Energy Efficiency and Sustainability in Data Centers

Simon Harrison-Michaels
harr3247@morris.umn.edu
Division of Computer Science
University of Minnesota, Morris
Morris, Minnesota, USA

Abstract

This paper summarizes the different strategies used to maximize energy efficiency in data centers. It explores techniques involving both hardware and software, and explains the future importance/consequences of energy efficiency and sustainability in data centers. It analyzes two tools designed to improve energy efficiency of data centers - FootPrinter and PACT - and examines why they are effective.

Keywords: data centers, power usage effectiveness, demand based switching, multiplicative saving

1 Introduction

What are data centers? Data centers are a type of building that contains servers which can store and manipulate large amounts of data. These servers require large amounts of energy to store and manipulate data. Some parts of servers (like processing units) also need to be kept at a consistent temperature to function correctly. This means that cooling systems are a necessity for data centers. There are two main points of energy consumption in a data center - IT equipment and cooling systems. We will explore both of these later on.

There has been an increase in the amount of attention directed towards data center energy efficiency. The global energy consumption of data centers is expected to increase substantially in the coming decades, in large part due to the growth of artificial intelligence. To put this into perspective, Figure 1 is an often-cited graph from a 2020 study undertaken by the company Huawei Technologies. It estimates the best-case and the expected future electricity usage of data centers from 2020-2030. The best case scenario requires an 86 percent increase in electricity usage, and the *expected* scenario requires a 225 percent increase in electricity usage.

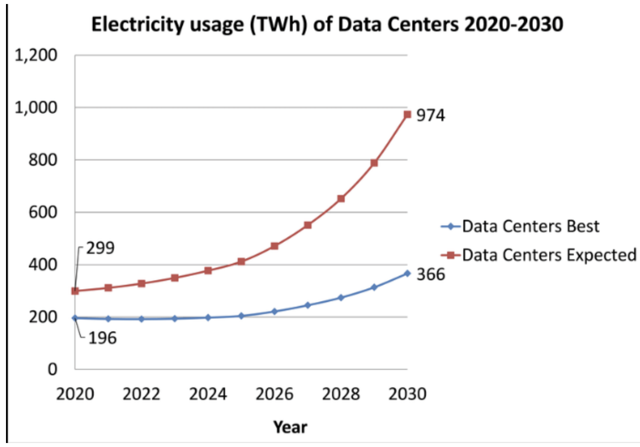

Figure 1. Projected Data Center Electricity Usage (taken from [5])

1.1 Functions of a Data Center

The basic purpose of a data center is to store and manage data. Data centers need to be reliable and consistent at all times. Data centers are typically open 24 hours a day, 7 days a week. They respond to requests for data, maximizing energy efficiency and minimizing latency issues. They often deal with sensitive and critical data which needs to be managed securely as well as distributed quickly, as the potential consequences for security breaches or latency issues can be disastrous. An example of this is the technology of medical devices. For example, types of robotic surgery devices communicate with data centers in real-time while doing surgery. Errors and latency issues in the data center will increase the probability of a mistake from the robotic surgery device [1]. Problems in data centers can cause real safety and/or security concerns in the real world.

1.2 Studies that will be focused on

New methods designed to improve energy efficiency in data centers are in high demand. In this area, two products that have been introduced recently are FootPrinter which is a software system that measures Operational Footprint designed at Vrije University in Amsterdam in 2024, [4] and PACT (Per Application Class Turbo Controller) which is a Demand Based Switching (DBS) software system designed in a joint effort between researchers at Stanford University and Google in 2020[2]. First this paper will go over the prerequisite ideas to understand FootPrinter, and then will analyze FootPrinter itself. Then, this paper will go over the prerequisite ideas to understand PACT, and then will analyze PACT.

2 Energy Consumption of a Data Center

Data centers currently produce roughly 2-3 percent of the world's total global carbon emissions [2]. However, this number is estimated to rise significantly in the coming decades. Most estimates conclude that data centers will produce close to 10 percent of the world's total global carbon emissions by

2040 [5]. Therefore, there has been much recent attention directed to reducing carbon emissions of data centers.

The vast majority of data center electrical energy expenditure can be divided into two categories - IT equipment (such as servers) and cooling systems. In the average data center, cooling systems will consume roughly 50 percent of the total electrical power. The other 50 percent will be split between the other tasks - powering the servers, network hardware, security systems, lighting, etc. Furthermore, many data centers require types of energy besides electricity. For example, liquid cooling systems require vast quantities of water.

2.1 Cooling Systems

Cooling systems are a necessary part of every data center. The main purpose of a cooling system is to conduct heat away from the heat-generating components - like processing units. Complete studies can be written about cooling systems - this paper will not go in depth. However, it is important to understand the basic purpose and components. Data center cooling systems utilize either liquid cooling or air cooling to reduce temperature - often a combination of the two.

Air cooling uses fans and HVAC systems to push cold air towards the heat-generating components. Then, the cold air absorbs heat and is transferred away from the heat-generating components. Although less efficient then liquid cooling, they are easier to install and maintain.

There are two main types of liquid cooling - direct-to-chip cooling and immersion cooling. Direct-to-chip cooling involves circulating coolant (usually water) through very small pipes close to the heat-generating components. Immersion cooling physically immerses processing units in water. Although immersion cooling is more prone to leaks and repair costs, it is more efficient than direct-to-chip cooling. Currently, immersion cooling is less commonly used than direct-to-chip cooling, due to the higher cost of implementation and the higher likelihood of errors. However, immersion cooling is thought of by some as the cooling method of the future.

3 Energy Efficiency Measurement Tools

3.1 Power Usage Effectiveness (PUE)

Power Usage Effectiveness (PUE) is a metric for measuring the efficiency of data center's power usage as a whole [4]. Power Usage Effectiveness is the ratio of the total amount of power used by a data center to the power used by the IT equipment. A Power Usage Effectiveness value of 1 would be ideal - signifying that all of the energy arriving to a data center goes directly into computing power. However, this is not possible in real-world data centers because things like cooling systems, security systems, and lighting require energy as well. The formula to calculate Power Usage Effectiveness is the total facility energy usage divided by the IT equipment energy usage.

$$PUE = \frac{TotalFacilityEnergyUsage}{ITEquipmentEnergyUsage} \tag{1}$$

The average PUE value has been declining from 2007-2022. That means that data centers have been getting better at using their available energy more efficiently. This is largely due to improvements in cooling systems, as well an increased focus in data center energy efficiency during that time. Figure 2 shows a graph from a study published in the Energy Informatics scientific journal of August 2023.

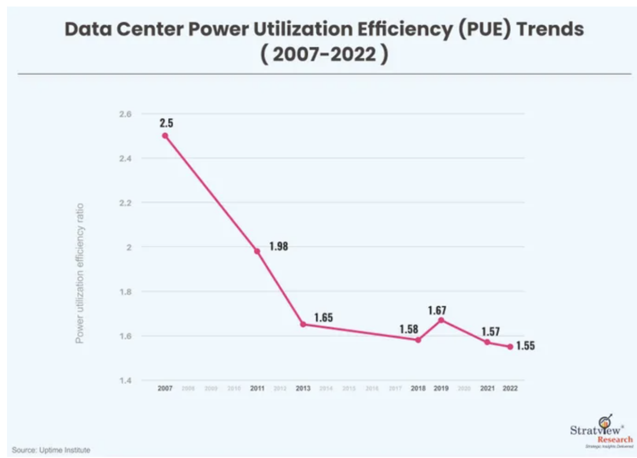

Figure 2. PUE trends 2007-2022(taken from 3)

As mentioned before, a Power Usage Effectiveness value of 1 isn't realistic in the real world. Therefore, A Power Usage Effectiveness value of 1.2 or less is generally considered to be good. [6]

Although the average Power Usage Effectiveness has been decreasing over the last two decades, it it has not quite reached 1.2 yet. However, significant progress has been made in this regard. Furthermore, the Power Usage Effectiveness measurement system is not without its flaws. Because the PUE value is simply the ratio of the total facility energy usage to the energy used by the IT equipment (servers), it doesn't account for operational energy usage outside of the IT equipment and the energy expended towards cooling systems. Furthermore, it ignores the supply chain of the energy itself - whether the energy is supplied from renewable or nonrenewable sources.

3.2 Carbon Intensity

To take into account the origin of the energy itself (coal, solar, etc.) a metric called carbon intensity is used. Carbon intensity is defined as the amount of carbon emitted per unit of energy used. Unlike Power Usage Effectiveness, Carbon Intensity takes into account the source of the energy itself. The source of the energy (coal, solar, natural gas, etc...) has a great impact on the amount of carbon emitted [4]. This can be tricky because data centers obtain power from a number of different sources, and many data centers get energy from the grid. Energy from the grid almost always comes from multiple energy sources. This complicates things as energy generated from renewable sources such as wind or solar can emit up to 20 times less carbon dioxide compared to energy generated from coal. Therefore, the carbon intensity of the grid is calculated by aggregating the different energy sources.

$$CarbonIntensity = \frac{TotalCO2Emission(g)}{TotalEnergyConsumed(kWh)} \tag{2}$$

3.3 Limitations with Existing Methods

A major limitation of existing methods is that they cannot accurately predict energy efficiency into the future. Power Usage Effectiveness and Carbon Intensity are helpful for evaluating past and current carbon emissions, but are unable to predict future carbon emissions - whether for a newly constructed data center or an existing data center implementing new technology. To predict future carbon emissions, data center designers and operators often need to choose between different techniques and methods without fully knowing their effects. Because of this, reducing carbon emissions can come down to trial and error, which is inefficient from every perspective.

4 FootPrinter

With the goal of developing an accurate model to simulate carbon footprint, researchers at the University of Vrije in Amsterdam designed the FootPrinter software system. It incorporates aspects of existing tools. For example, similar to the Power Usage Effectiveness metric it ultimately determines the total footprint of the data center, and similar to the Carbon Intensity metric it takes into account if energy is renewable or nonrenewable. Indeed, the formula to calculate Carbon Intensity is built in to the FootPrinter algorithm itself.

However, FootPrinter goes beyond reporting past performance because it is designed to simulate energy demand into the future. This is something that is quite difficult to do given all the different variables and moving parts associated with energy generation, the supply chain, and the data center itself. To simulate energy demand into the future, FootPrinter takes three pieces of information as input data - workload traces, hardware and environment specifications, and operational techniques [4].

4.1 FootPrinter Input Data:

4.1.1 Workload Traces. The big idea is behind workload traces is that as specific jobs are executed in a data center, a small amount of information is stored about that specific job. This data is called a workload trace. It contains things like timestamps for the job and resource usage (CPU, memory). In this way, a workload trace is similar to how a normal operating system tracks processes - it captures the details about a specific job's execution and necessary resources. This data is gathered when a task is finished - not when a task is started. Therefore, the main usefulness of workload traces is that FootPrinter can use them to extrapolate for future predictions.

4.1.2 Hardware and Environment Specifications. Hardware and environmental specifications represent data about the physical infrastructure of a data center itself. This includes information like the location of the data center, the types of hardware used (servers, storage devices, networking equipment), and environmental factors (cooling systems, power supply, and temperature control). These specifications help FootPrinter understand the physical setup and the conditions under which the data center operates

4.1.3 Operational Techniques. Operational techniques involve scheduling and resource allocation policies. In the study, the researchers describe operational techniques as "human-based constraints". This means that any regulation on the data center would fit into this category. An example of this would be if the cooling system needs to be maintained every so often, or if there is a set maximum value of total energy consumption per day. Any human-based restriction made on the hardware/software would fall into this category.

4.2 Results of FootPrinter

FootPrinter was tested over a 1 week period. FootPrinter predicts the next 30 seconds of energy consumption at a time. FootPrinter was able to simulate the power draw of a data center to an accuracy of 1.0 kW. Figure 3 is a graph published in the study. FootPrinter was able to predict power draw with a high degree of precision, never varying more than 1kW from the real data [4]. Comparing FootPrinter to the ground truth results in a MAPE total error of 3.15 percent, underestimation error of 3.19 percent, and overestimation error of 2.93 percent [4].

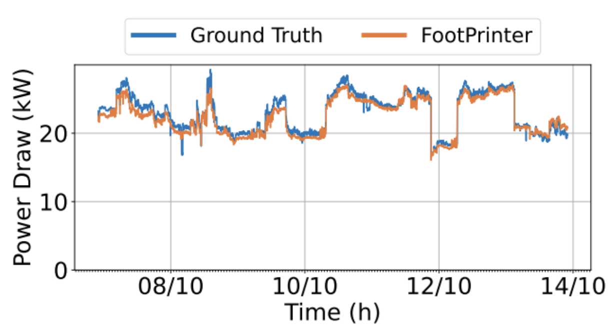

Figure 3. PUE trends 2007-2022(taken from [3])

5 Saving Power at the CPU Level

5.1 Multiplicative Saving

One of the easiest ways to save energy is to reduce the need at the CPU level. A big reason for this is the idea of multiplicative saving. Multiplicative saving means that saving energy at the CPU level can have a cascading effect towards the total energy efficiency. A simple example of this is that when less power is used at the CPU level, less power is required to cool that same CPU. This will in turn save power at the cooling/HVAC level, resulting in decreased consumption in the data center as a whole. It is possible that saving 1 watt of energy at the CPU level can save close to *3 watts* of energy in total facility energy consumption [3]. This is why it is important to save power whenever possible at the CPU level.

5.2 Idling CPUs

We've established how important it is to save power at the CPU level whenever possible. But how you do go about doing that? One of the simplest ways is to minimize the CPU's idle time. The precise definition of CPU idle time is the amount of time a processor spends not doing any particular tasks. If an idling processor is connected to power and ready to work then it is wasting usable power, which will translate to decreased energy efficiency. Similarly, a CPU that is powered down takes a small amount of time to be ready to do work. So - you might ask - what is the point of having idling CPUs at all? Could a data center function correctly with no idling CPUs? The answer is almost always no. This is somewhat of a tricky problem, because CPU idle time is sometimes necessary in many data centers.

For example, imagine a data center has optimized its CPU idle time so that every CPU is being used at all times. What would happen if suddenly, the amount of client-server connections required in the data center at that moment doubled? There would be no more free CPUs left to handle the requests. Therefore, in most cases there needs to be some level of CPU idling. This is especially true if your data center houses data where latency issues are unacceptable. An earlier example of a type of client requests where latency issues are unacceptable was robotic surgery. Even industries like e-commerce can be greatly effected by latency issues in data centers. Internal research from Amazon concluded that every 100ms of latency cost it 1 percent in sales [5].

5.3 What is Demand Based Switching?

Demand Based Switching (DBS) systems are a method of circumventing power drainage from idling CPUs. The big idea of demand based switching is to supply a CPU with only the power it needs and not more. DBS uses a processor to dynamically adjust multiple different CPUs' clock speed and voltage simultaneously to match the necessary workload at the time. If the data center's demand is high, DBS software will increase the clock speed and voltage of the CPUs to meet demand, and vice versa if the demand is low.

However, this becomes more complicated when you consider that demand based switching software itself requires processing capacity to work. Similar to how the weight of a rocket's fuel increases the amount of total fuel needed, the power consumption of demand based switching adds to the need for demand based switching. This means that complicated statistical analysis is necessary to determine the effectiveness of and proper development of demand based switching software systems.

5.4 How do Demand Based Switching systems work?

Demand Based Switching systems work by embedding sensors into equipment in order to collect current data about energy usage, processing load, and temperature. These sensors then send data back to a centralized software program, which analyzes the data received. Then, the centralized software system decides which system operations to alter. This could involve manipulating the number of available processing units, changing parameters of the cooling system, and adjusting the voltage and clock speed of the processing units.

6 PACT (Per Application Class Turbo Controller)

PACT (Per Application Class Turbo Controller) is a demand based switching system designed jointly between researchers at Stanford University and Google. The goal of PACT is to use demand based switching to decrease energy demand without creating latency issues in a shared cloud environment. It is true that this stated goal is the goal of *any* demand based switching system. However, the unique part of PACT compared to other demand based switching systems lies in its use of per-thread dynamic frequency scaling and its separation of latency sensitive and non-latency sensitive tasks. This means that demand based switching can be implemented on a thread to thread basis. Similarly, the separation of tasks into latency sensitive and non-latency sensitive (or best effort) tasks means that demand based switching will decrease total energy consumption while increasing latency issues as little as possible. PACT is made up of two parts - Turbo Control and CPUJailing.

6.1 Turbo Control

The purpose of Turbo Control is to dynamically adjust the parameters (like the clock speed and voltage) of each CPU core depending on the necessary workload at a particular moment. The ultimate goal of turbo control is to switch the frequency of each CPU core to match the requirements of the application currently running on that core, in order to decrease the number of idling CPUs [2]. One difference between turbo control and an average demand based switching system is that while an average demand based switching system will set CPU parameters based on data over a period of time, a turbo control demand based switching is good for providing short bursts of energy when required. This idea of "short bursts of energy" when required can be seen in the name - turbo control.

The advantage of turbo control is that it allows for a higher quantity of idling CPUs. This is because turbo control increases the capacity of the data center to handle periods of increased demand. An ability to allocate a higher percentage of processing units to a state of idling means less power consumed at the CPU level. This will cascade along with multiplicative saving to save more power at all levels. Indeed, the "ability to allocate a higher percentage of processing units to a state of idling" is a main component of the second big part of PACT - CPU Jailing.

6.2 CPUJailing

Similar to Turbo Control, the big idea of CPUJailing is to minimize the number of idling CPU cores in order to save power. One of the ways that it does this is by dividing tasks into two categories - latency sensitive (LS) tasks and best effort (BE) tasks. Latency sensitive tasks are tasks where latency issues are unacceptable. Our earlier example of robotic surgery would be considered a latency sensitive task. Best effort tasks are all tasks that aren't latency sensitive tasks. An example of a best effort task could be sending an email. This is because the email getting delivered a few milliseconds later will not cause safety or security issues. CPUJailing then uses this info (LS or BE) to make decisions. Simply put, CPUJailing controls the amount of idling CPU cores by prioritizing latency sensitive tasks over best effort tasks. Similar to the Turbo Control mechanism, CPUJailing dynamically switches the parameters of the CPUs in real time, running LS tasks on separate physical cores from BE tasks. If there are not enough idle cores available at a given time, CPUJailing will auto-disable itself. This means that CPUJailing will interfere with latency issues as little as possible, because if latency issues are severe in a given instance the mechanism of CPUJailing will simply shut off.

6.3 Testing of PACT

PACT was demonstrated to reduce power through two different production clusters - Cluster A and Cluster B. Each cluster contains approximately 10,000 servers running a combination of different software platforms common in data centers. Each clusters' servers responded to tasks for 10 total days, responding to a combination of latency sensitive and non-latency sensitive tasks. Clusters A and B were fed slightly different types of tasks - Cluster A responded to 45.8 percent latency sensitive tasks and Cluster B responded to 53.4 percent latency sensitive tasks. Similarly, each cluster iterated through four different settings - Baseline, Turbo Control, CPUJailing, and PACT. The "PACT" setting means that both the Turbo Control and CPUJailing mechanisms were being used by the IT equipment at the same time.

6.4 Results of PACT

Turbo Control and CPUJailing both individually reduced energy consumption, and significantly reduced energy consumption when combined together [2]. This can be seen in Figure 4 in which Turbo Control and CPUJailing both individually saved energy, but particularly so when both aspects of PACT were utilized together.

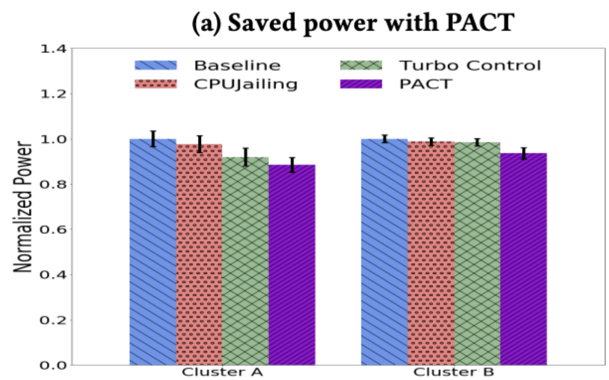

Figure 4. Normalized Power of Two Different Clusters (PACT) ([2])

7 Conclusion

In conclusion, the long-term effects of the FootPrinter and PACT papers will be determined by their ability to translate from the experiments and testing to real implementation. For example, both the FootPrinter and PACT papers were somewhat vague about the easiness of their installation and if they can be set up efficiently. However, the potential for predicting (in the case of FootPrinter) and reducing (in the case of PACT) energy consumption will bring attention to their products. The future will likely bring even more new techniques to increase energy efficiency. Things like the continued research of immersion cooling or nuclear energy powered data centers will likely be undertaken. The consequences of ignoring the energy efficiency of data centers will be experienced by the environment, and subsequently the people that inhabit it.

Acknowledgments

I want to thank my senior seminar advisor, Kristin Lamberty, for guidance and feedback during the development of this paper. I would also like to thank Elena Machkasova for leading the senior seminar course.

References

[1] W.J Gordon. Protecting procedural care—cybersecurity considerations for robotic surgery. *npj Digit. Med. 5*, 2022. URL https://doi.org/10.1038/s41746-022-00693-8.

[2] Kostis Kaffes, Dragos Sbirlea, Yiyan Lin, David Lo, and Christos Kozyrakis. Leveraging application classes to save power in highly-utilized data centers. In *Proceedings of the 11th ACM Symposium on Cloud Computing*, SoCC '20, page 134–149, New York, NY, USA, 2020. Association for Computing Machinery. ISBN 9781450381376. doi: 10.1145/3419111.3421274. URL https://doi.org/10.1145/3419111.3421274.

[3] Wania Khan, Davide De Chiara, Ah-Lian Kor, and Marta Chinnici. Exploratory data analysis for data center energy management. In *Proceedings of the Thirteenth ACM International Conference on Future Energy Systems*, e-Energy '22, page 571–580, New York, NY, USA, 2022. Association for Computing Machinery. ISBN 9781450393973. doi: 10.1145/3538637.3539654. URL https://doi.org/10.1145/3538637.3539654.

[4] Dante Niewenhuis, Sacheendra Talluri, Alexandru Iosup, and Tiziano De Matteis. Footprinter: Quantifying data center carbon footprint. In *Companion of the 15th ACM/SPEC International Conference on Performance Engineering*, ICPE '24 Companion, page 189–195, New York, NY, USA, 2024. Association for Computing Machinery. ISBN 9798400704451. doi: 10.1145/3629527.3651419. URL https://doi.org/10.1145/3629527.3651419.

[5] Wojciech Piatek, Ariel Oleksiak, and Micha vor dem Berge. Modeling impact of power- and thermal-aware fans management on data center energy consumption. In *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*, e-Energy '15, page 253–258, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450336093. doi: 10.1145/2768510.2768525. URL https://doi.org/10.1145/2768510.2768525.

[6] Sanyukta Raje, Hemant Maan, Suprotim Ganguly, Tanvin Singh, Nisha Jayaram, Girish Ghatikar, Steve Greenberg, Satish Kumar, and Dale Sartor. Data center energy efficiency standards in india. In *Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems*, e-Energy '15, page 233–240, New York, NY, USA, 2015. Association for Computing Machinery. ISBN 9781450336093. doi: 10.1145/2768510.2768524. URL https://doi.org/10.1145/2768510.2768524.