DIGITAL TWINS FOR DIGITAL INFRASTRUCTURE: THE KEY TO OPTIMIZING DATA CENTER OPERATIONS

Jonathan Koomey, Ph.D.* and Karl Freund†

*Koomey Analytics

†Cambrian-AI Research

650-517-3456, jon@koomey.com

512-632-3634, Karl.freund@cambrian-ai.com

Final, October 16, 2024

ABSTRACT

The modern data center lies at the heart of today's digital global economy, performing computing tasks like e-commerce, communications, search, financial modeling, and artificial intelligence (AI). Data centers undergo constant change, both in the workloads they run 24x7 and the hardware that runs those applications.

Lack of adequate planning and management can lead to under-provisioning, over-heating, and lost capacity, all of which undermine the profitability and sustainability of these critical facilities. Today's AI and high-performance compute nodes can exacerbate these problems.

Lost data center capacity is exactly analogous to what are often called "zombie servers" in data centers, which are servers using electricity but doing nothing useful. This time it's part of the data center itself (the cooling and power infrastructure) that is costing money (and lots of it) but not enabling any useful computing.

In this paper, we describe the challenges data center planners face and the potential for digital twins to help better manage data centers over their useful lives. Combining digital twins with computational fluid dynamics software (models that simulate and predict the behavior of airflow and heat in data centers) helps planners and managers save millions of dollars, reduce energy waste, increase profitability, improve data center reliability, predict failures, and lengthen the useful lifespan of costly data center equipment.

TABLE OF CONTENTS

1. INTRODUCTION

From the earliest days of electronic computing, delivery of power and removal of heat forced design and operational choices [1, 2, 3]. This report describes a technique for applying computer modeling, already widely used in the design of electronic equipment and data centers, to optimize the *operation* of data centers, which is now rarely done [4].

This technique creates "digital twins" that imitate the characteristics and performance of data centers operating in the real world [5]. Of course, just having an exact physical or digital replica of a complex system isn't enough. A twin becomes transformational when it can be paired with computer simulations of physical systems and when it is used to drive institutional change.

With the acquisition of Future Facilities in 2022, Systems Design embraced the use of digital twins for the design and operation of data centers. This paper explores the history of that concept and describes the rationales for and benefits of using such models in operations.

2. THE DIGITAL TWIN: A JOURNEY FROM NASA TO DATA CENTERS

The concept of digital twins, virtual counterparts mirroring physical systems, has captivated industries from aerospace to healthcare. Its journey, however, began with bold leaps into space.

NASA: Pioneering the Twin Concept

While the term "digital twin" wasn't coined until later, NASA's Apollo missions in the 1960s showcased the essence of the concept with its use of "physical twins." Ground control meticulously monitored spacecraft and tested scenarios using simulations that mirrored real-time conditions.

The Apollo 13 mission highlighted the critical role such twins could play. When an explosion crippled the spacecraft while 200,000 miles out, engineers on Earth used its twin to test solutions and guide the crew to a safe return. The use of the twin was widely credited with saving the astronauts [6].

A Step Toward the Digital Twin: Applying Computers to Electronic Design

To prevent computing devices from overheating decades ago, engineers designing personal computers and other electronic devices built simple physical models using cardboard or other materials, placing fans, chips, and other components inside the box to understand airflow, heat transfer, and temperature variations. Because of the expense and complexity of this crude approach, they could only test a few design variations before deciding on the one that worked best.

At some point, the industry realized that computer simulation of the heat and airflow within the box was possible, but it wasn't simple. The physics of heat transfer was well understood in theory. However, when dealing with manufactured components installed in complex systems, modeling tools and measurements had to be combined correctly to make accurate predictions possible.

Such tools allow a designer to simulate the effects of swapping out components without physically modifying the system. Modeling computer systems has the advantage of being fast and cheap, so designers can assess hundreds or thousands of combinations before deciding on a final design.

As computers became more powerful, designers began to create computer simulations that better represented the key elements of computer systems, modeling heat flows using computational fluid dynamics (CFD). By the late 1980s, these tools had become sophisticated enough for the computer industry to rely on them instead of physical models when designing equipment [7, 8, 9]. Only later did people realize that digital twins combined with CFD could also be powerful tools when applied to operations and management, not just equipment design.

Beyond Space: The Digital Twin as a Tool for Operations

In 2002, Dr. Michael Grieves formally introduced the digital twin concept as a product lifecycle management tool, laying the groundwork for broader applications [10, 11, 12, 13]. The concept is simple, but the details are not. The objective is to create a digital model of a physical system such as a car, an engine, a factory, or a data center and apply physics algorithms to simulate the rules of how the actual components and total system behave, with the goal of maximizing operational efficiency over the life of the device, product, or building [12, 13].

Tools like CFD and finite element analysis have long been used to simulate a physical system under various conditions and stresses, including in data centers. A complete digital twin increases the complexity as multiple components act and respond to changing conditions according to the laws of the real world. These twins can be used to simulate and analyze how a machine, or an entire data center, will behave when the components are changed and upgraded, avoiding costly mistakes before the device or data center is even built.

3. RATIONALE FOR DIGITAL TWINS IN OPERATIONS

Applying digital twins to data center operations solves business problems that originated decades ago. In the 1950s and 1960s, the dominant computer vendor (IBM) designed and operated computer systems for large corporations [14]. The company was a single point of contact for all issues (both software and hardware) related to these systems. While IBM's services were expensive, their customers paid them well to ensure that everything worked seamlessly. IBM (and other vendors with similar business models) ensured that changes to all parts of the system were made with full awareness of the interconnections between different parts of the system.

With the advent of more powerful personal computers and small servers, some companies began experimenting with installing their own computer systems to avoid paying the fees of the integrated providers. The wider use of client-server architectures in the 1990s and 2000s was one gateway to these apparent cost reductions, but companies adopting this approach (and most did) failed to understand the historical benefits integrated providers of computing services delivered to their customers.

The costs of implementation and the risk of failure shifted from the integrated providers to the customers designing their own systems. Moving away from a single provider of computing services meant that costs and risks were shifted to many disparate actors within these institutions who sometimes had conflicting incentives. Coordination became difficult or impossible and total costs did not necessarily go down. While there may have been enough advantages to justify this shift, it introduced other problems, the most important of which is the creation of stranded capacity as data centers change over time [4].

The fundamental problem: IT loads diverge from the original design and change over time

The modern data center is a complex system that consumes power to enable the modern digital world, but the technology that goes into a data center is constantly changing as new types of processors, networking, and storage are introduced. A data center designed in 2010 was filled with CPUs, spinning disk drives, and low-bandwidth Ethernet. Now, that data center is filled with multi-core CPUs, power-hungry GPUs, and solid-state storage, with new equipment constantly being added.

Data center designers face an even more complex challenge than the early PC designers. When a new computer design is finalized for mass production, designers know that the configuration of the components won't change once it's built. That makes designing the system to remove heat straightforward. Unfortunately, the modern data center is like a computer with hundreds or thousands of fans, power supplies, chips, and disk drives, with all these components changing over time in ways that are almost impossible to predict, reflecting the complexity of the operational challenge.

Adding to the complexity is that each data center is unique, so a model specific to each facility must be developed and calibrated. Furthermore, each new data center is designed with a set of assumptions about computing equipment in the facility. Still, with few exceptions, the computing equipment installed once the facility is commissioned has little or no resemblance to the initial design assumptions.

A poster-child technology that challenges data center designers and operators is the Nvidia H100 GPU SXM. Each GPU has a thermal design power (TDP) of about 700W, which characterizes the maximum power dissipation per GPU in operation. In addition, other components needed to manage the GPUs, store data, and transfer data contribute to the total electricity use.

The total full-load power consumption of an H100-based NVIDIA DGX server node with 8 H100 GPUs, 2 CPUs, NICs, and SSDs running inferencing workloads is between 4 and 6 thousand watts (kilowatts or kW).[1] When installing a high-power device like the H100, data center managers and planners must consider power and cooling availability, the power density available at the individual rack, and the effect of such high-density deployments on cooling and reliability for all systems in the data center.

Technology and workloads change constantly. Last year's (2023) explosion in GPUs for AI highlighted the need for a dynamic platform to predict and manage data center changes over time. The power density for AI can be much higher than the 8 kW per rack for which most conventional

---

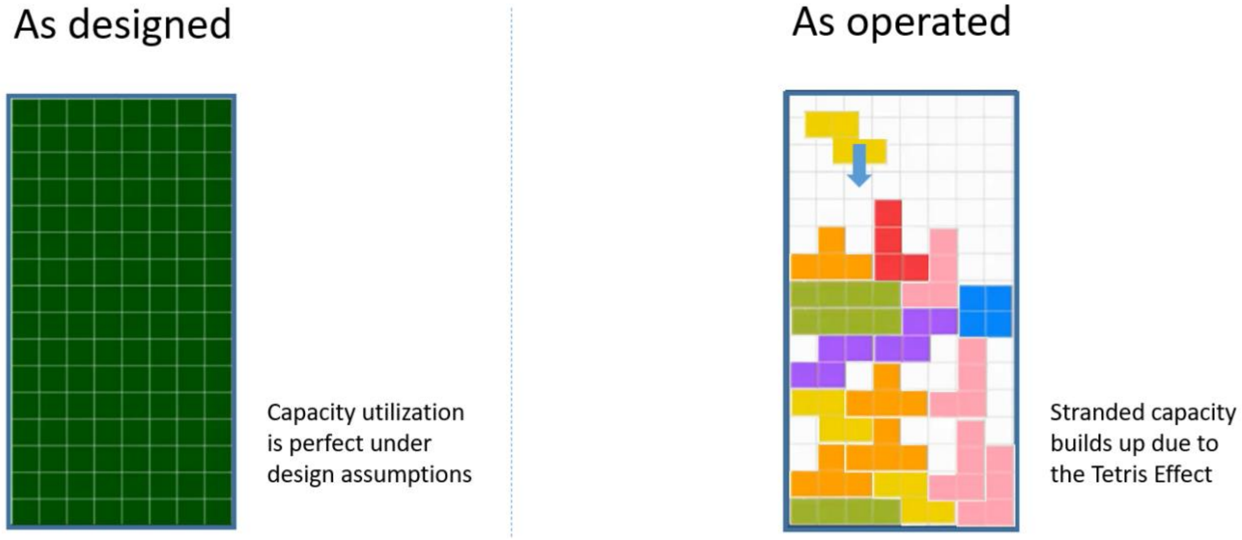
[1] The power use is measured and covers the system running a range of different AI inference applications. The related ML Perf benchmarks for training of AI systems do not yet include power measurements. https://mlcommons.org/benchmarks/inference-datacenter/. Recent measurements by Brookhaven National Laboratory for the most intense training workloads show active power around 8 kW for H100 nodes.

data centers were designed. Cooling requirements for GPUs are pushing beyond what air-cooled systems can handle; hotspots become entire *hot aisles*, reducing equipment lifespan and causing outages.

When IT loads deviate from the original data center design, stranded power and cooling capacity are the result. A simple analogy helps explain the problem. Most people are familiar with the game of Tetris $^{\text{TM}}$, in which blocks fall at a regular pace, and the player's task is to place those blocks in the correct orientation, filling up the space as thoroughly as possible.

In the simplest case, the blocks are of uniform size and shape (i.e., they conform precisely to what data center designers specified initially), and it's easy for the user to fill up the space completely. The example on the left-hand side of Figure 1 illustrates this case. On the right-hand side, the Tetris$^{\text{TM}}$ player cannot make the shapes fit perfectly because their shapes are random, and they just keep coming. That leaves gaps (white space) between the shapes, which represent *lost capacity* in the data center. White space above the colored bricks represents *unused capacity*.

Figure 1: Lost capacity as illustrated by the game of Tetris



Source: Cadence, 2024.

In virtually all companies, computing deployments are driven by perceived business needs as defined by executive leadership and implemented by the IT department, while cooling and power are (often incorrectly) assumed to be available for all possible deployments. The facilities department responsible for power and cooling is given the difficult task of managing changing computing devices. Still, they have little to no influence or control over when or how such devices are deployed, leading to lost capacity.

The dispersal of responsibility over computing services in modern deployments has another significant downside: In most companies, nobody tracks total costs and risks for changes in computing deployments. Each department knows one part of the costs, but only rarely is anyone tracking total costs because they accrue in separate budgets in different parts of the company. Ideally, the CFO should compile and track those data, but this ideal rarely happens in practice [15].

Let's say executives commission a 1 MW data center (i.e., a data center with 1 MW available to power IT loads in that facility). As soon as that data center goes online, the available capacity will be less than 1 MW if actual IT deployments differ from design assumptions. After a few years of operation, lost capacity can be a quarter to a third of the original design capacity. That means up to one-third of the capital invested in the original data center is no longer producing business value. Still, the stranded capacity and associated costs are invisible to management because total costs aren't carefully tracked.

Cultural differences and a lack of focus on total costs hamper department communication. This lack of communication makes optimization even harder. That's yet another reason why modern data center organizations need to change how they operate [15, 16, 17, 18, 19]. A digital twin creates a common language and framework for optimizing total costs and risks inside an organization.

The solution: creating and updating a digital twin for each data center

Stranded data center capacity is similar to the roughly 30% of servers in data centers that are using electricity but creating no useful computing output [20, 21], which are often called "zombie" servers. Lost infrastructure capacity is the exact analog to zombie servers, but for power and cooling. In both cases, companies pay for data center equipment that they can't use without applying the proper tools and management procedures.

The solution to the stranded capacity problem is the creation of a digital twin, combined with the ability to do predictive modeling to identify and reduce lost capacity. Using the digital twin, data center operators can test the effect of IT equipment on data center efficiency and reliability before deployment, minimizing the loss of capacity.

This approach is not as simple as deploying new equipment. It requires changes in management, operations, data collection, and communications between departments. Once these changes are made, however, data center organizations will reap the benefits of increased business agility, speed of deployment, efficiency, reliability, security, and capital utilization [5].

The key elements of successful digital twin deployments become clear from a description of the process for such deployments in existing facilities. The first step is always to inventory equipment in the data center and then to create digital representations of all existing equipment. The inventory includes IT equipment, such as servers, storage units, and network equipment, as well as cooling units, fans, pumps, and backup batteries. It also includes the physical characteristics of ducting, piping, cabling, and rooms.

To make the inventory useful, the digital twin needs data and models to characterize the behavior of all components in isolation and in use. This step requires up-to-date libraries of equipment characteristics, characterizing power use, fluid flows, and waste heat removal. In the best of all possible worlds, equipment manufacturers would supply these data as a matter of course. When manufacturers don't supply them, modelers need to estimate the key relationships and create simple models for that equipment.

The next step is to *calibrate* the digital twin to measurements. That means measuring temperature and air flows throughout the data center under different operating conditions. The model is then tuned (calibrated) to match the measurements of the current configuration. The calibration process ensures that the model matches the real characteristics and conditions in the data center, and it can be tested against the actual performance of the data center as conditions change. There are many subtleties to such measurements, but the physics is well understood [22] and trained data center operators can often collect such data with existing sensors supplemented with a few additional instruments [23].

Once the model is calibrated, it can be used to test different possible deployments virtually. Libraries of equipment characteristics are applied to each possible deployment, and the digital twin model is run to see what effect deployments would have on efficiency, reliability, and operating conditions [22].

One challenge is that the rigorous use of a digital twin requires continuous updating to reflect changes in the facility. That means tracking new installations of equipment and then re-calibrating to reflect the new conditions. It's an ongoing process that is essential to maintaining the accuracy of the predictive model [22].

A benefit of digital twins is that their proper use induces re-evaluation of incentives, institutional structures, and procedures so that they can more effectively assess the total costs, benefits, and risks of proposed IT deployments inside data centers. The digital twin provides a common language and framework for structured data center decision-making that simply doesn't exist inside most organizations.

4. REAL-WORLD APPLICATIONS

One case study of the application of a digital twin to a data center [24] showed that changing controls strategy, modifying air flows, and increasing water temperatures, all based on the results of the digital twin model, led to £380,000 per year in energy savings and a significant reduction in stranded capacity, resulting in a simple payback time of less than a year.[2]

Table 1 summarizes the details for three companies currently using digital twins. The health care and financial companies use the digital twin for both design and operations, while the hyperscale company uses it for initial design and major redesigns every few years. A common theme across these three firms is that digital twins allowed them to move away from rules of thumb (like preserving a 20% capacity buffer) and towards more accurate physics-based assessments that allowed for higher capital utilization.

The financial and healthcare firms adopted digital twin technology to cope with rising power demand at the rack level and to aid in consolidating and rationalizing their data center floor area. The hyperscale firm used the digital twin to design their facility and plans to update the model in a few years when it’s time to refresh the IT equipment (hyperscale facilities usually keep IT loads close to design specifications and don’t change them much until they swap out the IT equipment wholesale).

---

[2] £380,000 is about $600,000 US at 2012 exchange rates of 1.5852 US$/£. https://www.exchange-rates.org/exchange-rate-history/gbp-usd-2012

Table 1: Some companies now using digital twins for data center design and operations

	Units	Financial firm	Health care firm	Hyperscale data center
Number of sites	#	6	4	1
Number of digital twins	#	33	10	1
Electrically active data center floor area	Square feet	437,000	250,000	52,000
IT load	MW	50	16.7	3.5
Implied IT power density	W/sf	114	66.8	67.3
First year using digital twins		2020	2018	2023
Number of digital twin users	#	66	130	5

The rollout of the digital twin for one of the data halls owned by the financial firm illustrates the potential benefits of applying the digital twin to expanding useful capacity in existing data centers. The data center owner rolled out digital twins for more than 400,000 square feet of electrically active floor area in 2020 and early 2021, including one data hall with about 2.9 MW.

Of that 2.9 MW, about 2.4 MW of IT load was already deployed, implying that about 17% of the "as built" capacity was untapped. Analysis using the digital twin showed that deploying 0.5 MW of IT load would be possible with modest changes in airflow and cooling strategies (the existing deployment relied on rules of thumb, not predictive modeling, so it left some capacity unused).

The analysis also showed that an additional 0.25 MW of IT load could be deployed in the existing hall if the operators made more extensive airflow and cooling changes and added additional power delivery capability. In this case, modestly reducing air flows and adding blanking panels to change where air was supplied *increased* the available cooling capacity.

Capturing this additional capacity required an iterative process as well as coordination between different parts of this institution. The company needed to lay out the proposed new IT deployment in the digital twin to analyze air flows and then modify the deployment to reflect the physics of the air flows. It was not as simple as just adding power delivery capacity, and the simulations using the digital twin allowed that optimization and gave confidence that the new deployment would operate with high reliability.

5. RECOMMENDATIONS

Businesses install computers to solve business problems, but because of time constraints, technical complexity, and the fragmented nature of most data center organizations, the way computing devices are chosen, deployed, and operated is far from optimal (and quite different from the assumptions embodied in data center design). That means that the expensive capital represented by the data center itself is, in most facilities, being used at far below its maximum potential. This stranded capital is often one-third or more of the total, so capital assets worth tens or hundreds of millions of dollars are generating no financial return.

The data center industry is only now beginning to face up to this challenge. Digital twins, combined with predictive modeling, allow each data center to track its equipment, match that equipment inventory with measurements of airflow, temperature, and energy use, and analyze in advance how different possible configurations of equipment would affect stranded capacity. If such tools are properly used, a significant fraction of stranded capacity can be recaptured, but it will take changes in the way many of these facilities are designed, built, and operated. It will also require management changes–it's not just (and not even primarily) a technical problem [25].

Most existing data centers have significant stranded capacity because they haven't taken advantage of the new tools and changed their management practices. In fact, it's the change in institutional structure that is the most critical factor in driving lower cost and higher efficiency, but this is rarely within the power of data center operators to affect. That's why it's imperative that senior management in corporations with significant data center operations understand what's at stake so they can drive these changes from the top down.

The use of digital twins for data centers allows data center operators to identify stranded capacity and unlock it, thus installing more value-generating computing equipment and reducing the cost per compute cycle (which, rather than energy savings, is the ultimate goal). To accomplish this goal in most organizations requires the consolidation of budgets and authority under one manager and one department, the development of a common language around data center investments, the creation of an accurate and calibrated digital twin for each data center facility, the use of digital twins for analyzing alternative investment scenarios to guide computing deployments, and a single-minded focus by all parts of the organization on minimizing total cost per compute cycle delivered.

Achieving low cost per computation requires the use of digital twins for data center operations to minimize stranded capacity. Haphazard installation of computing equipment and poor coordination between facilities and information technology departments will inevitably result in poorly utilized capital assets, and this waste can only be stopped when the size of the potential waste is quantified and management changes are implemented to stamp it out. Digital twins, combined with predictive modeling, are the key to accurately assessing and addressing this waste.

One urgent technical need is for standardized libraries of equipment data for use in digital twin models. When building a digital twin, accurate data on equipment geometry, air flows, heat output and many other characteristics are needed, but are often not available from equipment manufacturers. In that case, those creating the digital twin need to create those data themselves, which is inefficient and duplicative. If manufacturers produced such libraries every time they introduced a new piece of equipment, all data center designers and operators attempting to create a digital twin could use them. If a well-respected standards organization defined a vendor-neutral format, the whole industry would benefit.

6. CONCLUSIONS

Innovation in modern industrial societies has been driven in the past two centuries by what economists call "general-purpose technologies", which have far-ranging effects on the way the economy produces value. The most important of these were the steam engine, the telegraph, the electric power grid, the internal combustion engine, and most recently, computers and related communications technologies [26].

One of the characteristics of general-purpose technologies is their ability to accelerate their own development by reducing costs and increasing the rate of innovation. Because of computing's unique and expanding ability to collect, process, and analyze real-time data, it may be the most powerful general-purpose technology in history.

Digital twins are an example of how computing technology, combined with management changes, can drive continuous and rapid improvements in performance, energy efficiency, and profits. We strongly believe that these tools are essential for the next phase of data center industry innovation and that they should be universally applied by those designing and operating data centers.

Despite the advancements in applications creating and using digital twins, challenges remain:

- **Data Acquisition and Integration:** Accurately characterizing physical systems and feeding these characterizations into the digital twin can be complex.

- **Model Fidelity:** Creating a digital twin that truly reflects the behavior of the physical system requires sophisticated modeling techniques, advanced software, and significant computational power.

- **Security and Privacy:** Ensuring the security and privacy of sensitive data created in the digital twins is crucial. A trusted cloud or on-prem processing is mandatory.

However, the benefits of embracing digital twins are substantial:

- **Improved Efficiency and Productivity:** By virtually testing and optimizing systems before they are built, digital twins can significantly reduce development time and costs.

- **Enhanced Predictive Maintenance**: Real-time data analysis from digital twins can predict potential equipment failures, enabling proactive maintenance and preventing costly downtime.

- **Greater Innovation:** Digital twins create a safe space for experimentation and what-if analyses, allowing engineers to push boundaries without risks to the physical system.

The future of digital twins is promising, filled with possibilities for revolutionizing how we design, build, and operate systems. By harnessing the power of virtual systems, we can unlock a new era of efficiency, reliability, innovation, and sustainability.

ACKNOWLEDGMENTS

We would like to acknowledge the financial support of Cadence Design Systems for writing this white paper. We are grateful to Sherman Ikemoto and Hassan Moezzi of Cadence for sharing insights and data over the years.

REFERENCES

1. McCartney, Scott. 2001. *ENIAC: The Triumphs And Tragedies Of The World's First Computer*. New York, NY: Berkley Publishing Group. [https://amzn.to/3IhFiHj]

2. Bell, C. Gordon, J. Craig Mudge, and John E. McNamara. 1978. *Computer Engineering: A DEC View Of Hardware Systems Design*. Bedford, MA: Digital Equipment Corporation. [https://amzn.to/3uJW3rC]

3. Haigh, Thomas, Mark Priestley, and Crispin Rope. 2016. *ENIAC In Action: Making And Remaking The Modern Computer*. Cambridge, MA: The MIT Press. [https://eniacinaction.com]

4. Koomey, Jonathan. 2013. "Modeling the Modern Facility." In *Data Center Dynamics (DCD) Focus*. November/December. pp. 102. [https://www.datacenterdynamics.com/en/news/modeling-the-modern-facility/]

5. Athavale, J., C. Bash, W. Brewer, M. Maiterth, D. Milojicic, H. Petty, and S. Sarkar. 2024. "Digital Twins for Data Centers." *Computer*. vol. 57, no. 10. pp. 151-158.

6. Lovell, Jim, and Jeffrey Kluger. 2006. *Apollo 13*. New York, NY: Mariner Books. February 20. [https://amzn.to/4btyuDm]

7. Lasance, Clemens J. M. 1990. *Optimization of Forced Convection Cooled Electronic Systems: An Experimental Study*. Proceedings of the International Heat Transfer Conference. January. [http://dx.doi.org/10.1615/IHTC9.480]

8. Lasance, Clemens J. M. 1991. *Numerical simulation of natural convection heat transfer in a television set*. Proceedings of American Society of Mechanical Engineers, Heat Transfer Division (ASME HTD).

9. Lasance, Clemens J. M. 1994. *Thermal Management of Air-Cooled Electronic Systems: New Challenges for Research*. Proceedings of Dordrecht: Springer Netherlands.

10. Grieves, Michael. 2011. *Virtually Perfect: Driving Innovative and Lean Products through Product Lifecycle Management*. Cocoa Beach, FL: Space Coast Press.

11. Grieves, Michael. 2014. *Digital Twin: Manufacturing Excellence through Virtual Factory Replication*. [http://innovate.fit.edu/plm/documents/doc_mgr/912/1411.0_Digital_Twin_White_Paper_Dr_Grieves.pdf]

12. Jones, David, Chris Snider, Aydin Nassehi, Jason Yon, and Ben Hicks. 2020. "Characterizing the Digital Twin: A Systematic Literature Review." *CIRP Journal of Manufacturing Science and Technology*. vol. 29, 2020/05/01/. pp. 36-52. [https://www.sciencedirect.com/science/article/pii/S1755581720300110]

13. Sharma, Angira, Edward Kosasih, Jie Zhang, Alexandra Brintrup, and Anisoara Calinescu. 2022. "Digital Twins: State of The Art Theory and Practice, Challenges, and Open Research

Questions." *Journal of Industrial Information Integration*. vol. 30, 2022/11/01/. pp. 100383. [https://www.sciencedirect.com/science/article/pii/S2452414X22000516]

14. Cortada, James W. 2019. *IBM: The Rise and Fall and Reinvention of a Global Icon*. Cambridge, MA: The MIT Press. [https://direct.mit.edu/books/monograph/4177/IBMThe-Rise-and-Fall-and-Reinvention-of-a-Global]

15. Koomey, Jonathan. 2015. "Facing the efficiency scandal: Enterprise data centers are shockingly wasteful." In *DCD Focus*. September. pp. 21.

16. Koomey, Jonathan, and Patrick Flynn. 2014. "How to run data center operations like a well-oiled machine." In *DCD Focus*. September/October 2014. pp. 81.

17. Schuetz, Nicole, Anna Kovaleva, and Jonathan Koomey. 2013. *eBay: A Case Study of Organizational Change Underlying Technical Infrastructure Optimization*. Stanford, CA: Steyer-Taylor Center for Energy Policy and Finance, Stanford University. September 26. [http://www.mediafire.com/view/8ema554a2ho9ifj/Stanford_eBay_Case_Study-_FINAL-130926.pdf]

18. Koomey, Jonathan. 2016. "Three Pillars of Modern Data Center Operations." In *Data Center Knowledge*. February 2. pp. [https://www.datacenterknowledge.com/archives/2016/02/02/three-pillars-modern-data-center-operations]

19. Koomey, Jonathan. 2016. "Applying the Scientific Method in Data Center Management." In *Data Center Knowledge*. March 9. pp. [https://www.datacenterknowledge.com/archives/2016/03/09/applying-scientific-method-data-center-management]

20. Koomey, Jonathan, and Jon Taylor. 2015. *New data supports finding that 30 percent of servers are 'Comatose', indicating that nearly a third of capital in enterprise data centers is wasted*. Oakland, CA: Anthesis Group. [https://www.mediafire.com/file/yczd2zmda79bcok/Case-Study_DataSupports30PercentComatoseEstimate-FINAL_06032015.pdf/file]

21. Koomey, Jonathan, and Jon Taylor. 2017. *Zombie/comatose servers redux*. Bay Area, CA: Koomey Analytics and Anthesis. April 3. [https://www.mediafire.com/file/g99fiybafzq7lp2/2017_Anthesis-Koomey_Zombie+Server+Report_Final_20170329v2.pdf/file]

22. Alissa, Husam A., Kourosh Nemati, Bahgat G. Sammakia, Tom Wu, and Mark J. Seymour. 2016. *Management and Predictions of Operational Changes and Growth in Mission Critical Facilities*. Proceedings of the 32nd Annual IEEE Semiconductor Thermal Measurement and Management Symposium (SEMI-THERM). San Jose, CA: March 14-17. [http://www.semi-therm.org]

23. Stanley, John, and Jonathan Koomey. 2009. *The Science of Measurement: Improving Data Center Performance with Continuous Monitoring and Measurement of Site Infrastructure*.

Oakland, CA: Analytics Press. October 23. [http://www.mediafire.com/file/8avb8s2hd64yvgo/scienceofmeasurement091028final.pdf/file]

24. Cane, Gary. 2012. *At the End of the Day, It's Lost Capacity: A Road Map to a Fully Deployed Data Centre*. London: CBRE Global Corporate Services. [https://www.prweb.com/releases/new_white_paper_by_cbre_provides_roadmap_to_a_fully_deployed_data_centre_using_predictive_modelling/prweb11348110.htm]

25. Koomey, Jonathan. 2014. *Bringing Enterprise Computing into the 21st Century: A Management and Sustainability Challenge*, March 17, 2014 [http://www.corporateecoforum.com/bringing-enterprise-computing-21st-century-management-sustainability-challenge/]

26. Bresnahan, Timothy F., and M. Trajtenberg. 1995. "General purpose technologies 'Engines of growth'?" *Journal of Econometrics*. vol. 65, no. 1. 1995/01/01/. pp. 83-108. [http://www.sciencedirect.com/science/article/pii/030440769401598T]