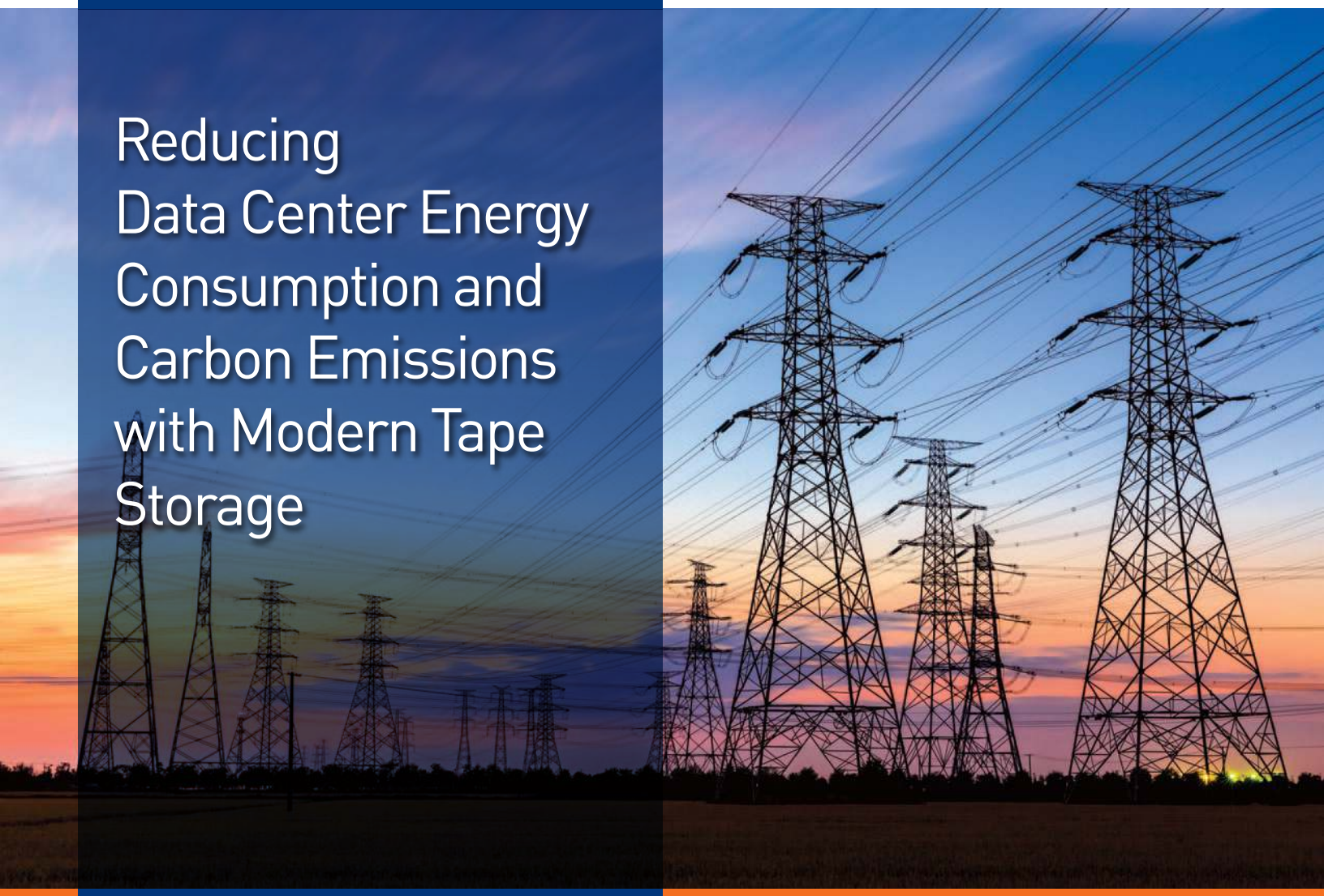
Brad Johns Consulting LLC

Reducing Data Center Energy Consumption and Carbon Emissions with Modern Tape Storage



Brad Johns
brad.johns@bradjohnsconsulting.com
November 2020

Executive Summary

Reducing carbon emissions is a significant global challenge. Many companies have decided they must incorporate carbon reductions into their strategies and have announced green initiatives. Researchers estimate that data centers consume 1.8% of all electricity in the United States. Studies also estimate storage systems incorporating hard disk drives consume approximately 19% of the total power within the data centers. Industry analysts estimate that 60% of the data stored on disk storage is infrequently accessed. Using the LTO Total Cost of Ownership (TCO) tool, we estimate that by moving 10 PB of "cold data" that is growing 35% annually from disk to tape storage, an 87% reduction in carbon emission and an 86% reduction in TCO can be achieved over ten years. IT organizations have a significant opportunity to achieve meaningful carbon emissions reductions while lowering operational and capital expenses.

Introduction

Organizations across all industries are concerned about global warming and are actively looking for ways to reduce carbon emissions. Data centers are a large user of electrical power. A recent article in Science magazine[1] estimated that data centers use as much as one percent of all the world's electric power. Within the data centers, data storage is a significant portion of total energy usage. Disk systems are the primary driver of storage energy consumption. However, industry studies indicate that much of the data residing on disk is infrequently accessed. By identifying "cold data" and moving it to modern tape storage, organizations can dramatically reduce energy consumption and associated carbon emissions while also lowering data center capital and operational expenses.

Focus on Carbon Emissions

Global warming is a tremendous concern today, and enterprises worldwide and in widely different industries have announced significant initiatives to reduce carbon emissions. Companies are aggressively seeking ways to reduce carbon emissions generated by their products, production, and supply chains. *Forbes* magazine[2] recently highlighted 101 companies committed to reducing their carbon emissions. For example, Microsoft announced this summer the Transform to Net Zero initiative with eight

[1] Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. 2020. "Recalibrating global data center energy-use estimates." Science, February 2020, p984

[2] Blake Morgan "101 Companies Committed To Reducing Their Carbon Footprint." *Forbes* magazine, August 26, 2019

other founding companies. Amazon created a Climate Pledge Fund that started with $2 billion in funding to support the development of sustainable technologies and services to enable them and other companies to be net-zero carbon by 2040. Verizon is on track to be carbon neutral by 2035 and issued a $1 billion Green Bond to invest in innovative solutions to accelerate its efforts. Walmart launched Project Gigaton to reduce one gigaton of greenhouse gas emissions from the company's supply chain by 2030. Delta Airlines committed $1 billion to mitigate all emissions from its fleet.

Data Center Energy Use

Concern over the rapid growth of data center energy usage in the early 2000s resulted in several rigorous government studies and white papers. A 2016 study[3] (A. Shehabi et al. 2016) noted data center power consumption had increased as much as 90% in the 2000-2005 period, slowing to 24% in the 2005-2010 period. Further, they estimated that data centers consumed 1.8% of all power in the United States. Their 2018 paper (A. Shehabi et al. 2018) forecasted that data center power requirements would grow 5% during the 2010-2020 decade despite the rapid expansion of cloud computing, analytics, IoT, and video during this time. They found several factors contributed to improving data center energy efficiency, including more cores/server, server virtualization, higher capacity HDDs, and advances in Power Usage Effectiveness or PUE.

[3] A. Shehabi et al.,"United States data center usage report" (Lawrence Berkeley National Laboratory, LBNL-1005775, 2016)

PUE is the ratio of total power consumed by the data center divided by the servers, storage, and internet equipment power. Cooling is a significant component of total power consumption as data center equipment generates considerable heat that must be removed. A typical PUE is two; for every watt of energy expended on equipment, another watt is used for cooling and other infrastructure components. Advances in cooling and infrastructure power management have allowed large data centers, including the hyperscale center, to reduce this ratio. They found as workloads moved from smaller environments to hyperscale data centers, which incorporate leading-edge power management techniques, overall data center power consumption was lowered. However, with data center workloads projected to grow significantly in the coming years, it will take significant management focus to continue to mine these efficiencies.

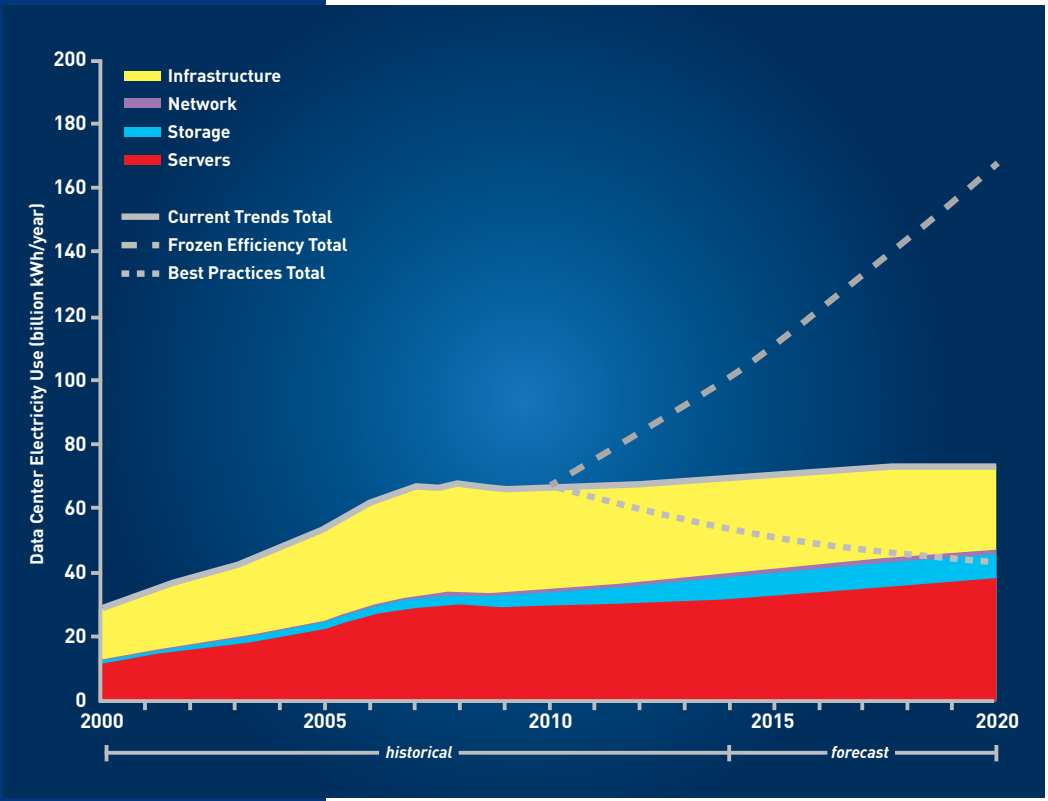
The rapid growth of data center energy consumption has also raised concerns in the environmental community. Greenpeace published a report in February 2019[4] highlighting the amount of electricity used by data centers in the Virginia Data Center Alley and their use of conventional fuels. They found that data center demand in Virginia was the largest globally and was projected to grow at 10% per year by 2021, adding 2.4 terawatt hours of electricity demand. Further, they estimated the potential electricity demand of both existing data centers and those under development in Virginia to be approaching 4.5 gigawatts or about 39.5 terawatt hours of annual energy usage. This is roughly the same power output as nine large (500-megawatt) coal power plants. However, Greenpeace is primarily focused on increasing the use of renewable energy rather than addressing the demand drivers of data center power consumption.

[4] Greenpeace. 2019. "Clicking Clean Virginia The Dirty Energy Powering Data Center Alley." https://www.greenpeace.org/usa/wp-content/uploads/2019/02/Greenpeace-Click-Clean-Virginia-2019.pdf

Data Storage Energy Consumption

Data storage is a significant component of data center power consumption. Data center energy studies define data storage as disk systems composed of Hard Disk Drives (HDD) or Solid State Disk (SSD). In their 2018 paper (A. Shehabi et al. 2018) estimated the total 2020 United States data center power consumption will be 73 billion kilowatt hours (kWh) based on the current trend, the solid line in Figure 1. The current trends scenario includes historical and projected equipment shipments with the expected efficiency improvements and is the most likely case. The "Frozen Efficiency" alternative holds energy efficiency at 2010 levels while the "Best Practice" alternative assumes widespread adoption of the most efficient technologies and best management practices. In the current trend's scenario, storage consumes approximately 19% of the data center power, including its share of cooling and other infrastructure energy consumption. Applying this percentage to the estimated 2020 data center power consumptions means storage will consume about 14 billion kWh. If this power is provided by a utility using natural gas generation[5], almost 6.5 million tons of carbon dioxide will be created by disk storage.

Figure 1 – Shehabi 2018 –
Current trends of U.S. data center equipment



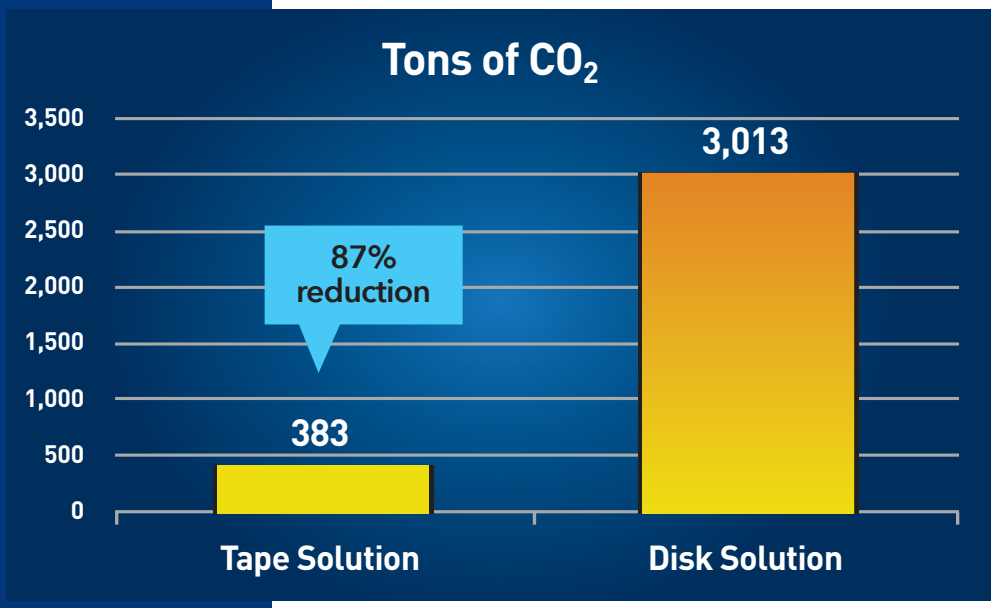
[5] U.S. Energy Information Administration. 2020. How much carbon dioxide is produced per kilowatt-hour of U.S. electricity generation? FAQ

Carbon Emission Reduction Opportunity

IDC projects[6] that the amount of corporate data would grow to 7.5 ZB by 2025, doubling every two to three years through 2025. The growth is fuelled by IoT, analytics, 5G networks, and video. Much of this data is unstructured, and as it grows older, it is infrequently accessed. They estimated up to 60% of information is seldom accessed, meaning that the expectation of access diminishes after 30 days. Moving this cold data to tape storage represents a tremendous carbon reduction opportunity.

To illustrate the carbon reduction benefit, let us start with a hypothetical example. First, we identify 10 PB of cold data residing on disk growing at 35% per year. Next, using the publicly available Total Cost of Ownership (TCO) tool from the LTO Consortium[7], we calculate the ten-year energy cost of both a tape storage solution and a disk storage solution. We then convert the respective energy costs into energy consumption based on an estimated price of $.12/kWh. In this example, the tape solution uses dramatically less power. The disk storage systems require over 6.5 million kWh of energy over ten years. The tape storage system supporting the same capacity requires just .83 million kWh. If a natural gas plant generates the power, the disk system will have produced 3,013 tons of carbon dioxide versus 383 tons for the tape system. Storing the inactive data on tape storage produces 87% less carbon dioxide. Figure 2 illustrates these reductions.

Figure 2 – Tons of CO_2 – 10 PB growing at 35% Annually for 10 Years



[6] Goodwin, Phil. 2019. Tape and Cloud: Solving Storage Problems in the Zettabyte Era of Data. IDC.

[7] LTO Ultrium TCO Calculator for Data Storage solutions. https://www.lto.org/resources/tcotool/

Reduction in Total Cost of Ownership (TCO)

Carbon dioxide emissions reductions are dramatic, but so are power cost reductions. The estimated cost of the tape solution's power over the ten years is $100,000 versus $786,098 for the disk solution, a decrease of 87%. However, power costs are not the only cost reductions. The ten-year TCO of tape storage is also significantly lower. The LTO TCO tool provides a comprehensive cost comparison that includes acquisition, maintenance, software license, extended warranty, administration, connectivity, floor space, and power costs. The estimated TCO for the tape solution is $4.8 million versus $33.8 million for the disk storage solution, an 86% reduction.

Summary

As organizations look to reduce energy consumption and carbon emissions, moving inactive data to tape storage is an exciting opportunity. Besides lowering data center energy usage and costs, modern tape provides other substantial benefits for inactive data, including high capacity, 30-year media life and the highest reliability rating of any storage medium. Also, since tape media can be easily removed from the network and stored off-line, it can be protected from malware via “air gap" protection. A review of data storage within the data center and the identification of static data in the enterprise, while not effortless, may lead to significant reductions in energy consumption and carbon dioxide emissions and help mitigate climate change. Of course, every organization will be different, but the opportunity is certainly worth exploring for any organization storing petabytes of data.

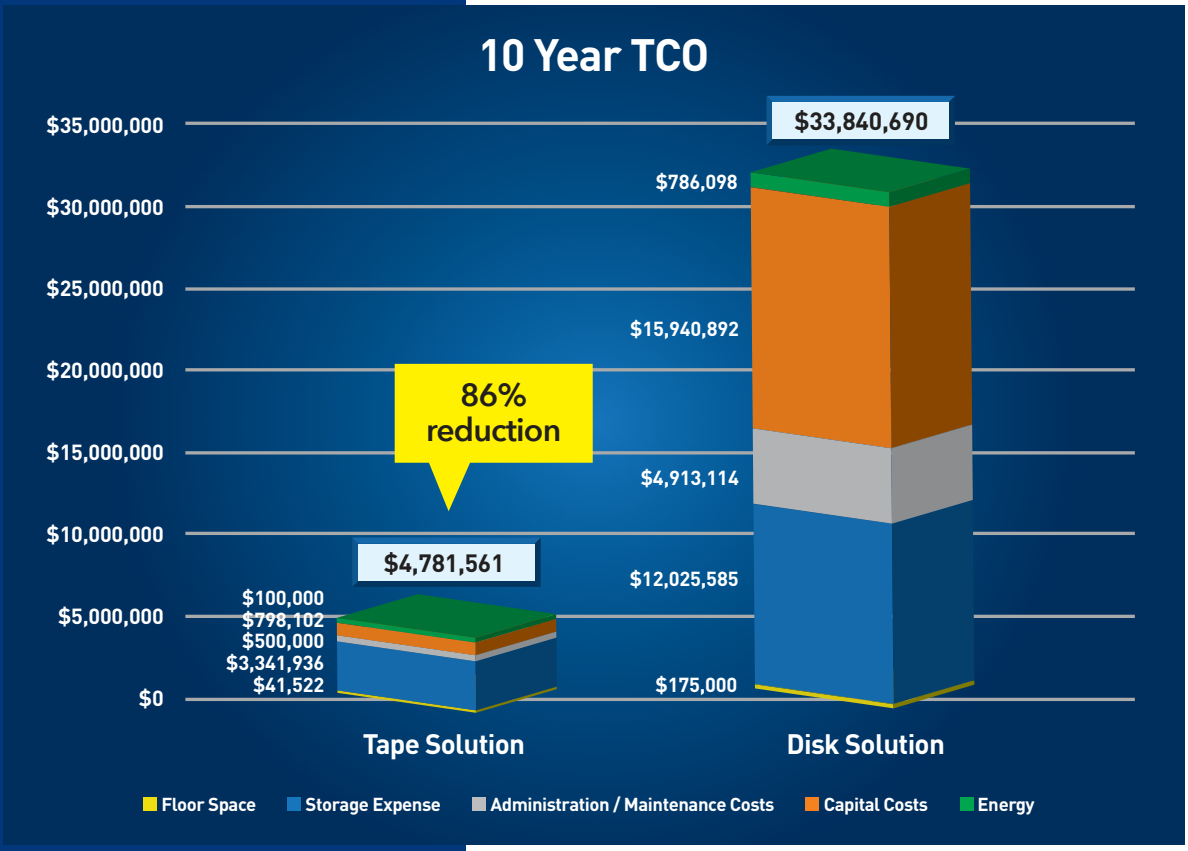


Figure 3 – Ten-year TCO for 10 PB growing at 35% annually

References

A. Shehabi et al. 2018. "Data center growth in the United States: decoupling the demand for services from electricity use. *Environmental Research Letters.*

A.Shehabi et al. 2016. United States data center energy usage report. Berkeley National Laboratory.

Eric Masanet, Arman Shehabi, Nuoa Lei, Sarah Smith and Jonathan Koomey. 2020. *Recalibrating global data center energy-use estimates.* Science 984.

Goodwin, Phil. 2019. Tape and Cloud: Solving Storage Problems in the Zettabyte Era of Data. IDC.

U.S. Energy Information Administration. 2020. *How much carbon dioxide is produced per kilowatt hour of U.S. electricity generation?* FAQ, U.S. Energy Information Administration.

U.S. Energy Information Administration. 2020. *Average Price of Electricity to Ultimate Customers by End-Use Sector, by State June 2020 and 2019.* Table, U.S. Energy Information Administration.

Regarding the information used in this report

Brad Johns Consulting L.L.C. believes that the information in this report was accurate as of the date of publication. Information is provided "AS IS" without warranty of any kind.

Trademarks and Special Notices

About the Author

Brad Johns is the owner and President of Brad Johns Consulting LLC. He has over forty years' experience in the Information Technology industry. His firm specializes in storage industry economic analysis and consulting. He started his information technology career with the Data Processing Division of IBM in 1978 and held several sales, consulting, and marketing management positions. He joined the IBM Storage Systems Division in 1997 and was responsible for product management and marketing until his retirement in 2010. He holds a Masters degree in Business Administration and a bachelors degree in Economics from the University of Arizona. His website is bradjohnsconsulting.com