HIGH SCHOOL EDITION
Journal of Student Research
Volume 12 Issue 4 (2023)

Future of Blockchain: Data Storage, Carbon Calculation, Accounting, and Emissions Trading

Melissa Fan

Conestoga High School

ABSTRACT

The paper explores various aspects of data storage technologies, including Solid-State Storage, Cloud Computing, Edge Computing, and Blockchain, in the context of their implications, advantages, and limitations. Each technology's impact on data storage, performance, security, scalability, and environmental considerations is examined. Carbon Footprint (CF) is introduced as a measure of the environmental impact of data storage methods. The paper then delves into carbon calculation approaches, emphasizing the importance of accurate measurement in a computing environment. The integration of Carbon Footprint and Blockchain technology is discussed, presenting a framework for managing carbon emissions in data storage. Carbon accounting methodologies, including spend-based, activity-based, and hybrid methods, are detailed along with their applications in estimating greenhouse gas emissions. Finally, the paper explores the intersection of Carbon Emission Trading and Blockchain, highlighting how Blockchain's transparency and security attributes can address challenges within carbon trading systems.

Solid-State Storage

Evolving from disk drives, Solid-State Drives (SSD) were found to consume less power and be more resilient against physical damage. Unlike the traditional hard disk drives consisting of spinning disks and mechanical parts, solid-state storage makes use of integrated circuits capable of storing data even when power is removed. SSDs primarily consist of NAND flash memory chips, cells organized into binary digits. By controlling the electrical charge in the cells, the data can be manipulated or read from memory.

By storing data with moving parts, Solid-State Drivees eliminated the need for mechanical components to read and write data, therefore also eliminating the worry of mechanical failures such as disk crashes or motor failures. Latency, the amount of time it takes for a specific operation to be completed after a request has been made, originally a weakness of disk drives due to constant timing delays, has been improved upon within the SSD model. Reports indicate that SSDs deliver data roughly 50 times faster than their predecessor.

Disk drives lack the ability to overlap computing abilities, as each request must be processed one after the other without a way to perform multiple access operations simultaneously. In traditional hard disk drives, the read and write head must physically move to the correct location to access the specified data, resulting in each access operation needing to wait for the earlier one to complete before starting. With such serialized latency, when multiple requests are issued, response times increase which negatively affects the overall performance of the disk. Solid-State Drives, on the other hand, can support multiple devices, allowing for simultaneous operations.

However, while solid-state drives come in various capacities, they are unable to meet large storage needs. Solid-state drives have a limited number of cycles for each memory cell. Data in solid-state drives are stored locally, and internet connection is a requirement to access the stored material.

Additionally, solid-state drives do not inherently offer data backup and redundancy. In the case of a drive failure or corruption of data, there is no inherent protection against data loss, and data recovery can prove to be challenging.

Cloud Computing

Cloud computing was first proposed in August of 2006 by the CEO of Google. Cloud computing is defined as a model that allows for easy and instant access to computing resources over a network, including data storage. Cloud computing consists of on-demand self-service that allows for instant user access to computing resources, rather than being delayed with a need of interaction with the service providers. [3] With cloud computing, data and applications can be accessed from any location with an internet connection, enabling easy collaboration and access to data on various devices with reliance on local or physical storage.

Without physicality involved in Cloud data storage, data can be replicated across multiple data centers as well as protected using automatic backup methods, a previously mentioned weakness of solid-state data storage. Cloud service providers become the ones to handle the maintenance, updates, and other administrative tasks of the data, relieving users of the responsibility of managing and maintaining physical hardware. [3]

Users can change their level of use of the computing resources as needed, rather than being limited to what a physical storage method would require. Unlike traditional methods, cloud computing allows for on-demand scalability, enabling users to quickly allocate or de-allocate resources based on their personalized needs. With cloud computing, users only pay for services and resources used, reducing total costs and eliminating the need for upfront hardware investments. [4]

However, as time has progressed, cloud computing is no longer able to fulfill data storage demands. With the increasing popularity and development of the Internet, the total number of devices connected has increased as well, generating copious amounts of data that cloud computing is unable to keep up with. As data transmission volume increases, it will inevitably result in reduced performance, resulting in larger loads of network bandwidth and delay in the final transmission. With certain situations that rely on fast and accurate real-time feedback, cloud computing will be unable to meet the requirements.

Security has also been a challenge in adopting Cloud computing, as this data storage method involves sharing of computing resources. While this may be an advantage in that multiple users can use the same physical infrastructure while maintaining privacy and security for their data, cloud consumers do not have control over the infrastructure. Despite the precautions that Cloud computing provides, threats such as data loss, phishing, and botnet continue to pose significant security risks to the data and software integrity of organizations.

Edge Computing

Edge computing performs calculations at the edge of the network, bringing the user closer to the source of the data, allowing for faster and more efficient data processing.

Compared to cloud computing, edge computing focuses on smaller-scale analyses, and has greater suitability towards local services, responsible for the data located closer to the information source, allowing data to be stored and processed locally without the need for the data to be updated to the cloud. Because the data is processed locally on edge devices or edge servers, closer to the source of data generation, it significantly reduces the time it takes for data to travel to a central cloud server and back, leading to lower latency and faster response times.

Edge computing also provides improved security for important data, as data is processed and stored locally rather than sent to a central cloud server, as in cloud computing. By storing the data locally, the risk of data breaches during data transmission is reduced, and it gives users greater control over their data privacy and security.

A limitation of edge computing appears in the edge nodes, as they are unable to accommodate heavyweight software because of hardware limitations.[6] Edge devices typically have less processing power compared to centralized servers or data centers, restricting the complexity and scale of computations that edge computing can perform. Edge analytics require more lightweight algorithms capable of handling machine learning and data processing tasks with greater efficiency.

In terms of data abstraction, edge computing may provide unreliable reports due to low precision sensors, hazardous environments, and unstable wireless connections. If too much raw data is lost due to filters, certain applications and services may be unable to retain sufficient knowledge from the original data, however keeping large quantities of the raw data would be challenging for the data storage. [7]

With edge computing, it can be difficult to precisely identify the reason for a service failure, as all applications are connected. If several applications share the same data resource, if one fails, the whole system often crashes as well. It can then be difficult to determine the exact source of the failure or alert the user if a certain part of the system becomes at high risk of failure. [7]

Blockchain

Traditional data storage methods, including those mentioned previously, rely on centralized servers controlled by a single entity. Blockchain, however, operates on a decentralized network of computers (nodes). Each node contains a copy of the entire blockchain, ensuring that data is not stored in a single location vulnerable to attacks.

With the rise of blockchain, contracts labeled "Smart Contracts" have been created. These are self-executing contracts with the terms of the agreement directly written in the code. Smart contracts can automate processes related to data storage, retrieval, and sharing. Smart contracts have become increasingly popular due to their enhanced usability through integration of blockchain technology. [17]

Blockchain technology enables the creation of contracts without the need for intermediaries such as notaries that were previously essential for establishing trust. [17] Every participant in a blockchain network has access to the entire ledger, and all transactions are recorded in a transparent manner. This transparency is particularly useful for auditing purposes, as any changes made to the data can be easily traced and verified. Blockchain technology also uses cryptographic hashes to secure data. With the extra measure, data stored on the blockchain can be verified for integrity, ensuring that it has not been tampered with. Additionally, consensus mechanisms like proof-of-work or proof-of-stake contribute to the security of the network.

Carbon Footprint of Data Storage

Carbon Footprint (CF)is the total amount of greenhouse gas emissions that are accumulated during a product's life cycle. Directly defined, it is the carbon dioxide equivalent mass based on 100 years GWP (the amount of greenhouse gases contributing to global warming and climate change within a 100-year time frame). [9] CF has become widely used as an indicator of greenhouse gas emissions, used to evaluate and mitigate increasing emissions, as well as used to further awareness of climate change challenges and environmental consequences of industrial issues, sustainability, research, and innovation. [9]

Using CF, companies can identify the sources of greenhouse gases that affect the climate most substantially and use that analysis to increase production efficiency while also reducing emissions. [9] CF standards applied by companies may differentiate themselves from competitors by demonstrating environmental responsibility, while also meeting the growing consumer demand for more environmentally friendly products. [9]

Carbon Calculation

Various approaches are employed to calculate carbon footprints. If the calculations are executed on local computer devices such as a laptop or a desktop computer, the tools must be situated on the user's end, meaning that the user's device performs the data collection. Users can use resources such as an online calculator or integrate software, package, or library in their code to monitor computational usage and use to generate carbon footprint estimates. [12]

The online calculator method is most flexible as it functions across any hardware setup, programming language, or field of research. It also has the capacity to anticipate carbon footprints for pre-planning and retrospectively analyze past computations. However, it can be inefficient dealing with large numbers of tasks and requires users to gather specific information regarding the executions including runtime, memory, core usage, and more. [12]

An embedded package in the code presents potential for greater accuracy, as it monitors metrics in real-time and unlike the online calculator, is capable of efficiently handling numerous tasks. However, this advantage depends upon the availability of a tool tailored to the specific combination of programming language, hardware, and research field. [12]

By integrating the concept of Carbon Footprint and Blockchain data storage, a framework can be established to manage carbon emissions. This framework illustrated by Liu, Chag, Huang, and Lu [11] consists of three layers: calculation, blockchain, and integration.

The calculation layer outlines how an organization gathers data to calculate its carbon footprint, involving two methods: the conventional carbon footprint inventory and the automatic collection of electronic data, like IoT technology, which encompasses raw materials, energy usage, and resource consumption (water, electricity, fuel, or gas). [11] The blockchain layer serves three main purposes: certification, consensus, and tracking. The data stored in the blockchain layer, if not encrypted, can be made public to establish consensus between the public, the industry, and local or central governments. Blockchain allows for the public to have easy access to product information and take relevant actions towards carbon emission reduction. The final layer, integration, serves to retrieve CF details from each organization either through blockchain or traditional reporting. [11]

Carbon Accounting

Carbon accounting is built upon two distinct sets of information: **business data** and **emission factors**.

Business data consists of the actions and undertakings of an enterprise. This information comprises:

1. **Spend data**: reflecting the monetary transactions directed to company X for specific goods or services. [19]
2. **Activity data:** quantifying the volume of fuel in liters or materials in kilograms procured during transactions.[19]

Emissions factors specify the quantity of greenhouse gas emissions linked to a specific unit of business data. Once all required data is accumulated, it can then be translated into emissions estimates. The methodology for achieving this outcome varies based on the approach applied.

Different Approaches:

Carbon accounting uses two methodologies to calculate an organization's greenhouse gas (GHG) emissions: **spend-based** and **activity-based**. The hybrid approach combines these two methodologies.

Spend-based method: value of a purchased product or service is multiplied by an emission factor, usually representing emissions per unit of currency, to yield an emissions estimate. [19] Emission factors in the spend-based approach are often drawn from environmentally extended input-output (EEIO) models, depicting resource flows among economic sectors. This allows the determination of average emissions associated with each monetary unit disbursed to companies within specific industries and regions. [19]

Activity-based method involves quantifying the quantity of a specific product or material purchased by a company. [19]

Hybrid method: combines both spend-based and activity-based data.

While activity-based data generally yields more accurate emissions estimates than spend-based data, it may be less readily available and time-intensive to gather. Thus, the hybrid model's strategy can maximize the utilization of available activity-based data and subsequently applying spend-based techniques to estimate remaining emissions. [19]

Carbon Emission Trading and Blockchain Implementation

Carbon emission trading emerging from the 1997 Kyoto Protocol, is designed to encourage reduction of greenhouse gases including carbon dioxide and control emission levels. Nations are limited to producing certain amounts of carbon dioxide, and they can trade excess carbon dioxide with other nations.

Where the system was designed to foster environmental protection, lack of transparency within the trading system has led to corruption and manipulation. An integral aspect of carbon emissions trading is having a carbon allowance, a predetermined amount of greenhouse gas emissions permissible to avert global warming and climate change. The process of calculating a country's carbon allowance is highly complex and considered projection rather than set quantity, making it greatly vulnerable to artificial inflation and manipulation. [15]

Companies often control a country's carbon allowance; however, the company's actions are not monitored, and their various systems and decisions are not transparent to accessible to others. [15]

Blockchain may emerge as a solution to the challenges facing carbon emission trading. Blockchain technology is a trusted platform for permanently and securely storing and validating data records. [16] Blockchain would provide a completely transparent operational framework that facilitates interactions between parties without necessitating a third party or intermediary. At any point, nodes can enter or exit the carbon trading network without causing any disruption to the operation or continuity of other nodes or ongoing processes within the blockchain. All transactions will be fully visible to any node within the blockchain, and the blockchain retains comprehensive historical records of all processed data. [15]

Despite its full transparency, the acquired information is maintained in an anonymous state through cryptography, and once a block is appended to the chain, it remains unchangeable and immune to erasure. [15]

References

[1] K. Cao, Y. Liu, G. Meng and Q. Sun, "An Overview on Edge Computing Research," in IEEE Access, vol. 8, pp. 85714-85728, 2020, doi: 10.1109/ACCESS.2020.2991734.
[2] E. Seppanen, M. T. O'Keefe and D. J. Lilja, "High performance solid state storage under Linux," 2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST), Incline Village, NV, USA, 2010, pp. 1-12, doi: 10.1109/MSST.2010.5496976.
[3] T. Dillon, C. Wu and E. Chang, "Cloud Computing: Issues and Challenges," 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Perth, WA, Australia, 2010, pp. 27-33, doi: 10.1109/AINA.2010.187.
[4] Kim, W. (2009). Cloud computing: today and tomorrow. The Journal of Object Technology, 8(1), 65. https://doi.org/10.5381/jot.2009.8.1.c4
[5] Hayes, B. (2008). Cloud computing. Communications of the ACM, 51(7), 9–11. https://doi.org/10.1145/1364782.1364786
[6] B. Varghese, N. Wang, S. Barbhuiya, P. Kilpatrick and D. S. Nikolopoulos, "Challenges and Opportunities in Edge Computing," 2016 IEEE International Conference on Smart Cloud (SmartCloud), New York, NY, USA, 2016, pp. 20-26, doi: 10.1109/SmartCloud.2016.18.

[7] W. Shi, J. Cao, Q. Zhang, Y. Li and L. Xu, "Edge Computing: Vision and Challenges," in IEEE Internet of Things Journal, vol. 3, no. 5, pp. 637-646, Oct. 2016, doi: 10.1109/JIOT.2016.2579198.
[8] M. Satyanarayanan, "The Emergence of Edge Computing," in Computer, vol. 50, no. 1, pp. 30-39, Jan. 2017, doi: 10.1109/MC.2017.9.
[9] Scrucca, F., Barberio, G., Fantin, V., Porta, P. L., & Barbanera, M. (2020). Carbon Footprint: Concept, Methodology and Calculation. In Environmental footprints and eco-design of products and processes (pp. 1–31). https://doi.org/10.1007/978-981-15-9577-6_1
[10] C. Honée, D. Hedin, J. St-Laurent and M. Fröling, "Environmental performance of data centres - A case study of the Swedish National Insurance Administration," 2012 Electronics Goes Green 2012+, Berlin, Germany, 2012, pp. 1-6.
[11] Liu, K., Chang, S., Huang, W., & Lu, I. (2019). The framework of the integration of carbon footprint and Blockchain: Using blockchain as a carbon emission management tool. In Springer eBooks (pp. 15–22). https://doi.org/10.1007/978-981-13-1181-9_2
[12] Lannelongue, L., Inouye, M. Carbon footprint estimation for computational research. *Nat Rev Methods Primers* **3**, 9 (2023). https://doi.org/10.1038/s43586-023-00202-5
[13] D. Shakhbulatov, A. Arora, Z. Dong and R. Rojas-Cessa, "Blockchain Implementation for Analysis of Carbon Footprint across Food Supply Chain," 2019 IEEE International Conference on Blockchain (Blockchain), Atlanta, GA, USA, 2019, pp. 546-551, doi: 10.1109/Blockchain.2019.00079.
[14] Li, J., Wei, Y., Liu, L., Li, X., & Yan, R. (2022). The carbon footprint and cost of coal-based hydrogen production with and without carbon capture and storage technology in China. *Journal of Cleaner Production*, 362, 132514. https://doi.org/10.1016/j.jclepro.2022.132514
[15] Sadawi, A. A., Madani, B., Saboor, S., Ndiaye, M., & Abu-Lebdeh, G. (2021). A comprehensive hierarchical blockchain system for carbon emission trading utilizing blockchain of things and smart contract. *Technological Forecasting and Social Change*, 173, 121124. https://doi.org/10.1016/j.techfore.2021.121124
[16] Murray, M. (2019). Tutorial: A descriptive introduction to the blockchain. *Communications of the Association for Information Systems*, 464–487. https://doi.org/10.17705/1cais.04525
[17] Nofer, M., Gomber, P., Hinz, O. et al. Blockchain. Bus Inf Syst Eng 59, 183–187 (2017). https://doi.org/10.1007/s12599-017-0467-3
[18] P. Tasatanattakool and C. Techapanupreeda, "Blockchain: Challenges and applications," 2018 International Conference on Information Networking (ICOIN), Chiang Mai, Thailand, 2018, pp. 473-475, doi: 10.1109/ICOIN.2018.8343163.
[19] Farbstein, E., Vallinder, A., & Buchmann, L. (2023). Carbon accounting, explained. Normative. https://normative.io/insight/carbon-accounting-explained