that each of this material needs for extraction, we can see that even if the steel account for the biggest mass percentage, the exergy consumed to extract it account for less than 30% of the total exergy consumption of material gathering. The opposite happens with aluminum, which mass does not reach 20% of the total but consumes more than half of the total exergy consumption (Hannemann et al., 2010).

All in all, (Hannemann et al., 2010) estimates that the environmental impact of the material gathering process is 2.287,4 MJ, which accounts for 14,58% of the total impact of the server.

(Chang et al., 2012) also analyzed the embedded exergy that materials used for the server with an existing material assessment model. In this assessment, silicon is the material that prevails, consuming almost 50% of the total embedded energy. Steel and aluminum are other materials that consume a big part of the embedded exergy. The PCBs are the components that contain most of these materials making them the component that has the highest environmental impact. In this assessment, the impact of manufacturing is also included in the embedded exergy consumption, making the data of material extraction impossible to know. However, it can be assumed that the materials that consume a big part of the energy in manufacturing processes are also energy costing when gathering them.

Taking this into account, if we look at the materials that servers are made of, Silicon, Aluminum, Steel and Copper are the main materials used. As they have a big exergy cost, the environmental cost of gathering and recycling these materials is very high.

As stated before, some researchers believe that taking an architecture-centric approach is better when analyzing the material impact. Usual material assessment is made by disassembling the components and examining their chemical composition. On the one hand, if this approach is used, material assessments to new servers are going to be difficult, as the whole process will have to be repeated. Also, when applying optimization changes to a component, it is hard to understand how it affects the materials in the server and its exergy consumption (Chang et al., 2012).

Seeing the results from (Chang et al., 2010), material gathering accounts for 6% of the total environmental impact of the server. PCB is the component that has the biggest exergy consumption, which is partly made using copper. Memory, chassis, and power supply are other components that have the highest exergy consumption. (Chang et al., 2012) made a similar study 2 years later, obtaining very similar results.

On the other hand, (Peñaherrera et al., 2019) also divides the server into components to analyze the materials that each of them has. The Cumulated Energy Demand (CED) of the material parts accounts for 5% of the total used energy. The component that most CED had is the PCB, which contains plastic, copper, aluminum, and ceramic and takes around 4% of the CED in the server's entire life cycle.

In the following table, the effect that the material gathering process has on the final environmental impact is presented, showing the results of different authors that have been examined.

Table 6 Impact of the material gathering process

Author	Impact of material gathering in comparison with total	Value	Type
(Hannemann et al., 2010)	14,58 %	2.287,4 MJ	Exergy
(Peñaherrera et al., 2019)	5%	-	CED
(Chang et al., 2010)	6%	1.416 MJ	Exergy

All in all, material selection and the process of extraction of it should be assessed with more deepness in the future, as it has notable importance in the environmental impact of the server. Even if it varies, as not all models use the same materials and process, the impact of material extraction can go from 5% to 15% of the total exergy consumption or the total CED of the server.

5.1.2 Manufacturing

Manufacturing is another phase that is often not considered but must be assessed. In this phase, the materials that have been extracted and transported are used to assemble the server and prepare it for sale. When assessing manufacturing environmental impact on other devices such as cell phones, the conclusion is that this phase accounts for 40-50% of the total impact of the device (Chang et al., 2012).

(Hannemann et al., 2010) accounts that the exergy consumption of the manufacturing processes can be 485,2 MJ. Accounting for between 1 and 3 percent of the total exergy consumption of the server. On the other hand, (van Heddeghem et al., 2012) calculates that the environmental impact of the manufacturing stage of a server is 500 kg of CO-eq.

As previously mentioned, (Chang et al., 2012) and (Stutz et al., 2012) assess the manufacturing impact together with the material extraction phase.

(Peñaherrera et al., 2019) states that the CED of the assembly of the servers corresponds to 4% of the total CED of the server life cycle. They also state that the percentage of energy consumption of the manufacturing process depends highly on the server's lifetime as if the servers have a short lifetime, the importance of manufacturing impact highly increases.

A big part of the embedded exergy is destroyed when manufacturing because high-quality energy is needed on some complicated processes to manufacture electronic components and numerous chemicals. The manufacturing processes consume 13% of the total exergy consumption of the server and the materials are divided in the following way; Silicon stands almost for 75% of the exergy consumption and PCB consumes around 10% of the total manufacturing exergy consumption (Chang et al., 2010).

In the following table, the effect that the manufacturing process has on the final environmental impact is presented, showing the results of different authors that have been examined.

Table 7 Impact of the manufacturing process

Author	Impact of manufacturing in comparison with total	Value	Type
(Hannemann et al., 2010)	1-3%	485,2 MJ	Exergy
(Peñaherrera et al., 2019)	4,9%	-	CED
(Chang et al., 2010)	13%	3.068 MJ	Exergy
(van Heddeghem et al., 2012)	8,77%	500 kg COeq	Env. impact

In conclusion, the impact of the manufacturing process can vary between 1% and 13% when talking about energy consumption and accounts for 8,77% when talking about environmental impact. The variability resides in the different processes, materials, and machinery that are used in this phase.

5.1.3 Transport

After gathering the resources, materials have to be transported to the country where the manufacturing factories are. The previously mentioned materials are usually gathered from countries like China and Russia. Manufacturing of these devices takes place in south-east Asian countries, making the use of some means of transport more suitable than others.

Regarding the environmental impact that different transportation methods have in the environment, planes are the more harmful ones, as it consumes 10 times more (22,41 kJ/km) exergy than trucks. Train and ships are considerably less harmful to the environment, as they consume 0,253 and 0,296 kJ/km (Hannemann et al., 2010).

(Hannemann et al., 2010) estimates that the environmental impact of the transport phase accounts for 0,48% of the total, which results in energy consumption of 75,1MJ.

On the other hand, (Stutz et al., 2012) explains that in the case of a dell rack server, 90% of the transport is made by Truck and the rest of the road is transported by plane when delivering a server in the U.S. after manufacturing it in Austin, TX. They also state that regarding the total environmental impact, transport does not have a big effect on it, but it must be considered too.

An examination of the servers' manufacturers' logistics was made and different transporting modes were identified for different stages. The subassemblies were usually shipped by truck or water to the factory while the final product was usually transported by plane. This is why the biggest impact takes place in the last phases of the logistic process (Weber, 2012).

(Weber, 2012) estimates that the environmental impact of transport of a server is 37 kg of CO_2, which would be around 0,5% of the total environmental impact of the server. They also compare the impact of logistics with a desktop computer, concluding that the big difference is due to the heaviness of the device being delivered through air. This heaviness comes from the fact that the rack chassis is much bigger and it also has multiple hard drives and power supplies.

In the following table, the effect that the transport process has on the final environmental impact is presented, showing the results of different authors that have been examined.

Table 8 Impact of the transport process

Author	Impact of transport in comparison with total	Value	Type
(Hannemann et al., 2010)	0,48%	75,1 MJ	Exergy
(Chang et al., 2010)	< 1%	236 MJ	Exergy
(Weber, 2012)	0,56%	37 kg CO_2	Env. impact

To sum up, the exergy consumption of the transport phase can go from 0,48 to 0,99% of the total exergy consumption and when talking about the environmental impact, it accounts for 0,56% of the total. As previously said, this can vary due to the mean of transport, geographical location, or weight of the device.

5.1.4 Operational

Even if the objective of this project is to account for the environmental impact that servers have through their whole life cycle, it cannot be denied that due to its properties of use and the long use time that these devices have, compared to devices such as desktop PCs, the operational phase is the one in which servers have the biggest environmental impact (Hannemann et al., 2010).

The investigations that have been carried out used different data when doing the analysis. Factors such as time of use and peak of load are assets that have to be taken into account, as this affects considerably the results of the studies.

Starting with the study that was made by (Stutz et al., 2012), different assumptions were made to carry the study. Firstly, they estimated the lifetime of the device to be 4 years in continuous execution, and server operation was 50% in idle workload and the other 50% in full workload. For the environmental impact, it was assumed that the server used the US grid mix source.

With all these properties defined, the server produces 6.360kg of CO2eq and the use phase accounts for more than 90% of the total life cycle environmental impact, resulting in 5.960kg of Co2eq.

Moving into the study made by (Peñaherrera et al., 2019), they also specify some assets before calculating the environmental impact of the usage stage. An average load of 70% is estimated and usage time is almost the whole year, estimating that the server remains unavailable for 12 hours per year. Apart from that, in this particular study cooling is also taken into account when assessing the life cycle impact.

This study also uses the terms CMD (Cumulated Material Demand) and CED (Cumulated Energy Demand) taken from the study carried out by (Giegrich et al., 2012) intending to quantify the energy and material demand of the server. Operation CED is quantified to 50,7% of the total CED of the server.

Moving into the work that (Hannemann et al., 2010) developed, they also defined some assets when they made up the research. On the one hand, they defined the lifetime as 3 years and an uptime of 95%. On the other hand, a study was made to quantify the maximum load of the server and resulted in 265 W.

Taking 50% of the peak server as a constant server usage, the operational phase accounts for 75,93% of the total exergy consumption. Even if this study's objective is to assess the environmental impact of the server, (Hannemann et al., 2010) consider that adding the external cooling is pertinent as this part has a very big impact. If external cooling running at 100% is considered, cooling exergy consumption moves from being 6%, which is the consumption from the internal cooling, to 63%. This increases the total exergy consumption by 150%.

Even if it is true that this study is limited to server only environmental impact, the debate of whether it is appropriate to only assess the device's impact without considering the rest of the components is legitimate and should be addressed in further studies. The infrastructure used in the rest of the phases of the life cycle is also considered, but (Hannemann et al., 2010) argue that the time spent in these stages is not big enough to have a big impact on the total life cycle impact.

Moving into the next study, similar assets to the previous study were established. An operation time of 3 years was established and a 99,99% of uptime to do the study. In this study, they also include the cooling system of the data center in the study. The results that came out showed that operational exergy consumption accounts for 53% of the total (Chang et al., 2010).

As for now, the environmental impact of the use phase has been assessed. However, the idleness of a big part of the servers that are operating in the data centers has not expressly been mentioned. Zombie servers are servers that use big proportions of idle resources which are usually consuming energy but not delivering useful information (Panneerselvam et al., 2017).

For the analysis of the impact of zombie servers, three servers were defined. A and C were defined as mid-spec servers and B was defined as a high-spec server. Apart from that, the memory and the CPU environmental impacts were also assessed separately.

The carbon emission created by the power consumption of the idle CPU resource time is 1.213, 3.390, and 1.115 kg of CO2, respectively. Regarding memory, the average CO2 emissions are 414, 1.157, and 381 kg of CO2.

With these results, two conclusions can be made. On the one hand, it is clear that CPU use time has a much higher environmental impact than memory. On the other hand, servers having higher specs are also bigger energy consumers than the lower ones, making their impact on the environment bigger (Panneerselvam et al., 2017).

(van Heddeghem et al., 2012) also presents an estimation of the impact of the servers in the use phase considering that the server is operative for 4 years. The work shows that the impact is 5.200kg CO2-eq. On the other hand, (Chang et al., 2012) also estimated the amount of the exergy spent in the operational phase, resulting in 54% of the total exergy consumption.

According to (Weber, 2012), the average environmental impact of the servers' use phase was 6.238 kg CO2e. However, the use phase in different geographical zones was also assessed. The biggest environmental impact was identified in North America, 7.917 kg CO2e as the electricity emission rate was 0,72kg CO2/kWh. After that, Asia, Europe, and South America were also assessed giving results of 5.962, 5.084, and 4.241 kg of CO2 in the use phase.

As it has been said in the previously mentioned study, the environmental impact of servers' operational phase varies highly depending on the geographical zone in which the measurements have taken place.

In the following table, the effect that the operational process has on the final environmental impact is presented, showing the results of different authors that have been examined.

Table 9 Impact of operational process

Author	Impact of operational phase in comparison with total	Value	Type
(Hannemann et al., 2010)	75,93%	11.908,8 MJ	Exergy
(Peñaherrera et al., 2019)	50,7%	-	CED
(Chang et al., 2012)	54%	-	Exergy
(Chang et al., 2010)	53%	23.600 MJ	Exergy
(Stutz et al., 2012)	>90%	5.960 kg Co2eq	Env. impact
(Weber, 2012)	94,41%	6.238 kg CO2eq	Env. impact
(van Heddeghem et al., 2012)	91,23%	5.200 kg CO2eq	Env. impact

On the whole, it is clear that the operational phase is the one in which the biggest impact takes place. The exergy consumption goes from 50,7 to 75,93% of the total exergy consumption. On the other hand, the environmental impact of this phase goes from 90% to 94,41% of the total environmental impact of the server. The big variability can be due to different configurations, use times, or components of the server.

5.1.5 End of life

End of life is the procedure that comes after stopping using the server. There are different options when assessing the different parts and materials of the server. In European countries and the U.S, the regulation says that at least 75% of the materials that have been used on it have to be recycled. The rest must be incinerated so that energy that has been used can be recovered from it.

That happens in the case of Dell rack servers, as 75% of the server is recycled, its negative environmental impact is 86kg of CO_2eq, as this is saved from next server productions and saves 1,35% of the total environmental impact. Recycling a big part of the used material has a lot of advantages. On the one hand, it helps to reduce the environmental impact of the servers' EOL (End Of Life). On the other hand, it also avoids the environmental impact of extracting new material for future servers, as recycled material can be used (Stutz et al., 2012).

However, in the recycling processes machinery is used to do so but the environmental impact of these processes is a little part of the total impact of the server. For instance, (Hannemann et al., 2010) show that the recycling processes only take 0,06-0,02% (8,7 MJ) of the total exergy consumption of the server life cycle.

(Chang et al., 2010) also shows that the effect of this process on the overall environmental impact is very low. Regarding their data, recycling specific exergy consumption can not be concluded, but it can be extracted that transport and recycling together make 1% of the total exergy consumption.

Even if the negative impact of recycling is very low, some alternatives have also been studied and will be presented in the next chapter.

In the following table, the effect that the end-of-life process has on the final environmental impact is presented, showing the results of different authors that have been assessed.

Table 10 Impact of EOL process

Author	Impact of end of life phase in comparison with total	Value	Type
(Hannemann et al., 2010)	0,06-0,02%	8,7 MJ	Exergy
(Chang et al., 2010)	< 1%	< 236MJ	Exergy
(Stutz et al., 2012)	- 1,35%	86kg CO_2eq	Env. impact

Regarding the last phase of the server's life cycle, the exergy consumption can go from 0,02% to 0,99% of the total. On the other hand, its environmental impact in relation to the total environmental impact is -1,35%. This impact can vary depending on the localization in which the study has been developed, as different recycling policies are in force, or the machinery that is used in the process.

5.1.6 Total environmental impact

After discussing the total environmental impact of each one of the life cycle phases, the total impact of the servers will be presented. Each one of the analyzed researches presents the environmental impact with different units or values (exergy, CO_2 kg, etc.)

The location where the research has been carried has to be taken into account, as energy consumption can be the same, but depending on the country the carbon footprint changes.

(Stutz et al., 2012) states that the total environmental impact of the server is approximately 6.360kg CO2eq when being in operation in the U.S. The researchers propose improvements that could be done by modifying utilization and refresh rates, but this will be discussed in the next chapter. To put this data into perspective, the environmental impact can be compared to driving 21.500km with a SUV (Sport Utility Vehicle) car (Stutz et al., 2012).

(Hannemann et al., 2010) calculates the total exergy consumption to be 1.5684,2 MJ, which would approximately produce 2.614,03kg of CO2, taking into account the previously mentioned exergy/co2 conversion by (Hyser et al., 2011).

(Chang et al., 2012) estimates that the total lifetime exergy consumption of a server is 25,4GJ, which is equal to 25.400 MJ. Taking into account that 1kWh equals 3,6 MJ, the exergy consumption can be represented in 7.055,56 kWh. This would result in an approximate environmental impact of 4.233,336kg of CO2.

(Weber, 2012) also assesses the total environmental impact of the server, calculating that it is 6.607 kg of CO2e. Lastly, (van Heddeghem et al., 2012) estimated that the total environmental impact of a server is 5.700 kg of CO2.

In the following table, the total environmental impact of a server will be presented.

Table 11 Total environmental impact

Author	Environmental impact
(Hannemann et al., 2010)	2.614,03 kg CO2
(Chang et al., 2012)	4.233,336 kg CO2
(Stutz et al., 2012)	6.360 kg CO2eq
(Weber, 2012)	6.607 kg CO2e
(van Heddeghem et al., 2012)	5.700 kg CO2eq

As it can be seen, the results of different articles differ, as the values that each one of them presents are very different from the others. This can be due to a lot of factors. First, the use of different models of servers when doing the experiments may alter the environmental impact. Also, the methodology used in the different research could have resulted in this difference. As previously mentioned, the geographical zone in which the study has been carried can also affect the outcome as it can vary in the shipping and the emission-energy ratio.

Table 12 Life cycle phases impact and total environmental impact

	Material		Manufacturing		Transport		Operational		EOL		Total
Author	%	Value (MJ)	%	Value	%	Value	%	Value	%	Value	Value
(Hanneman n et al., 2009)	14,58	2.287	1-3	485,2 MJ	0,48	75,1 MJ	75,93	11.908 MJ	0,06 - 0,02	8,7 MJ	2.614 kg CO2
(Chang et al., 2010)	6	1.416	13	3.068 MJ	<1	236 MJ	53	23.600 MJ	< 1	<236 MJ	-
(Chang et al., 2012)	-	-	-	-	-	-	54	-	-	-	4.233 kg CO2
(Stutz et al., 2012)	-	-	-	-	-	-	>90	5.960 kg CO2eq	1,35	86 kg CO2eq	6.360 kg CO2eq
(Weber, 2012)	-	-	-	-	0,56	37kg CO2	94,41	6.238 kg CO2eq	-	-	6.607 kg CO2eq
(van Heddeghem et al., 2012)	-	-	8,77	500kg CO2	-	-	91,23	5.200 kg CO2eq	-	-	5.700 kg CO2eq
(Peñaherrer a et al., 2019)*	5	-	4,9	-	-	-	50,7	-	-	-	-

*Uses CED instead of exergy for calculating the energy consumption.

All in all, the total environmental impact that servers have during their whole lifecycle is estimated to be between 2.614,03 and 6.607 kg of CO2eq. Apart from that, in this summary, we can clearly see the importance of each phase in the final impact of the device. For instance, Material Gathering and Manufacturing phases can have an impact of almost 15% compared to the total impact of the server.

5.2 Improvements

Moving into the optimization techniques that can be applied to the servers, the collected articles and works present a considerable quantity of them. Some of them try to optimize exergy consumption by applying different techniques while others try to do it by changing the design of the server.

5.2.1 Design and operational improvements

To start with the improvements in server systems, design improvements have been considered and studied in the reviewed works. (Chang et al., 2012) studied three possible design optimizations. Firstly, the baseline is a rack container design that includes 1056 servers and has cooling from server fans and external cooling.

Secondly, the Low-power blade server architecture (LP-blade) proposed the use of low-power server components that goes along with I/O workloads. It uses internal cooling and the same external cooling as the previous design.

Thirdly, Dematerialized system architectures (Demat), is a design that finds to include disaggregation, dematerialization, and free cooling. Disaggregated servers divide functionality components into blocks improving sustainability by allowing efficient design and reuse of components. Dematerialized design reduces the material that is used to build the server. Free cooling consists of using the air from outside to cool the infrastructure without using internal fans.

All these design changes can be translated into improvements in the server's environmental impact. On the one hand, the baseline design is the one that consumes the biggest quantity of exergy, 25 TJ (Tera Joules). On the other hand, the LP-blade design consumes only 10,5 TJ, which reduces by 58% the impact that the baseline design had. Finally, Demat design is the best design of the three, only having an impact of 7,4 TJ which is translated into a reduction of 71% in comparison with the baseline design. With the last one, 500 to 600 metric tons of coal could be saved over the server's life cycle (Chang et al., 2012).

There are a lot of techniques that help to improve the effectiveness of the servers when operating. On the one hand, we have the energy proportionality (EP) technique in which the objective is to make the activity of the system and the energy consumption go hand in hand (Chang et al., 2010).

Consolidation is another of the used techniques. Normally, servers have a very low load during their operation time and peaks are unsynchronized. To lower the utilization exergy, the virtual machines from several servers are all grouped in one server. This makes the load of one specific server raise but decreases the number of operating servers (Chang et al., 2010).

Fourthly, low power (LP) server solution is also a used technique in the data centers. This is made using energy-efficient and low-power processors (Chang et al., 2010).

(Chang et al., 2010) studied the effect that these practices have in different situations in the total and operational exergy. When looking at total exergy consumption, EP works better than consolidation when the workload is not inconsistent and is high. This changes when the load goes under 50% of the capacity of the server, as the reduction of needed material concerning fewer servers makes embedded energy fall considerably. Regarding operational exergy, this does not happen, as perfect energy proportionality always outrages consolidation.

Regarding LP, after a certain value of resource activity, it is always better than the other two solutions. Even so, there is a difference when considering the whole life cycle or only the operational stage, as more server quantity is required when using this technique and the embedded exergy suffers an increase.

When applying this type of improvement, researchers always prioritize the progress in the operational stage. However, these optimizations can unexpectedly affect other life stages' environmental impact too (Chang et al., 2012).

For instance, LP-Blade design reduces embedded exergy consumption by 38%. This is mainly due to the infrastructure reduction and reduced material when producing LP servers. This type of server has a smaller silicon quantity and a material reduction in printed circuits, which is the part that higher impact has in the manufacturing process (Chang et al., 2012).

The same happens in the inverse direction, where the Demat improvement was intended to optimize the embedded exergy consumption but it also reduces the infrastructure-operation consumption. For example, the removal of the backplane reduces the consumption by 11% as there is a higher airflow. The use of shared fans between the servers also contributes to an improvement of 28% in efficiency (Chang et al., 2012).

In conclusion, making a complete study of the whole life cycle of the server is important, as this will give the researcher a complete insight into the impact that different optimization methods have on the environmental impact.

5.2.2 Refreshing cycles

Refreshing cycles are also a hotly debated topic in the sustainability improvements that are discussed inside this field. Lately, keeping the hardware devices as long as possible has become a very popular practice but its consequences are not always positive, as it depends on a lot of factors.

To study the impact of these refresh rates, (Stutz et al., 2012) defined two customer types. One that refreshes servers every 4,5 years and another one that does the same once every 3 years. It was also estimated that performance increased 30% every generation (3 years) and the power use decreased 10% because of the reduction of server quantity.

The results show that due to a decrease in power use applying the aggressive refresh cycle saves 1700kg of CO2eq over 10 years. If the power use decreased 20% every generation, this savings would become 6100kg of CO2eq.

In another scenario, in which performance improves by 30%, instead of 20% in each generation, an aggressive refresh cycle can save 11000 kg of CO2eq in 9 years, which is 11% of savings.

In conclusion, applying an aggressive refresh strategy results in improvements if the power use of servers decreases at least 10% every 3 years. The improvement is even bigger if, with the improvement of the systems, the numbers of servers are reduced and the performance keeps being the same.

(Vereecken et al., 2012) have also assessed the refreshing rates optimization method. In their case, two scenarios have been defined. On the one hand, there is a scenario where the number of servers remains constant, and on the other, the capacity of the set of servers remains constant but the number of servers can decrease, because of the improvement in the performance of new hardware pieces. The increase in capacity is based on Moore's law, which says that every 18 months processing capacity of servers doubles.

In the first scenario, they state that it is important to keep the available hardware in use as long as possible, as replacing them will increase the footprint because of the non-use-phase footprint of replaced servers. Also, considering that server power rises every year, the low number of replaced servers makes the overall power consumption decrease.

In the other case, as capacity increases, fewer servers need to be installed year after year. As capacity is growing faster than power consumption, old servers outweigh lower power consumption in the long term. So, if capacity keeps constant, the carbon footprint decreases when they are substituted in the optimal time of 2 years.

Continuing, if both techniques are compared when the refresh cycle rate is every 8 years, the constant server scenario footprint is lower than the constant capacities one. But, if this situation is assessed in the longer term, the constant number of servers footprint increases yearly due to the server consumption of new devices increasing. In the other case, the footprint decreases year after year due to the diminution of the required equipment. On the other hand, if the server refreshing rate is of 2 and 4 years the second scenario prevails in carbon emission optimization. With this, we can reiterate that the non-use-phase footprint is also important in the life cycle footprint of the server, as even if the new server's energy consumption rises, the fact of producing fewer servers makes the footprint fall dramatically.

5.2.3 Load migration

Load migration is a technique used to minimize the carbon footprint that certain energy consumption can generate. Depending on the moment of the day and the station, the production of renewable energy sources can vary, and consequently, the carbon footprint can fluctuate. For avoiding a big environmental impact, the load is passed from one data center to another that has a lower carbon footprint at that moment.

Cloud computing helps this mechanism to be applied, as the load can be transported from one center to another, avoiding the peak emission moments of both of the regions. (Dandres et al., 2016) takes Ontario and Alberta as the scope territories for the study. On the one hand, Ontario mainly relies on nuclear power for energy generation. On the other hand, Alberta uses natural gas and coal for energy generation purposes, having a high greenhouse gas emission per kWh.

Due to the energy generation models in both territories, the load is usually hosted in the Ontarian territory 71% of the time. Using cloud computing, the study successfully achieves to reduce the GHG emissions by 23% and 44% compared respectively to Ontario and Alberta. This stresses the importance of assessing the regional power generation sources before deploying a cloud computing system (Dandres et al., 2016).

6 Conclusion

This project was made to answer the following research question: *Which is the environmental impact of producing and operating server infrastructure?* For this, a systematic literature review was done following a very structured methodology and interesting results were gathered from it.

To start with the conclusions of the results that have been obtained, the research question that was first defined has to be addressed. The results have answered this question, showing the server's environmental impact in all the stages of its life cycle.

Even if depending on the assets that each study has defined results vary considerably, the results show that the impact of this type of ICT device in the environment is very big and is increasing yearly. Meaning that this should be addressed so that the increasing demand in cloud computing and this type of technology affects the environment as little as possible.

The operational phase came out to be the most influential one in the impact that the server has. Even if this was an expected result, giving the type of use that is given to this type of device, this does not take out the importance of addressing it.

Regarding the other life stages of the server, the influence that they have was bigger than expected at the first stage. This is not usually addressed in the LCA due to the complexity of tracking all the process that comes before the delivery of the device. However, a researcher should always include this part in their research as it gives a real perspective of the causes of the environmental impact.

To reduce this impact and build a more sustainable server infrastructure in the following years, different techniques have been presented in this document. All of them show a big capacity of saving a big quantity of energy and should be addressed in upcoming research.

7 Discussion and Future work

In this part of the report the methodology used and the improvements that could be done to it will be addressed. Apart from that, societal and ethical considerations will be made. Finally, future work that can be done in this field will be presented.

7.1 Implementation of the methodology

The systematic literature review was carried on following strictly the standards marked by the scientific community. More specifically following guides of computer science researchers that had a background in this type of methodology. To do so, fast research was done at the start of the project in order to gather knowledge about the different aspects to take into account.

Although the systematic literature review process is usually presented as a linear process, in practice it has a lot of different setbacks that convert it into a circular or iterative process. For instance, deciding the research question already took several days of work as other topics that were considered were impossible to do because of the current situation of the pandemic. Other topics were also brought down because of duplication of works of previous students.

Regarding the definition of the protocol that was defined to do the searches, this process was more time-consuming than expected. On the one hand, choosing which databases were going to be used took a lot of time as at first there were a lot of suitors. For choosing the correct ones, literature was read and trial searches were made if the scope of the project was going along with the database field.

Choosing the right search strings was also a time-consuming process, as different tries were made to understand which words were more effective on finding significant literature while avoiding false positives and false negatives as far as possible. However, as false positives were preferred to false negatives, very general word combinations were used to make sure that all the useful information was considered.

Applying inclusion and exclusion criteria however were very fast and effective, as a lot of literature was found about the usual practices in this aspect.

After executing the search, each result was assessed with the process that was explained in the methodology part. The chosen researches were downloaded and Mendeley Desktop software was used to manage these documents, facilitating a lot of the process of handling it. After that, even if it was more time-consuming, each article was read and a content table was made to facilitate the process of extracting the information from the documents.

Finally, the extracted information was written and developed in a presentable way in the report.

All in all, a way to improve would be to put in more working time at the start, considering that the process of preparing all the methodology was more time-consuming than expected. More communication with the supervisor of the project would be also a good improvement, starting to make weekly online meetings from the beginning of the project, for example.

Regarding the research question, one of the debates that came when doing the review was if only assessing the server energy consumption was effective and reasonable. In a data center, which is the building in which the enterprise servers are, there are a lot of different components that also consume big quantities of energy.

Knowing this, assessing the environmental impact of all components altogether could be more effective. On the one hand, regarding the optimization techniques, collective research would be more convenient, as a design change in the servers can reduce the footprint of the server but it can increase the effort that external cooling system must do increasing the overall environmental impact.

On the other hand, seeing a bigger picture of the environmental impact, preferences could be made, as other systems or components can have a bigger environmental impact than the devices studied in this review.

7.2 Ethical considerations

Since this review is not primary research, a lot of ethical problems regarding methodology can be discarded. However, some aspects could be mentioned.

To start, all the scientific literature that has been used in the review is of the public domain or accessible from the Skövde university database. This means that no restricted or unwanted information has been used that could tarnish the image or reputation of a company, organization, or any stakeholders.

All the data that has been presented have a scientific warranty as it has been taken from databases that publish only high-quality peer-reviewed literature. Regarding the methodology, all scientific standards have been followed to design it and execute it as high-quality scientific guidelines have been followed to do so.

Finally, all the results that have been obtained from this review have been presented in a completely objective and impartial way, avoiding bias from the reviewer.

7.3 Societal aspects

First, bringing all this information together in one study or review makes it more accessible for the people that have an interest in this field. The environmental impact that servers have through their whole lifecycle should be public and available to every individual that uses this type of device so that they are aware of the effect that it has on the environment.

Second, this work also helps in creating a base study for future researchers that will tackle this topic, giving a real insight into the situation. By facilitating research in this area more and more sustainable practices will be developed and the objective of doing the technological transformation while being sustainable will be even more feasible.

7.4 Future work

Looking into the future, different work can be done in this field. On the one hand, knowing that the presented optimization and improvement techniques can have a very positive impact, digging in this should be one of the main objectives.

Apart from the existing optimization techniques, which have been proven to be very efficient, new improvement techniques should also be researched to have a big range of practices that can reduce the energy use of servers to the minimum.

Even if research has to be done in the field of ICT devices, it is undeniable that with the growing increase of demand for these devices a long-term transition to greener energy sources should be one of the top priorities.

With both solutions, the environmental impact of server infrastructure would considerably decrease, reducing a big amount of the environmental impact of infrastructures such as data centers.

References

Chang, J., Meza, J., Ranganathan, P., Bash, C., & Shah, A. (2010). Green server design: Beyond operational energy to sustainability. *HotPower 2010 - 2010 Workshop on Power Aware Computing and Systems*, *1*, 1–5.

Chang, J., Meza, J., Ranganathan, P., Shah, A., Shih, R., & Bash, C. (2012). Totally green: Evaluating and designing servers for lifecycle environmental impact. *International Conference on Architectural Support for Programming Languages and Operating Systems - ASPLOS*, 25–36. https://doi.org/10.1145/2150976.2150980

Cook, G., & Horn, J. Van. (2011). How dirty is your data? *Telecommunications (Americas Edition)*, *40*(9), 24–26.

Dandres, T., Samson, R., & Lemieux, Y. (2016). *The Green Sustainable Telco Cloud centres*. *2*(c), 13–20.

Davison, R. M., de Vreede, G. J., & Briggs, R. O. (2005). On Peer Review Standards For the Information Systems Literature. *Communications of the Association for Information Systems*, *16*(January). https://doi.org/10.17705/1cais.01649

Gantz, B. J., Reinsel, D., & Shadows, B. D. (2012). Big Data , Bigger Digital Shadow s , and Biggest Grow th in the Far East Executive Summary: A Universe of Opportunities and Challenges. *Idc*, *2007*(December 2012), 1–16.

Giegrich, J., Liebich, A., Lauwigi, C., & Reinhardt, J. (2012). Indikatoren / Kennzahlen für den Rohstoffverbrauch im Rahmen der Nachhal- tigkeitsdiskussion. *Institut Für Energie- Und Umweltforschung Heidelberg GmbH*.

Hannemann, C. R., Carey, V. P., Shah, A. J., & Patel, C. (2009). Lifetime exergy consumption as a sustainability metric for enterprise servers. *2008 Proceedings of the 2nd International Conference on Energy Sustainability, ES 2008*, *1*, 35–42. https://doi.org/10.1115/es2008-54181

Hannemann, C. R., Carey, V. P., Shah, A. J., & Patel, C. D. (2010). Lifetime exergy consumption of enterprise servers. *International Journal of Exergy*, *7*(4), 439–453. https://doi.org/10.1504/IJEX.2010.033413

Hazel, S., & Brittany, M. (2020). Greenwashing in the Information Industry. *The IJournal: Graduate Student Journal of the Faculty of Information*, *5*(2). https://doi.org/10.33137/ijournal.v5i2.34413

Hintemann, R. (2015). Energy consumption of data centers continues to increase in 2014. *Borderstep Institute for Innovation and Sustainability*, 4–6.

Huijbregts, M. A., Rombouts, L. J. A., Hellweg, S., Frischknecht, R., Hendriks, A. J., Meent, D. van de, Ragas, A. M. J., Reijnders, L., & Struijs, J. (2005). *Is Cumulative Fossil Energy Demand a Useful Indicator for the Environmental Performance of Products*.

Hyser, C., Gmach, D., Ml, U., Chen, Y., & Suryanarayana, V. (2011). Improving server power management in research and development data centers. *Compute 2011 - 4th Annual ACM Bangalore Conference*, 1–6. https://doi.org/10.1145/1980422.1980428

IEA, I. E. A. (2020). *Electricity Market Report – December 2020*.

Jørgensen, S. E., & Svirezhev, Y. M. (2004). Work, exergy and information. *Towards a Thermodynamic Theory for Ecological Systems*, 95–126. https://doi.org/10.1016/b978-008044166-5/50005-7

Lannoo, C. B., Lambert, S., Heddeghem, W. van, Pickavet, M., Tudelft, F. K., Koutitas, G., Niavis, H., Certh, A. S., Till, M., Fischer, A., Meer, H. de, Passau, U. N. I., Ulanc, P. A., Epfl, P., Viet, N. H., Uio, T. P., & Uam, J. A. (2013). Overview of ICT energy consumption. *Network of Excellence in Internet Science, D8.1*, 1–59.

Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science*, *9*, 181–211. https://doi.org/10.28945/479

Li, C., Hu, Y., Zhou, R., Liu, M., Liu, L., Yuan, J., & Li, T. (2013). Enabling datacenter servers to scale out economically and sustainably. *MICRO 2013 - Proceedings of the 46th Annual IEEE/ACM International Symposium on Microarchitecture*, 322–333. https://doi.org/10.1145/2540708.2540736

Lindsey, R. (2020). Climate Change: Atmospheric Carbon Dioxide. *Noaa/Nasa*, 6–10.

Malmodin, J., & Lundén, D. (2018). The energy and carbon footprint of the global ICT and E & M sectors 2010-2015. *Sustainability (Switzerland)*, *10*(9). https://doi.org/10.3390/su10093027

Marwah, M., Maciel, P., Shah, A., Sharma, R., Christian, T., Almeida, V., Araújo, C., Souza, E., Callou, G., Silva, B., Galdino, S., & Pires, J. (2010). Quantifying the sustainability impact of data center availability. *Performance Evaluation Review*, *37*(4), 64–68. https://doi.org/10.1145/1773394.1773405

Montoya, D., Rodríguez, M. A., Zavala, M. A., & Hawkins, B. A. (2007). Contemporary richness of holarctic trees and the historical pattern of glacial retreat. *Ecography*, *30*(2), 173–182. https://doi.org/10.1111/j.0906-7590.2007.04873.x

Okoli, C., & Schabram, K. (2012). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*, *10*(2010). https://doi.org/10.2139/ssrn.1954824

Pandey, D., & Agrawal, M. (2011). *Carbon footprint : current methods of estimation*. 135–160. https://doi.org/10.1007/s10661-010-1678-y

Panneerselvam, J., Liu, L., Hardy, J., & Antonopoulos, N. (2017). Analysis, Modelling and Characterisation of Zombie Servers in Large-Scale Cloud Datacentres. *IEEE Access*, *5*, 15040–15054. https://doi.org/10.1109/ACCESS.2017.2725898

Peñaherrera, F., Hobohm, J., & Szczepaniak, K. (2019). LCA of Energy and Material Demands in Professional Data Centers: Case Study of a Server. In *Sustainable Production, Life Cycle Engineering and Management*. Springer International Publishing. https://doi.org/10.1007/978-3-319-92237-9_9

Ren, C., Wang, D., Urgaonkar, B., & Sivasubramaniam, A. (2012). Carbon-aware energy capacity planning for datacenters. *Proceedings of the 2012 IEEE 20th International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunication Systems, MASCOTS 2012*, 391–400. https://doi.org/10.1109/MASCOTS.2012.51

Shah, A. J., & Meckler, M. (2009). *An exergy-based framework for assessing sustainability* *of IT systems*. 1–10.

Stutz, M., O'Connell, S., & Pflueger, J. (2012). Carbon footprint of a dell rack server. *Electronics Goes Green 2012+, ECG 2012 - Joint International Conference and Exhibition, Proceedings*, *710*.

Talens Peiró, L., Polverini, D., Ardente, F., & Mathieux, F. (2020). Advances towards circular economy policies in the EU: The new Ecodesign regulation of enterprise servers. *Resources, Conservation and Recycling*, *154*(November 2018), 104426. https://doi.org/10.1016/j.resconrec.2019.104426

U.S. EPA. (2014). *Pollution Prevention Greenhouse Gas (GHG) Calculator Guidance*. *October*, 1–5. https://www.epa.gov/sites/production/files/2014-12/documents/ghgcalculatorhelp.pdf

van Heddeghem, W., Vereecken, W., Colle, D., Pickavet, M., & Demeester, P. (2012). Distributed computing for carbon footprint reduction by exploiting low-footprint energy availability. *Future Generation Computer Systems*, *28*(2), 405–414. https://doi.org/10.1016/j.future.2011.05.004

Vereecken, W., Vanheddeghem, W., Colle, D., Pickavet, M., Dhoedt, B., & Demeester, P. (2012). The environmental footprint of data centers: The influence of server renewal rates on the overall footprint. *Lecture Notes in Electrical Engineering*, *113 LNEE*(January), 823–831. https://doi.org/10.1007/978-94-007-2169-2_98

Weber, C. L. (2012). Uncertainty and Variability in Product Carbon Footprinting: Case Study of a Server. *Journal of Industrial Ecology*, *16*(2), 203–211. https://doi.org/10.1111/j.1530-9290.2011.00407.x

Wiedmann, T., & Minx, J. (2008). A definition of "carbon footprint." *Ecological Economics Research Trends*, 1–11.

Wohlin, C. (2014). Guidelines for snowballing in systematic literature studies and a replication in software engineering. *ACM International Conference Proceeding Series*. https://doi.org/10.1145/2601248.2601268

Wohlin, C., Runeson, P., Höst, M., Ohlsson, M. C., Regnell, B., & Wesslén, A. (2012). Experimentation in software engineering. In *Experimentation in Software Engineering* (Vol. 9783642290). https://doi.org/10.1007/978-3-642-29044-2