Future Data Center Energy Efficiency and Consumption

Alexander B. Penney[1]; Zhen Zhang[1]
[1]Dept. Of Electrical and Computer Engineering

Research Thesis Question

- What would large data centers look like in 2050, from an energy usage perspective?

Abstract

Presently, total energy consumption of all data centers (DCs) is estimated to be 1% of global electricity consumption. Consequently, facilities are motivated to maximize energy efficiency. Significant progress has been made to optimize the cooling, lighting, and power conditioning systems, thereby directing as much as 94% of power into the computing equipment. This poster presents a projection of data center power consumption and energy efficiency based on historical trends in server energy efficiency. An additive model with numerous IT equipment power metrics is used to estimate the aggregated energy consumption of a facility. Server-side Java Operations/Watt (ssj_ops/W) is collected over the years 2010 to 2022 and used to project the growth of server energy efficiency. We evaluate the claims of Koomey's law, which states that the energy efficiency of computation generally doubles every 2.6 years, compared to the growth of server energy efficiency over the past decade and draw conclusions on the expectation of data centers efficiency in the next 25 years.

Introduction

Data centers contain a system of information technology (IT) equipment consisting of servers, external storage devices, and network devices. Infrastructure systems such as cooling, lighting, and power conditioning enable the daily operation of the facility but are not directly related to the computing process, and their power consumption has mostly leveled off in recent years [1]. The categorization of electrical equipment is shown in Fig. 1. Prior work on projections up until 2020 have predicted lower PUE, and increased efficiency for IT infrastructure [2].

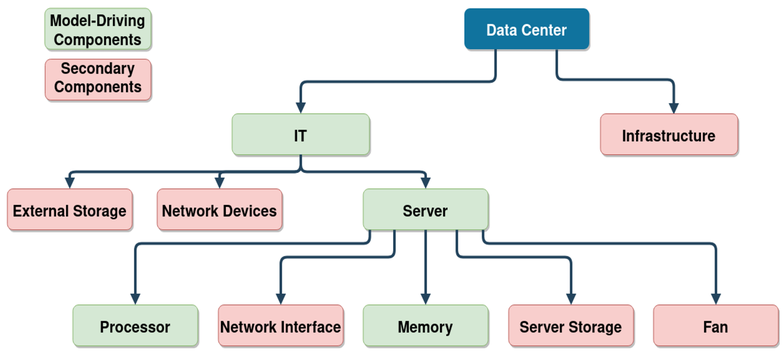

Figure 1. Categorical Chart of Data Center IT and Infrastructure Equipment

Methods and Assumptions

1. The study will only examine the energy efficiency of computing in future data centers.
2. Power usage effectiveness (PUE) is assumed to be constant, evidenced by the converging trend with diminishing returns in recent years [3]. Thus, change in energy consumption of infrastructure equipment is assumed to be proportional to that of IT equipment, and is not evaluated individually in this study.
3. The additive model sums the power draw of individual IT equipment: network devices, storage devices, and the server. One million servers is assumed for a facility. Individual growth rates of power consumption for each category of equipment are used to project current and future trends.
4. Number of server-side Java operations/Watt (ssj_ops/W) is a benchmark metric from Standard Performance Evaluation Committee (SPEC) [5] to evaluate energy efficiency of different servers with different instruction set architectures. All servers are instructed to run a computationally heavy program until a set amount of power is drawn from the power supply, thus offering a standardized comparison across different servers.

Results

By analyzing historical industry data, generating future trend profiles, and calculating projections, we have derived initial results for both baseline energy consumption and performance per watt metrics for the year 2050.

1. Individual Servers are 4X more energy efficient in computation.
2. External Storage and Network Devices use 15% larger share of total IT consumption.
3. Computational Capacity with normalized power consumption will increase 4X

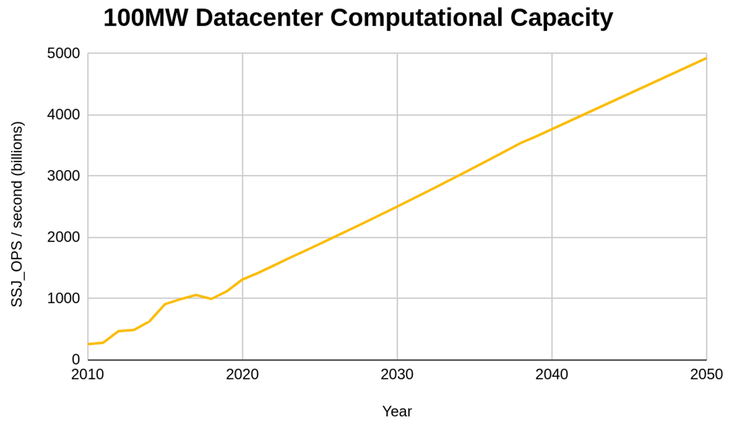

Figure 4. Line Plot of Projected Data Center Computational Capacity for a normalized power consumption of 100MW - a typical figure for a hyperscale datacenter.

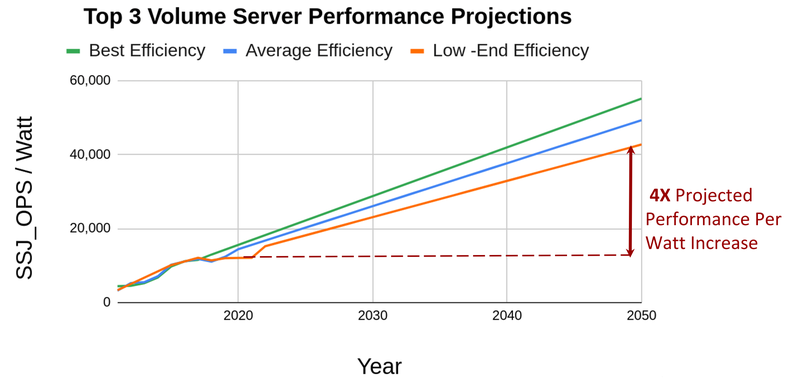

Figure 2. Projection of Energy Efficiency in ssj_ops/W for three case scenarios with varying efficiency growth rates. Conservatively, the lowest efficiency shows ~4x performance per Watt.

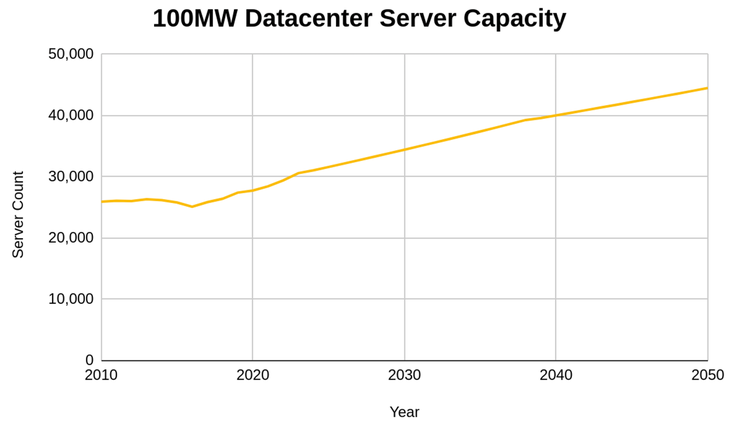

Figure 3. Line Plot of Projected Data Center Server Capacity

Conclusions and Discussion

Based on historical trends of data center IT equipment from 2010 to 2022, two projections were concluded regarding data center power consumption and server energy efficiency. From the growth rate of energy efficiency of servers, Koomey's law is expected to drastically slow down. The energy efficiency of computation will not double every 2.6 years, but rather show 4x growth after 25 years, which is a doubling every 12 years. For future work, the additive model may be expanded to include more growth rates for improved granularity, as well as taking into account increased server virtualization and active/idle consumption dynamics.

Contact

Alby Penney
Dept. of Electrical and Computer Engineering
Email: apenney@andrew.cmu.edu

Acknowledgements

Special thanks to Dr. Amritanshu Pandey for establishing this research project and providing invaluable guidance.

References

1. M. Dayarathna, Y. Wen and R. Fan, "Data Center Energy Consumption Modeling: A Survey," in IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 732-794, Firstquarter 2016, doi: 10.1109/COMST.2015.2481183.
2. Shehabi, A., Smith, S.J., Horner, N., Azevedo, I., Brown, R., Koomey, J., Masanet, E., Sartor, D., Herrlin, M., Lintner, W. 2016. United States Data Center Energy Usage Report. Lawrence Berkeley National Laboratory, Berkeley, California. LBNL-1005775.
3. Efficiency – Data Centers – Google. Google Data Centers, 2022, www.google.com/about/datacenters/efficiency/.
4. Masanet, E., Shehabi, A., Lei, N., Smith, S., and J.G. Koomey (2020). "Recalibrating global data center energy use estimates." Science, Vol 367, Iss 6481.
5. Power and Performance Benchmark Methodology v. 2.2, Standard Performance Evaluation Corporation, Rev. 1158, 2014.