Chapter 5

ENVIRONMENTAL IMPACT OF CLOUD COMPUTING: AN EXAMINATION ON ENERGY CONSUMPTION AND CARBON FOOTPRINT

Tolgahan Zorlu ÇELEBİ[1]

1. INTRODUCTION

To be able to reach as many users as possible and provide a flexible service the Internet providers created the Cloud Computing with requiring very few resources than on-site servers. Within following years, Cloud computing has quickly become the most popular technology in just a few years. From Google's release in 2003, to Amazon becoming available for use in 2006, and to AT&T Synaptic Hosting offering their services, cloud computing has advanced from being used only within a company's IT system to being a public service. It has also became a way to generate income by cost savings, and has expanded from being used by internet service providers to being used by telecommunications companies. (Ling Qian, Cloud Computing: An Overview)

The spread of high-speed networks has led to a worrying trend. The increase in network utilization, characterized by large numbers of concurrent e-commerce transactions and millions of daily web requests, is a significant phenomenon worthy of academic attention. The current increased demand is being met by extensive data centers in which hundreds and thousands of servers are integrated with additional infrastructure such as cooling facilities, storage units and network systems.

Several well-known Internet companies, like Google, Amazon, eBay, Facebook, Twitter and Yahoo, have set up numerous extensive data centers around the world. Cloud computing is the provision of virtualized data centers and applications as services. The operation of these units requires a significant amount of energy (Murugesan,.G.R.Gangadharan, 2012: 315-316) For example, a standard data center with 1000 racks requires around 10 megawatts (MW) of electricity for its operation, resulting in increased operating costs (Murugesan, G.R. Gangadharan,

[1] Software Developer, Detaysoft , tolgahan.celebi@detaysoft.com, ORCID iD: 0009-0006-1661-8261

2012: 315-316) The energy consumption for a data center is a significant factor in both the operating and acquisition costs. Additionally, an April 2007 Gartner report predicted that the information and communications technologies (ICT) sector accounts for approximately 2% of global carbon dioxide (CO2) emissions, a figure equivalent to that of the aviation industry. This study examines the environmental impact of cloud computing and compares the energy consumption and carbon fooprint of companies such as Google AWS Azur2. Understanding Cloud Computing

Cloud computing is a term that has been defined by various organizations, including Microsoft and the National Institute of Standards and Technology (NIST). Microsoft Azure (n.d.) describes cloud computing as the process of delivering various computing services over the internet, known as *"the cloud."* This includes services like servers, storage, databases, networking, software, analytics, and intelligence. The benefits of this model include rapid innovation, resource flexibility, and cost savings. With cloud computing, businesses typically pay only for the services they use, which can lead to reduced operating costs, more efficient infrastructure management, and the ability to scale based on business requirements (Microsoft Azure, n.d.).

On the other hand, NIST (n.d.) provides a more technical perspective. They define cloud computing as a model that offers on-demand network access to a shared and configurable pool of computing resources. These resources, such as networks, servers, storage, applications, and services, can be provisioned swiftly and require minimal interaction with service providers. NIST further breaks down the cloud model into five essential characteristics, three service models, and four delivery models (NIST, n.d).

2.1. Features of Cloud Computing:

Cloud computing has emerged as a transformative force in the digital landscape, underpinned by five essential features. Firstly, the **On-Demand Self-Service** aspect ensures that users can effortlessly access vital computing services such as server time and storage. This eliminates the need for manual intervention and direct communication with service providers. Instead, users can manage their cloud services through dedicated online portals, allowing them to monitor, modify, or even remove services as per their requirements, as highlighted by the NIST Cloud Computing Definition (NIST, n.d.).

Secondly, the concept of **Shared Resource Pooling** emphasizes the consolidation of multiple resources into a comprehensive pool. This approach is crucial for the scalability of cloud services, enabling users to adjust resources based on their needs. By centralizing both virtual and physical resources, the cloud ensures that users remain oblivious to the actual location of these resources, enhancing both user experience and security (Hiran, 2019).

Broad Network Access further amplifies the versatility of cloud computing. It ensures that cloud-stored resources are accessible across a myriad of devices, from smartphones to desktops. The expansive nature of cloud networks means resources are always within reach, provided there's an internet connection (cloudopedia, 2016).

The **Measured Service** feature underscores the efficiency of cloud services. Providers adopt a metered approach, billing users based on specific metrics such as hours of usage or data transferred. This data-driven approach ensures optimal resource allocation, allowing providers to meet user demands while maintaining cost-effectiveness (SSH Academy, 2023).

Lastly, **Rapid Elasticity** defines the adaptability of cloud services. They can be scaled up or down, sometimes even automatically, based on user demand. This dynamic nature of cloud services ensures users can tailor their service usage and costs without being bound by rigid contracts or incurring extra charges. The cloud's flexibility eliminates the need for direct hardware investments, allowing users to leverage the expansive resources of their cloud provider, as defined by the NIST Cloud Computing Definition (NIST, n.d.).

2.2. Service Models for a Cloud Computing System

Cloud computing refers to a broad concept where services are provided through the internet. These services are divided into three main categories or types of cloud computing: infrastructure as a service (IaaS), platform as a service (PaaS) and software as a service (SaaS).

Cloud computing has revolutionized the way businesses and individuals access and store data. As technology continues to evolve, various service models have emerged to cater to the diverse needs of users. These models offer different levels of control, flexibility, and management, depending on the specific requirements of the user. Let's delve into the three primary service models of cloud computing to understand their distinct features and applications.

There are basically three service models for a cloud computing system, namely, Infrastructure-as-a-Service, Platform-as-a-Service and Software-as-a-Service.

Software as a Service (SaaS) is a model where applications are delivered to users through web browsers. It employs a multi-tenant architecture, meaning multiple users share the same resources. For users, this translates to no upfront costs for server infrastructure or software licenses. Providers benefit from the economies of scale, as maintaining a single application is more cost-effective than traditional hosting. Salesforce.com stands as a prominent example of SaaS in the enterprise domain. Moreover, SaaS applications, such as Workday or even sophisticated platforms like SAP/HANA and ORACLE NETSUIT, have become prevalent in sectors like human resources and enterprise resource planning, offering features pertinent to various academic and business contexts (Knorr & Gruman, 2008).

Platform as a Service (PaaS) offers a development environment as a service. Here, users can craft applications using the provider's tools and infrastructure. These custom applications are then made available to end-users via the internet. While PaaS might impose certain design limitations, akin to building with a specific set of Lego pieces, it ensures reliability and seamless integration (Knorr & Gruman, 2008).

Lastly, **Infrastructure as a Service (IaaS)** provides users with fundamental computing resources, including processing power, storage, and networking capabilities. Users have the freedom to deploy any software, from operating systems to specific applications, on this infrastructure. However, they don't control the underlying cloud infrastructure. Instead, they manage the software, storage, and certain networking components (NIST definition of Cloud Computing, 2011).

2.3. Deployment Models

Cloud computing, with its transformative impact on technology infrastructure, offers a range of solutions tailored to the diverse needs of businesses and individuals. The evolution of cloud services has led to the development of distinct deployment models, each designed to address specific requirements related to ownership, accessibility, and functionality. Delving into the core of cloud computing, we find four primary deployment models:

The **Private Cloud** is tailored for a single organization, catering to its multiple internal consumers, such as different business units. While the infrastructure is exclusive to the organization, its management, operation, and ownership can be either internal, outsourced to a third-party, or a combination of both. Furthermore, the infrastructure's physical location can vary, being either on-premises or off-site (NIST definition of Cloud Computing, 2011).

The **Community Cloud** is designed for a specific group of users from organizations with shared objectives or concerns, such as similar security requirements or compliance needs. The infrastructure's ownership and management can be undertaken by one or more organizations within the community, an external entity, or a blend of both. Its location can also be flexible, being either on-site or off-site. (NIST definition of Cloud Computing, 2011).

In contrast, the **Public Cloud** offers its infrastructure for open use by the general public. This model can be owned and operated by various entities, including businesses, academic institutions, or government agencies. The infrastructure itself is housed within the premises of the cloud service provider.

Lastly, the **Hybrid Cloud** is a composite model, encompassing multiple distinct cloud infrastructures, such as private, community, or public clouds. These entities are interconnected, either through standardized or proprietary technologies, facilitating the smooth transfer of data and applications. An example of this is cloud bursting, which ensures optimal load distribution across clouds. However, it's crucial to note that even though these clouds are interconnected, each retains its distinct identity (NIST definition of Cloud Computing, 2011).

3. ENERGY CONSUMPTION IN CLOUD COMPUTING

Cloud infrastructure can be described as virtualized data centers and applications available as subscription-based services, The operation of this unit requires a significant amount of energy

As such Cloud infrastructure can be described as virtualized data centers and applications available as subscription-based services. The operation of this unit requires a significant amount of energy. For example a typical thousand-rack data center, requires approximately 10 megawatts of power to maintain operations (Riviore, 2007).

Therefore, expenses related to energy consumption represent a significant portion of both the upfront and ongoing costs of a data center. At the same time, this high amount of energy consumption creates environmental impacts.

Furthermore, a study conducted by Gartner in April 2007 estimated that the information and communications technologies (ICT) sector is responsible for approximately 2% of the total global release of carbon dioxide (CO_2) into the atmosphere, a figure similar to CO_2 -Footprint corresponds to the aviation industry. (Green Peace, 2010). In conclusion the problem of energy consumption and carbon emissions in cloud infrastructure has become a significant environmental concern (Riviore, 2007).

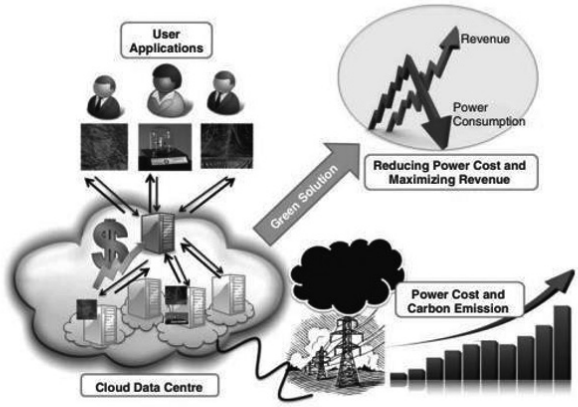


Figure 1: Cloud and environmental sustainability

Company Scorecard

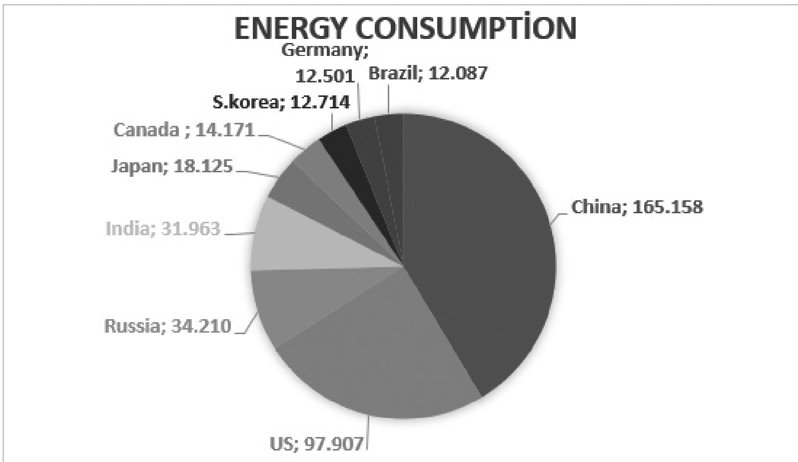
Company	Clean Energy Index	Coal	Nuclear	Energy Transparency	Infrastructure Siting	Energy Efficiency & GHG Mitigation	Renewables & Advocacy
Akamai	NA	NA		A	C	B	D
amazon.com	13.5%	33.9%	29.9%	F	F	D	F
Apple	15.3%	55.1%	27.8%	D	F	D	D
DELL	56.3%	20.1%	6.4%	C	C	C	D
facebook	36.4%	39.4%	13.2%	D	B	B	C
Google	39.4%	28.7%	15.3%	B	C	B	A
hp	19.4%	49.7%	14.1%	C	D	B	C
IBM	12.1%	49.5%	11.5%	C	D	C	D
Microsoft	13.9%	39.3%	26%	C	D	C	C
ORACLE	7.1%	48.7%	17.2%	D	D	C	D
rackspace HOSTING	23.6%	31.6%	22.3%	C	C	C	C
salesforce	4%	33.9%	31%	B	C	C	C
twitter	21.3%	35.6%	12.8%	F	D	F	D
YAHOO!	56.4%	20.3%	14.6%	C	B	B	B

Figure 2: Energy consumption data from the leading cloud computing service providers. Segmented by energy production type (Greenpeace international, 2012)

3.1. Power Consumption

Determining the amount of energy required to sustain the ever-expanding digital sphere, as well as measuring the IT industry's proportional contribution to global greenhouse gas (GHG) emissions, represents a daunting challenge. The factors responsible for this challenge include the rapid and expansive development of the sector, the wide range of devices and energy delivery mechanisms involved, as well as the frequent and dynamic changes in both technology and business methodology in this area. Precise answers to these inquiries cannot yet be found. However, a key factor contributing to this inaccuracy is the lack of transparency observed among leading IT companies. The phenomenon of collective confidentiality in the cloud computing industry appears to be driven both by fears of exposing operational competitive advantages or disadvantages between leading cloud providers and by the motivation to change the narrative about the sector's dependence on non-renewable energy sources such as electricity. B. to suppress fueling coal and other environmentally harmful forms in order to maintain a positive image among the public and employees. (Greenpeace international, 2012).

Figure 3: Total electricity consumption quadrillion kWh (US Energy information Administration, 2021)



Graphic 1: Top countries that consume the most electricity worldwide. (U.S. Energy Information Administration, 2021)

The table above lists the countries with the highest electricity consumption in the world in quadrillions per kWh. But the data that, makes this data meaningful to us is the energy consumption data of cloud systems as a result of research conducted by Greenpeace in 2007. Cloud systems overall outpaced many countries around the world in 2007, consuming 623 billion of energy per kilowatt hour (in billions of kilowatt hours per hour) and ranking on the list of highest electricity consumption in the world after America, China, Russia and Japan. Despite implementing efficient data center designs with exceptionally high utilization rates, such measures can only serve to mitigate harmful carbon dioxide emissions rather than completely eliminate them. This situation can be further examined by analyzing Figure 2. Figure 2 classifies the energy sources used by leading cloud service providers, categorized based on their procurement methods. The use of carbon-based energy sources, nuclear energy and green energy is examined in this analysis. Many companies are currently using carbon-based energy sources instead of those that come from the environment. The present observation can be attributed to the fact that cloud providers prioritize reducing electricity costs over reducing CO2 emissions. It is obvious that none of the cloud data centers listed can be considered environmentally sustainable.

Unfortunately, the most recent data available for this study was for 2007. Even comparing this old data with the current data reveals the terrible reality.

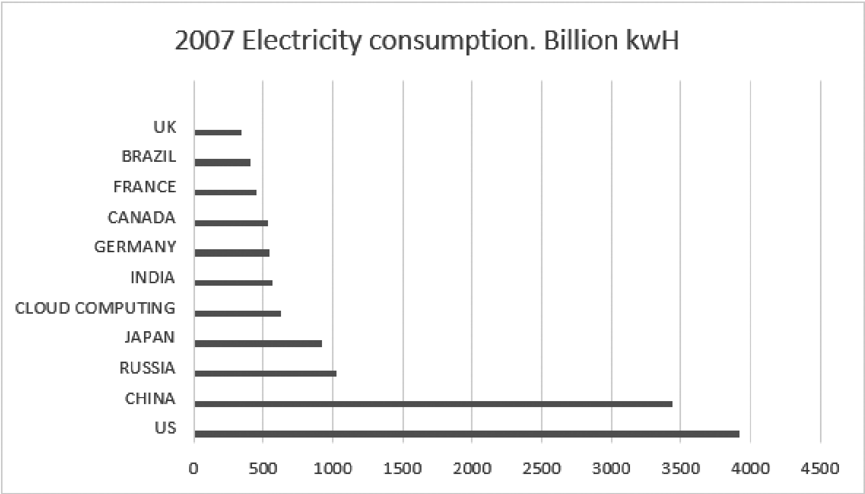


Figure 4: Chart showing energy consumption data worldwide, including cloud computing and countries. (Greenpeace international, 2012)

One of the key factors affecting energy consumption is the formulation and execution of software applications. Cloud computing has the ability to accelerate the delivery of applications that can impact either the end user or the cloud provider through the use of Software-as-a-Service (SaaS). The way energy is used in a given scenario depends on the type of specific application being implemented. If an application has long execution times and requires high CPU and memory usage, it is likely to have a significant impact on the energy consumption during its operation. It can therefore be assumed that there is a direct connection between energy consumption and the application profile. Optimizing resource allocation based solely on maximum CPU and memory utilization is expected to result in a significant increase in energy utilization beyond required levels. The apparent lack of energy efficiency in running applications can be attributed to poor design and implementation techniques. The present study shows that the proliferation of inefficiencies in applications, typically due to the use of suboptimal algorithms and inefficient allocation of shared resources, leads to increased CPU utilization and a consequent increase in energy consumption.

However, it is often observed that the integration of energy efficiency into the application design process is not given sufficient priority. The only notable exception is in certain areas such as embedded devices (e.g. the necessary electronic devices, commonly known as portable electronic devices). (Greenpeace international, 2010)

In 2020, the total amount of carbon dioxide equivalent (CO_2 equivalent) emissions generated by data centers and data transmission networks, including their gray emissions, was estimated at around 300 million tonnes. This figure represents 0.9% of energy-related greenhouse gas emissions or 0.6% of total greenhouse gas emissions, underscoring the important role of these facilities and networks in enabling digitalization. Since 2010, despite rapidly increasing demand for digital services, emissions have recorded a slight increase due to the adoption of energy efficient practices, the procurement of renewable energy by information and communications technology (ICT) companies and the extensive decarbonization programs of electricity grids in a variety of regions. (IEA Analysis, 2022). However, to meet the net zero scenario, it is imperative that emissions be reduced by fifty percent by 2030. Global data center electricity consumption was an estimated range of 220–320 TWh in 2021, representing approximately 0.9-1.3% of total global electricity demand. The stated value explicitly does not take into account the energy consumption attributable to cryptocurrency mining, which was estimated to be in the order of 100 to 140 terawatt hours in 2021. Since 2010, energy usage in data centers, excluding cryptocurrencies, has experienced a moderate increase, despite the significant acceleration in demand for data center services. This is largely due to advances in IT hardware and cooling technologies used to maximize efficiency, as well as the transition from less powerful small business data centers to more modern cloud and hyperscale data centers. (IEA Analysis, 2022).

The increasing demand for processing capacity in large data centers has led to a sharp increase in energy consumption in this area in recent years, with increases ranging between 10% and 30% per year. Total energy consumption of data centers, excluding cryptographic applications, is expected to show a slight increase in the coming years, although there is considerable uncertainty regarding long-term trends. Despite the moderate increase in global data center power consumption, some smaller countries with emerging data center markets are experiencing significant growth rates. (IEA Analysis, 2022).

In 2021, Apple, Google and Meta procured or generated a significant amount of renewable electricity, namely 2.8 terawatt-hours, 18.3 terawatt-hours and 9.4 terawatt-hours, respectively, allowing these companies to meet 100% of their operational needs To cover electricity needs, especially in their data centers, using renewable energy sources. In 2021, Amazon used a total of 30.9 terawatt hours (TWh) of energy, 85% of which came from renewable sources. The company has set itself the goal of switching to 100% renewable energy by 2025. Achieving full compliance between annual energy needs and the procurement of renewable energy or certificates does not necessarily mean that data centers are fully reliant on renewable energy sources. The irregularities associated with wind and solar sources may not match a data center's demand profile, resulting in renewable energy being sourced from projects located in different regions or grids distant from the data center's location. Renewable energy certificates are unlikely to facilitate the production of additional renewable energy, leading to inflated assessments of noticeable emissions reductions in the existing world. Google and Microsoft recently announced their 2030 goals to procure and synchronize zero-carbon electricity 24/7 on every grid where demand exists. A variety of companies are striving to obtain carbon-free energy 24/7. While a select number of network operators, such as BT, TIM and T-Mobile, have successfully achieved a 100% renewable energy portfolio, it is widely recognized that data transmission network operators generally lag behind data center operators in terms of procurement and adoption of sustainable energy practices. When comparing telecom operators and data centers, it is important to note that the latter tend to be larger and easier to relocate due to their centralized nature, while the former tend to have many smaller locations with limited site selection options. The quest to gain access to renewable energy is seen as a challenge in many markets, particularly in emerging and developing countries with underdeveloped energy markets.

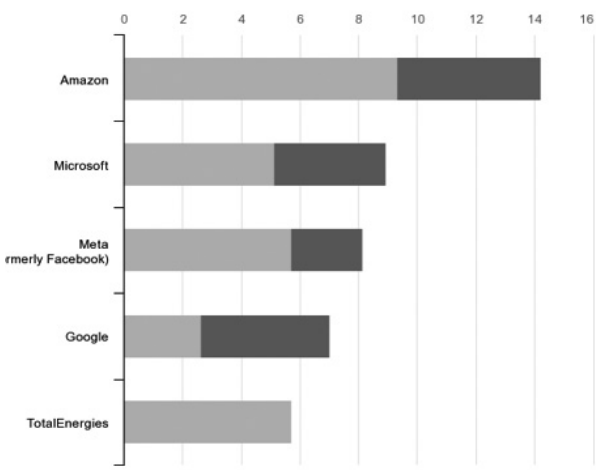


Figure 5: The five largest corporate buyers of power purchase agreements for renewable energies, 2010–2021

4. CARBON FOOTPRINT OF CLOUD COMPUTING

As the world becomes increasingly digital, the environmental impact of our technological advancements has come under scrutiny. One area that has garnered significant attention is the carbon footprint associated with cloud computing. This section delves into the intricacies of the carbon emissions linked to cloud computing, exploring its definition, measurement, and broader implications for our planet.

4.1. Exploring the Concept of Carbon Footprint

The term "carbon footprint" refers to the comprehensive total of carbon dioxide (CO2) emissions resulting from the actions of an individual or other entity, including structures, companies and nations. The above represents a broad area that includes both the direct release of emissions resulting from the combustion of fossil fuels used in manufacturing, heating and transportation, as well as the emissions associated with electricity generation Goods and amenities required is consumed. In addition, as part of carbon footprint analysis, it is often necessary to consider emissions of other powerful greenhouse gases such as methane, nitrous oxide, and chlorofluorocarbons (CFCs). (Britannica, 2023)

The concept of the CO2 footprint is closely linked to and derived from the previous idea of the ecological footprint. The concept was first introduced in the early 1990s by William Rees, a Canadian ecologist, and Mathis Wackernagel, a regional planner of Swiss origin who had a connection to the University of British Columbia. The term "Ecological footprint" refers to the cumulative terrestrial extent required to sustain a given entity or community of individuals in terms of resource use and the resulting environmental impacts. The scope of this matter covers a wide range of ecological impacts, including the use of water resources and the size of the land allocation dedicated to food production. In contrast, a weight-based metric is often used to quantify the carbon footprint, in tons of carbon dioxide or carbon dioxide equivalent, within a specified annual time frame. (Berl, 2010)

4.2. Measuring Cloud Computing Carbon Footprint

The emissions associated with cloud computing are often overlooked. Instead of reading these reports, we usually focus on popular activities like watching Instagram Reels videos, sharing photos, or sharing online assets.

According to analysis by The Shift Project, a non-profit organization based in Paris, streaming content on Netflix for half an hour produces about 1.6 kg.

of carbon dioxide emissions, the equivalent of traveling nearly 6 kilometers (4 miles) by car. Current studies indicate that current production is currently likely to be significantly lower. According to a report by Digiconomist, the annual CO2 emissions attributed to the Bitcoin mining process are on par with New Zealand. Specifically, the platform produces around 36.95 tons of carbon dioxide per year. According to the University of Cambridge, the energy consumption associated with cryptocurrencies exceeds the total annual energy consumption of the Netherlands. When analyzing all the emissions associated with cloud computing, certain comparative assessments are even more startling. The current share of greenhouse gas emissions caused by cloud computing is 2.5% to 3.7% of global emissions, exceeding emissions associated with commercial flights, which account for approximately 2.4%. Furthermore, this highlights the extent of the contribution that cloud computing makes to the overall emissions of various economic activities. (Lavi, 2022)

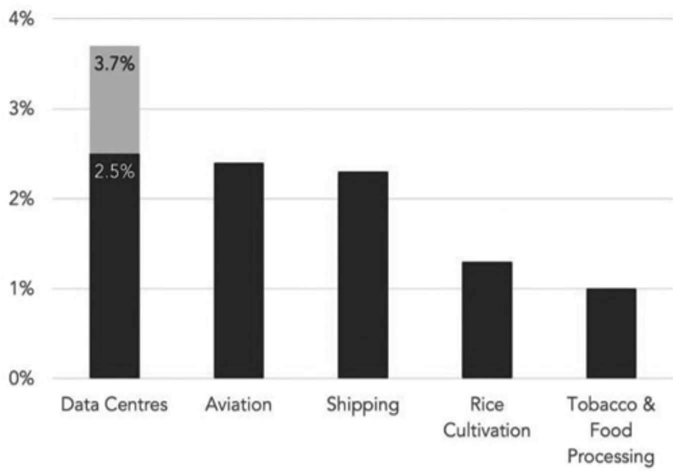


Figure 7: Global carbon emissions comparison by industry. (Lavi,2020)

It is estimated that data centers use around 200 terawatt hours (TWh) of energy annually. The above amount exceeds the combined energy consumption of several nations, including Iran. Furthermore, it represents half of the electricity used internationally for transportation, while accounting for only a minimal fraction of total global electricity demand (as indicated in the graph labeled “Energy Scale”). Amid concerns about the likelihood of impending energy demands, scientists at various scientific research institutions and engineers working for some of the wealthiest corporations around the world are currently exploring strategies to mitigate the environmental impact of data centers. Today's organizations are actively seeking to improve computing practices by implementing optimized processes, transitioning to renewable energy sources, and exploring advanced approaches to enable efficient cooling of data centers while promoting the recycling of residual thermal energy. Eric Masanet, an engineer at Northeastern University in Evanston, Illinois and co-author of a respected International Energy Agency (IEA) report on digitalization and energy, has emphasized the imperative need for conscientious regulation of energy consumption associated with information and communications technology (ICT). Regardless, Masanet believes that a vigilant approach to the problem could lead to competent management of energy consumption in the coming years. (Jones, 2018)

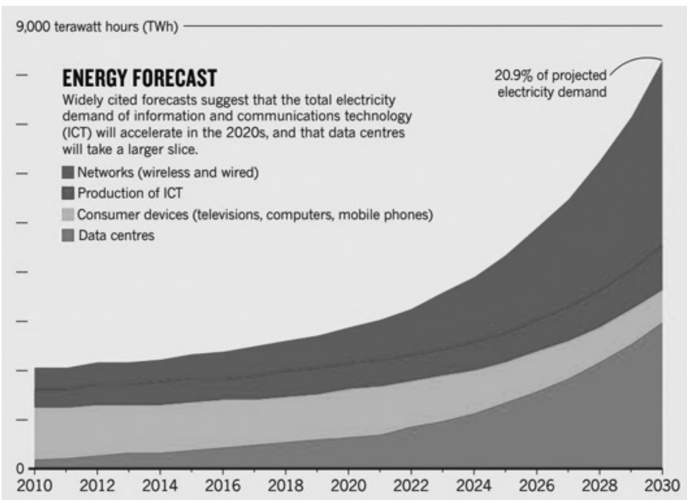


Figure 8: "Expected Case" projection by Anders Andrae, a specialist in sustainable ICT. (Jones, 2018)

5. GREEN CLOUD COMPUTING: TOWARDS SUSTAINABLE SOLUTIONS

While there is a heightened level of concern within the community about the potential increase in data center energy consumption due to cloud computing, it is important to recognize that there is a sustainable component to this technology. Cloud service providers use a variety of technologies and concepts to achieve improved utilization and efficiency beyond that of traditional data processing. One could argue that the use of cloud computing has the potential to reduce carbon emissions compared to traditional computing techniques. The real reason for this occurrence lies in the existence of an infrastructure capable of saving energy, coupled with the possibility of limiting the entire IT infrastructure by implementing multi-tenancy. However, since cooling the servers requires a high amount of energy, virtualization systems were implemented. The implementation of virtualization represents a crucial technological enabler for energy-efficient clouds. By leveraging economies of scale, this technological approach enables significant improvements in the energy efficiency of cloud providers. This involves the shared use of the infrastructure by a large number of units. The concept of virtualization encompasses the manifestation of a discrete logical aggregation or subset of computing resources that enables the achievement of benefits beyond those accessible with the original configuration.

By implementing server consolidation tactics, organizations can achieve significant financial benefits in terms of physical space, energy usage, and management overhead. This method involves consolidating underutilized servers by creating and deploying multiple virtual machines (VMs) running on a single physical server to improve overall server resource utilization. According to the findings of a report published by Accenture, the adoption of cloud computing has been facilitated by four key factors that have led to a decrease in energy consumption and carbon emissions in the information and communications technology (ICT) sector. (accenture, 2020):

5.1. Dynamic Deployment

In traditional contexts, data centers and private infrastructures have historically been maintained to meet the needs of the most extreme scenarios. As a result, information technology (IT) companies often resort to deploying excess infrastructure that exceeds their operational needs. There are numerous factors that contribute to the phenomenon of oversupply.

The complexity of predicting demand at a given point in time, particularly for web-based applications, as well as the requirement to ensure availability of services and maintain certain standards of service quality for end users are among the key reasons. To meet the challenge of meeting increased demand every year within a short period of time, maintaining a fleet of numerous servers throughout the year lacks the energy efficiency. Such approaches are another mistake in increasing energy consumption and causing environmental impacts. Unfortunately, companies resort to these methods in order to make profits and due to the lack of controls. For such scenarios, cloud infrastructure offers an easyto-manage solution. Data centers consistently manage their servers according to current demand, resulting in minimized energy consumption compared to the more traditional overprovisioning approach. This approach guarantees that only the necessary servers are active, resulting in significantly more efficient use of energy. (Murugesan,.G.R.Gangadharan, 2012)

5.2. Multi-Tenancy

By using multi-tenancy methodology, the cloud computing infrastructure effectively reduces overall energy consumption and the resulting CO2 emissions. Software-as-a-Service providers offer their services to multiple companies on a common infrastructure and software platform. This method is obviously more resource-efficient than installing numerous software copies on different infrastructure. But in business, demand patterns exhibit a significant degree of variability. With this in mind, implementing multi-tenancy on a single server can effectively balance peak loads and reduce the need for additional infrastructure. Lower demand fluctuations result in greater accuracy in demand forecasting and better energy saving results. (Murugesan ,.G.R.Gangadharan, 2012)

5.3. Server Utilization

As a rule, the on-site infrastructure has a significantly lower level of utilization, which is occasionally up to 5-10% below average utilization. The use of virtualization technologies allows hosting and running numerous applications on the same server while maintaining isolation, which in turn results in utilization rates of up to 70%. This results in a significant reduction in the number of actively involved servers. Despite the fact that increased server utilization is associated with increased power consumption, servers with high utilization are capable of processing larger workloads while consuming similar amounts. (Harnessing green IT)

5.4. Data Center Efficiency

By using the most energy efficient technologies, cloud service providers can significantly increase the Power Usage Effectiveness (PUE) of their data centers. Contemporary data center designs for major cloud service providers are capable of achieving a power consumption effectiveness ratio (PUE) of just 1.1-1.2, representing an approximately 40% improvement in energy efficiency compared to traditional data centers. Several strategies have been implemented in modern data centers to improve their power consumption effectiveness (PUE). These methods include the use of modular containers for server design and sophisticated cooling techniques using air or water.

Additionally, advanced power management techniques, including power supply optimization, have demonstrated notable PUE improvements. Additionally, the implementation of cloud computing technology makes it easier to move services to different data centers, all with improved Power Usage Effectiveness (PUE) metrics. Achieving this goal is achieved through the use of high-speed network infrastructures, virtualized services and the systematic quantification, monitoring and allocation of resources within the data center. Due to the unique characteristics of cloud infrastructure, it is possible for companies to reduce their carbon emissions by at least 30% per user by migrating their applications to the cloud. The above savings are due to the exceptional efficiency of large-scale cloud-based data centers. (Harnessing green IT)

Despite the progress that researchers and developers have made in optimizing specific elements of cloud computing for energy and performance efficiency, there remains a lack of comprehensive consideration of the environmental impact of this technology. Many efforts to ensure the sustainability of cloud computing have not taken into account the role of the network in this phenomenon. Certain studies are primarily concerned with workload distribution to enable energy efficient cooling without considering the impact of virtualization. Additionally, cloud providers are for-profit companies looking for viable solutions to reduce power consumption and thus reduce carbon emissions without compromising their market position. To address certain concerns, we propose a sustainable cloud framework that aims to take into account providers' goals while limiting the energy consumption of cloud systems. The diagram representing the overarching perspective of the green cloud computing architecture is shown in Figure 6. The goal of this architectural framework is to promote an environmentally friendly cloud ecosystem that takes into account the concerns and interests of both users and providers. (Harnessing green IT)

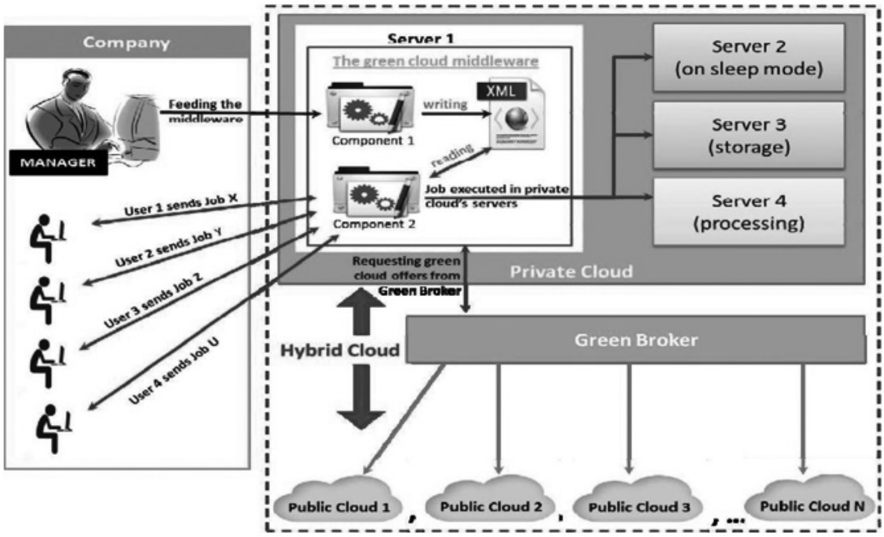


Figure 6: Green cloud architecture. (Integrated Green Cloud Computing Architecture)

Cloud providers have the opportunity to display their services as "ecological offers" in a publicly accessible directory. This directory is then retrieved by a green broker. The range of eco-friendly offerings includes a range of environmentally sustainable services, complemented by pricing structures and recommended time frames for using these services to reduce the release of carbon emissions. The current status of energy parameters related to the use of various cloud services was determined by Green Broker through consultation with the Carbon Emission Directory. The Carbon Emission Directory acts as a data reservoir for the energy efficiency of cloud services. The aforementioned details relate to two key aspects of the cloud data center service, namely its Power Usage Effectiveness (PUE) and cooling efficiency, as well as the costs associated with network usage. In addition, the analysis includes measuring the carbon emissions generated during electricity generation. Green Broker conducts a comprehensive assessment to determine the carbon footprint attributable to the provision of desired cloud-based services by relevant cloud service providers. The system then detects and retrieves a set of services that would result in the lowest carbon emissions. These services are then purchased on behalf of the end users. (Data Centres and Data Transmission Networks, IEA, 2022)

The main goal of the developed green cloud framework is to facilitate and maintain consistent monitoring and assessment of the total energy consumption resulting from the execution of user requests. The approach under consideration is based on two main components, namely the Carbon Emission Directory and the ecologically sustainable cloud services. These components serve the function of closely monitoring each cloud provider's energy efficiency benchmarks while providing significant incentive for cloud providers to promote the adoption of environmentally friendly practices in their offerings. Green Broker's central role is to monitor and select cloud services that meet the user's Quality of Service (QoS) requirements while ensuring that the carbon footprint associated with the user's service is minimized. In general, cloud computing users can effectively utilize various offerings such as Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). Therefore, providing these services requires a significant level of energy efficiency. Essentially, it is critical for each layer of the cloud infrastructure to adopt environmentally sustainable practices towards the cloud service provider. (Murugesan, G. R. Gangadharan, 2012)

SaaS: Because SaaS providers primarily provide software installed in their own data centers or source resources from IaaS providers, it is essential for these providers to develop models and metrics to measure the energy efficiency of their software design, implementation and to evaluate the use. The SaaS provider selects data centers that are not only energy efficient but also located in close proximity to its user base to optimize service delivery. Maintaining a sufficient number of replications of users' sensitive data is imperative, with an emphasis on energy-efficient storage solutions.

PaaS: PaaS providers typically offer a wide range of services tailored to application development needs and include a range of solutions. The platform acts as an intermediary to enable the progress of applications that aim for energy efficiency across the system. One strategy to achieve this goal involves the use of a range of energy profiling tools, such as JouleSort, among others. The above application serves as a benchmark for measuring energy efficiency by measuring the energy consumed during the execution of an external sorting operation. Additionally, there is an opportunity for platforms to incorporate various code-level optimizations to improve synergy with underlying compilers, making it easier to run applications in an energy-efficient manner. Cloud platforms have the potential to enable the deployment of user-centric applications in hybrid cloud environments while streamlining the application development process. To increase energy efficiency, the platforms use an application profiling mechanism. This requires identifying different components of the application or data to be processed as needed within the organizational spaces and within the cloud environment.

IaaS: Companies operating at this level play a crucial role in the overall success of green architecture. The above claim can be attributed to the fact that the Infrastructure as a Service (IaaS) layer not only offers independent infrastructure services but also enables the other services provided by cloud solutions. Advanced technological solutions have been deployed to improve the energy efficiency of the infrastructure in terms of information technology and cooling systems. The use of virtualization and consolidation has resulted in a reduction in energy consumption by deactivating idle servers. Multiple energy meters and sensors are integrated to determine the current energy efficiency of each Infrastructure as a Service (IaaS) provider and their corresponding facilities. The above data is often advocated by cloud service providers listed in the Carbon Emission Directory. Numerous green planning and resource allocation guidelines have been developed to ensure minimal energy consumption. In addition, the cloud service provider employs a variety of environmentally friendly initiatives and pricing strategies to encourage users to use its services during times of reduced energy consumption or during peak traffic times.

5.5. Environmental Cooling

As previously noted, the predominant factor contributing to the elevated energy consumption in Cloud Centers is the cooling mechanism. The energy demand for cooling intensifies, leading to a larger carbon footprint, as the utilized hardware generates heat. Contemporary enterprises have adopted diverse cooling strategies for their data centers, contingent upon their geographical locations, as illustrated in Figure 9.

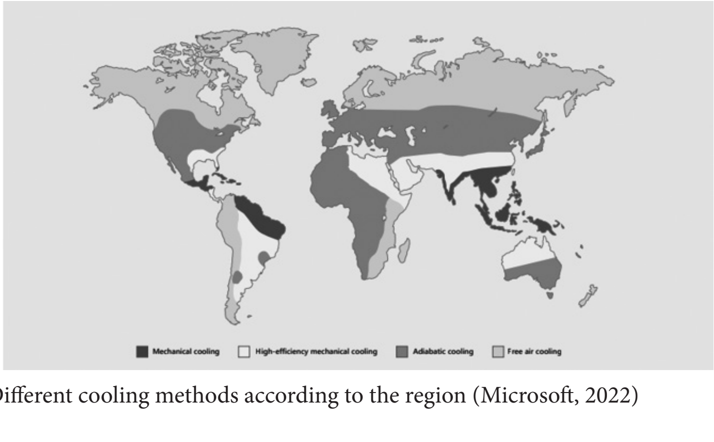


Figure 9: Different cooling methods according to the region (Microsoft, 2022)

Below, the carbon footprints of data centers belonging to 3 different companies are shown based on energy consumption data: Amazon AWS, Microsoft Azure and Google Cloud Platform.

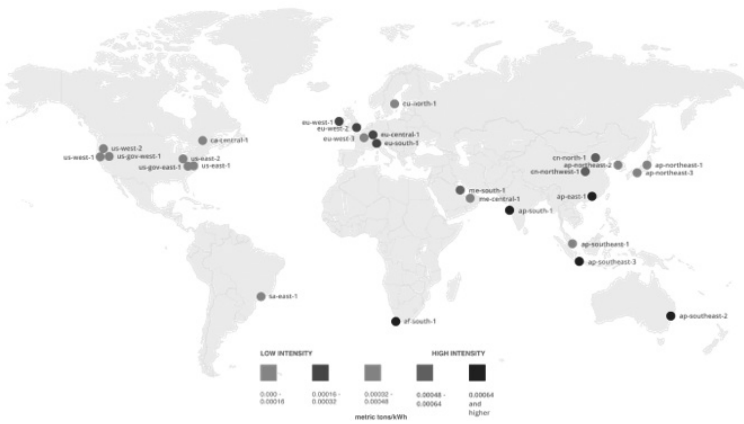


Figure 10: Amazon AWS carbon intensity map. (ccfp, 2023)

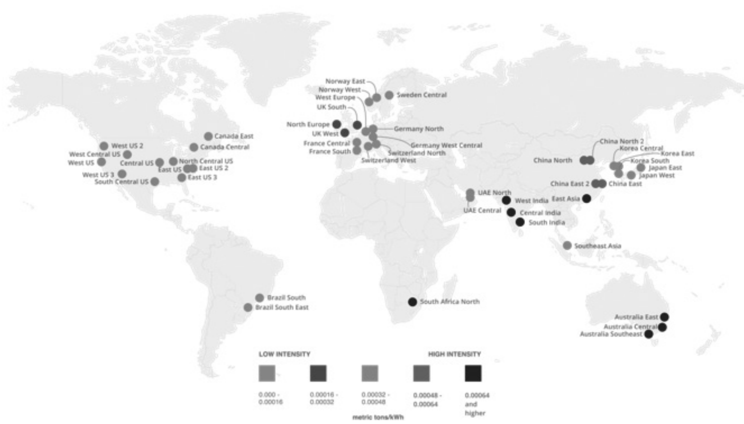


Figure 11: Microsoft Azure map of carbon intensity. (ccfp, 2023)

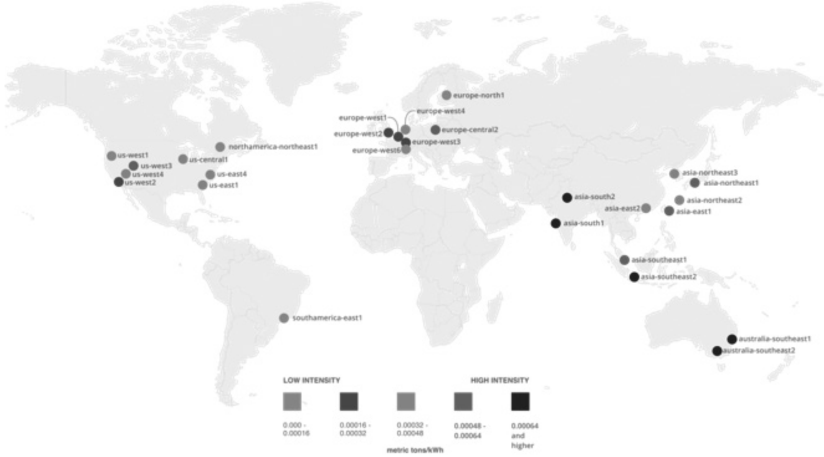


Figure 12: GCP map of carbon intensity. (ccfp, 2023)

It is shown in figure 10, figure 11 and figure 12 that, the carbon footprint of 3 different companies according to their electricity consumption in their data centers in various parts of the world. It is clearly seen that data centers located in warmer parts of the world consume more electricity, regardless of the company, due to the warmer environment. When we compare different companies in the same region, it can be seen that the main reason for this consumption is the energy spent for cooling. Based on this, using the earth itself to cool data centers is a great solution to reduce energy consumption. Environmental cooling can be achieved by locating these data centers in regions close to the poles. In this way, locating data centers in places where the environment is cold will cause serious decreases in the amount of energy used, and this decrease will reduce the carbon footprint.

6. RESULT

In the digital age, the environmental ramifications of our technological advancements have become a focal point of discussion. One such area that has been thrust into the limelight is the carbon footprint associated with cloud computing. The document delves deep into the intricacies of carbon emissions linked to this domain, offering insights into its definition, measurement, and broader implications for our planet.

The term "carbon footprint" encapsulates the total carbon dioxide (CO2) emissions resulting from the activities of entities ranging from individuals to nations. This encompasses both the direct emissions from fossil fuel combustion in processes like manufacturing and transportation, as well as those tied to electricity generation. Furthermore, the analysis often extends to emissions of other potent greenhouse gases, including methane, nitrous oxide, and chlorofluorocarbons (CFCs).

A pivotal revelation from the document is the discernible pattern in the carbon footprint of data centers across different global regions. Data centers in warmer regions tend to consume more electricity, predominantly due to the ambient temperatures and the energy expended on cooling. This insight underscores the potential of leveraging the Earth's natural cooling mechanisms as a sustainable solution to mitigate energy consumption in these centers.

Moreover, the document highlights the adaptability and efficiency of cloud services. With features like Rapid Elasticity, cloud services can dynamically scale based on user demand, ensuring optimal resource allocation and cost-effectiveness. This flexibility, combined with the metered approach of billing, positions cloud computing as a sustainable and efficient technological solution.

In essence, as cloud computing continues to evolve, understanding its environmental impact becomes paramount. The insights from this document underscore the need for sustainable practices and innovations to ensure that the growth of cloud computing aligns with our environmental objectives.

REFERENCES

accenture. (2020). The green behind the cloud. https://www.accenture.com/. https://www.accenture.com/content/dam/accenture/final/a-com-migration/manual/r3/pdf/pdf-135/Accenture-Strategy-Green-Behind-Cloud-POV.pdf

Cloud carbon footprint. (n.d.). https://demo.cloudcarbonfootprint.org

Cloudopedia. (2023). What is Broad Network Access? Cloudopedia. https://cloudopedia.com/broad-network-access/

Final version of NIST Cloud Computing Definition published | NIST. (2018). NIST. https://www.nist.gov/news-events/news/2011/10/final-version-nist-cloud-computing-definition-published

Greenpeace International. (2012). How Clean is Your Cloud? https://www.greenpeace.org/static/planet4-international-stateless/2012/04/e7c8ff21-howcleanisyourcloud.pdf

Greenpeace International. (n.d.). Make IT Green. Make IT Green. https://www.greenpeace.org/static/planet4-international-stateless/2010/03/f2954209-make-it-green-cloud-computing.pdf

Harnessing Green IT: Principles and Practices | IEEE eBooks | IEEE Xplore. (n.d.). https://ieeexplore.ieee.org/book/6542356

Hiran, K. K. (2019, May 1). Cloud Computing: Concepts, Architecture and Applications with Real-world examples and Case studies. Aalborg University's Research Portal.

https://vbn.aau.dk/en/publications/cloud-computing-concepts-architecture-and-applications-with-real-

Integrated Green Cloud Computing Architecture. (2012, November 1). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/6516364

International - U.S. Energy Information Administration (EIA). (n.d.). https://www.eia.gov/international/rankings/world?pa=44&u=2&f=A&v=heatmap&y=01%2F01%2F2021&ev=true

Jones, N. (2018). How to stop data centres from gobbling up the world's electricity. Nature, 561(7722), 163–166. https://doi.org/10.1038/d41586-018-06610-y

Lavi, H. (n.d.). Measuring greenhouse gas emissions in data centres: the environmental impact of cloud computing | Insights & Sustainability | Climatiq. https://www.climatiq.io/blog/measure-greenhouse-gas-emissions-carbon-data-centres-cloud-computing

Microsoft Azure. (n.d.). What is Cloud Computing. https://azure.microsoft.com/en-us/resources/cloud-computing-dictionary/what-is-cloud-computing

National Institute of Standards and Technology. (n.d.). The NIST Definition of Cloud Computing. https://www.govinfo.gov/. https://www.govinfo.gov/content/pkg/GOVPUB-C13-74cdc274b1109a7e1ead7185dfec2ada/pdf/GOVPUB-C13-74cdc274b1109a7e1ead7185dfec2ada.pdf

Rivoire, S., Shah, M. A., Ranganathan, P., & Kozyrakis, C. (2007). JouleSort. A Balanced Energy-efficiency Benchmark. https://doi.org/10.1145/1247480.1247522

Selin, N. E. (2023, September 21). Carbon footprint | Definition, Examples, Calculation, Effects, & Facts. Encyclopedia Britannica. https://www.britannica.com/science/carbon-footprint