Server Technology®
A brand of legrand®

stay powered. be supported. get ahead.

Flexibility and Sustainability for the Dynamic Data Center

Written by Marc Cram, CDCD

November 2018

Executive Overview

Dynamic data center environments are always in flux – changing hardware, changing workloads, changing layouts, changing goals.

Waste not, want not. Capital efficiency at times demands that we reuse existing infrastructure, often when something shiny and new is more attractive or more convenient.

There are many parts of the data center that can be reused or refitted to accommodate the latest and greatest servers, accelerators, storage, and networking. This white paper makes the case for reducing the electronic waste stream during the refresh cycle of the data center through reuse of the racks, PDUs, and other long-life items.

Introduction

Transformation is another word for change. And like it or not, the data center is no more immune to transformation and change than is our daily lives. The physical infrastructure, the software, and the demands placed upon them both vary moment by moment, day by day. In search of the holy grail of computing, data center managers seek the new, the innovative, the revolutionary, and embrace change. Yesterday's mainframe computing gave way to the PC, which drove the need for networks, and then servers in turn. Notebooks killed the desktop, and smartphones have in turn reduced the need for laptops. To have a truly smart phone, however, requires a continuous flow of data and AI that today both live in a remote data center. Smartphones both generate and consume massive amounts of information daily, forcing data centers to adapt along the way to the rising needs of the world's two billion smartphone users.

Moore's law delivered successive generations of faster, more efficient processors for both the smartphone and the data center server. Thanks to Moore's Law, three years (two CPU generations) between server replacement was once the norm for data center owners. A recent slowdown in the CPU development cycle[1], [2]due to issues with semiconductor scaling is leading to numerous changes elsewhere in the data center as a means of growing throughput and capacity. Through software, video GPUs have been adapted to the tasks of parallel processing and blockchain applications. Field Programmable Gate Arrays (FPGAs) that were once used only for prototyping and validating circuits have been added to the data center hardware mix as a tool for accelerating specialized compute loads. And those operators with the skills and the money have taken to designing custom silicon of their own (ASICs) to further speed up the specialized functions employed by Artificial Intelligence (AI) applications.

While the hardware side of the data center changed, the software running and living in the data center evolved even more rapidly. Thanks to advances in virtualization and containerization, many applications today operate independently of the underlying hardware. For those applications, modern software no longer cares what brand of hardware it is running on, or where that hardware is physically located. Hardware has been abstracted out of the software equation by the "software defined everything" (SDE) philosophy.

"The only way you survive is you continuously transform into something else. It's this idea of continuous transformation that makes you an innovation company."

Virginia (Ginni) Rometty, IBM

1 https://www.theinquirer.net/inquirer/news/3036660/intel-10nm-cannon-lake-processors-delayed-again-until-late-2019

2 https://www.eetimes.com/document.asp?doc_id=1333637 "Global Foundries Halts 7nm Work" by Rick Merritt

Meanwhile, increasing compute workloads, growing storage demands, new applications, and societal impatience led to a throwaway mentality for much of the infrastructure that makes up the data creation and consumption ecosphere. Servers, switches, storage, load balancers, racks, and cabling have in the past been thrown away or destroyed during data center refreshes and rebuilds. The WEEE Directive in the EU began changing the behavior of the data center industry starting in the early 2000s, and subsequent changes to WEEE that went into effect in 2014 further curtailed the amount of data center and personal electronics going into landfills around the world.

In this paper, we explore the roles that flexibility and reusability will play going forward in making decisions around engineering future data center designs, and how data centers continue to work on their environmental impacts by pursuing efficiency, reusability, and waste reduction.

It Takes a Cloud to be Mobile

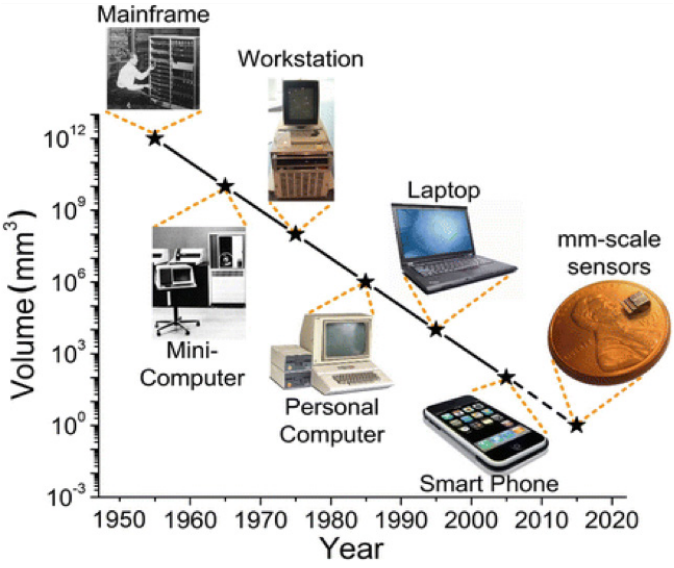
"Executive Jewelry" was once the name given to high-end ultrathin and small laptops available only to senior corporate management. Carrying a thin notebook was a sign of prestige. This is no longer the case with the widespread adoption by PC OEMs of Intel's Ultrabook platform and the migration to smartphones that began in 2007 with the launch of the first Apple iPhone.

In "Generational Change," a recent publication on TechRepublic by Patrick Gray[3], Gray asserts that

"Generational change leads to computing change – the heavy lifting of computing has moved away from the desktop to the cloud, a form of centralized computing that is reminiscent of the mainframes of the 1960s and 1970s, but at a different scale.

"One younger gentleman went so far as to note that "Laptops are for businesspeople and old guys." The general population, in particular the newer generations and the pool from which you're likely drawing your workforce, see the desktop not as the best general-purpose computing tool, but a relic of a bygone era."

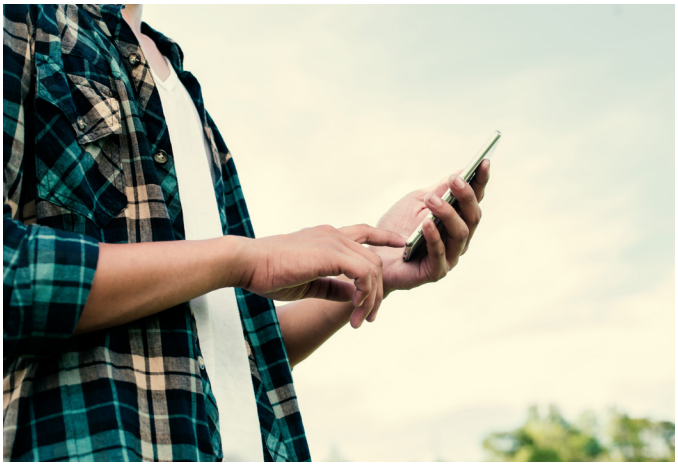
"Cloud has also created an economic disincentive to the historical cycle of upgrading corporate computers on a regular basis. If you don't need high-powered local computing, why upgrade? This is already causing the big technology companies to shift their efforts elsewhere, with desktop and laptop sales as the primary casualty of the cloud revolution."



3 https://www.techrepublic.com/article/the-desktop-is-finally-dead/?ftag=TRE684d531&bhid=280714555249799175524554821533368

With compute loads shifting away from the desktop to the data center, desktop and notebook PC refreshes are pushed out or foregone altogether. So where has the workload gone? The rise of virtualized and containerized software applications combined with the "everything-as-a-service" offerings available on Amazon, Microsoft, and Google cloud-based solutions has taken away the need for many enterprise data centers that were once viewed as being strategic for the success of many corporations. Cloud computing gives flexibility, speed, and scalability, along with providing an alternative to committing large amounts of money to capital investment required to support data center ownership.

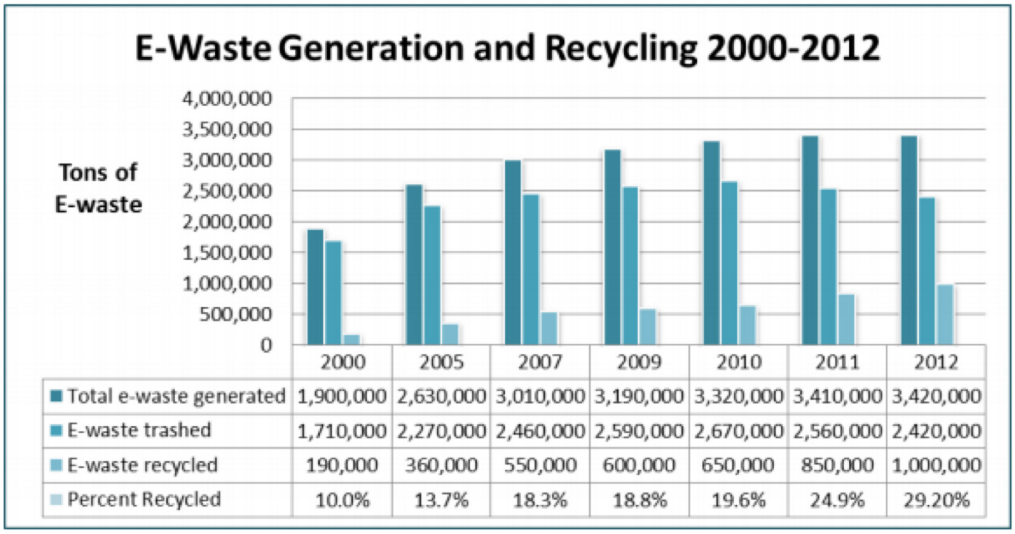
Today, disposable high-end smartphones that have a two to three-year life span are the new 'executive jewelry.' The growth of smartphone adoption has been heavily reliant on the growth in high availability cloud infrastructure. Whether a smartphone is running Android or iOS, smartphones are dependent on the constant availability of wireless internet access that is piped into large, centralized, cloud data centers belonging to Google, Microsoft, Apple, Facebook, Baidu, Alibaba, and Tencent to deliver on the promise of mobile-first applications that support content creation and consumption, social media, instant messaging, travel, banking, and e-commerce.



Hardware Refresh Cycles Impact the Dynamic and Cloudy Data Center

In 2000, the PC industry was already 18 years old. Dell had several years earlier introduced their line of rack-optimized PowerEdge servers for the data center. The 32-bit single core Pentium III and Pentium III Xeon were the processors of choice and servers offered up to 16GB of ECC SDRAM inside a 7U box.

Fast forward to 2018, and the PowerEdge brand is still offered, but thanks to the cumulative effects of Moore's Law, the processors and the memory have changed a bit. Today's top of the line PowerEdge supports Xeon Scalable Processors each with 28 cores/56 threads and have 48 DIMM slots that support up to 6TB of memory. Modern servers consolidate the work of more than 50 servers from the Y2K era, even before accounting for the impact of virtualization and containerization that enable an even greater number of processes and workloads. In 2018, more than one million new servers are being sold each quarter according to IDC and Forrester, principally into the hyperscale cloud service providers and large enterprise data centers.



	2000	2005	2007	2009	2010	2011	2012
Total e-waste generated	1,900,000	2,630,000	3,010,000	3,190,000	3,320,000	3,410,000	3,420,000
E-waste trashed	1,710,000	2,270,000	2,460,000	2,590,000	2,670,000	2,560,000	2,420,000
E-waste recycled	190,000	360,000	550,000	600,000	650,000	850,000	1,000,000
Percent Recycled	10.0%	13.7%	18.3%	18.8%	19.6%	24.9%	29.20%

Figure 1: EPA data from "Municipal Solid Waste Generation, Recycling and Disposal in the United States, 2012," Feb 2014; These EPA numbers are for "selected consumer electronics" which include products such as TVs, VCRs, DVD players, video cameras, stereo systems, telephones, and computer equipment.

How many generations of servers were thrown out by smaller data centers moving their applications to the cloud, or recycled in cloud data centers around the world in the time between 2000 and 2018? More than most of us in the industry care to admit. Figures on this are hard to come by, but if we conservatively assume a 5-year server lifespan, then at least 3 generations of servers were thrown out and entered the waste/recycling stream. It is feasible for some enterprise data centers that as many as five generations of servers were taken out of commission during that time frame, as it was once popular to replace servers every three years. But where did these obsolete systems go?

A 2014 compilation[4] of EPA released data by Electronics TakeBack Coalition (graph above) shows the magnitude of the E-Waste stream and the percentage of items recycled versus that going into the trash/landfill. The rise of conscientious companies coupled with growing regulatory requirements has diverted much of the flow of no longer useful electronics into recycling programs and away from landfills.

Despite the concerns over the waste stream, one of the many positives of data center refresh has been the reduction in PUE figures across the industry. Where it was once common to see a PUE figure in the range of 1.5 to 2.0, leading cloud providers today have data centers with PUE figures near 1.1, due principally to new data center designs and new generations of high efficiency servers that offer greater throughput while supporting higher operating temperatures.

4 http://www.electronicstakeback.com/wp-content/uploads/Facts_and_Figures_on_EWaste_and_Recycling.pdf

Driving Lower PUE

There are several data center trends that have contributed to improving (lower) PUE. Amongst them are:

New LED lighting

Higher operating temperatures

Free air cooling versus CRAC or CRAH

Higher levels of automation, requiring fewer people

Deduplication of data

Resilience through software via virtualization and containerization

Workload consolidation via virtualization and containerization

Adoption of flash storage versus rotating media

New CPUs, GPUs, RAM, and connectivity

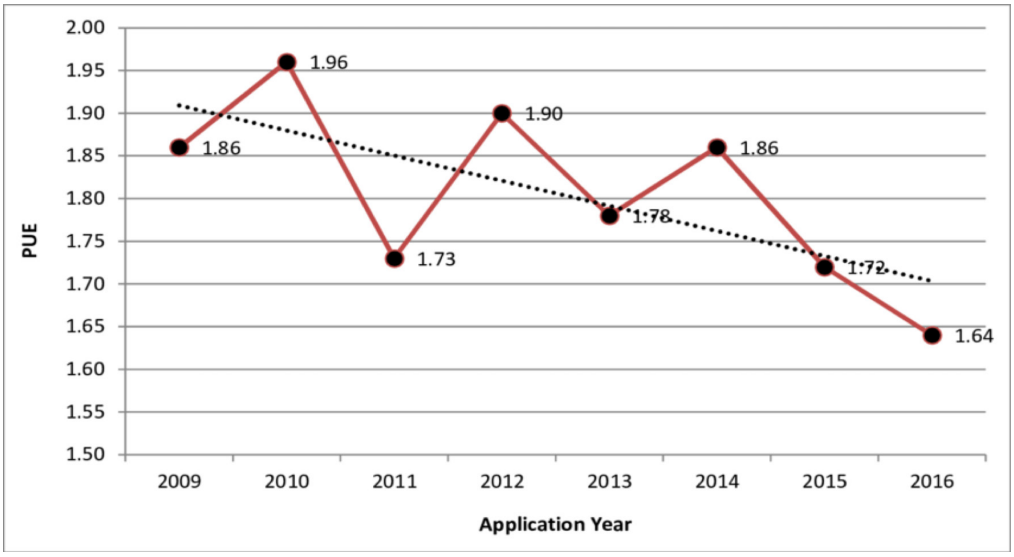


Figure 2 - Trends in Data Centre Energy Consumption

Flash storage has replaced rotating media for many applications, lowering power consumption and reducing data access times. Newer hyperscale data centers employ fewer people per server, thus requiring less light, heat, and cooling overhead required for the human occupants.

But "the best never rests," and hyperscale data center operators like Google are no exception. Not satisfied with the prior 5 years of nearly flat PUE scores (see Figure 3, below), in 2018 Google deployed AI software to operate and optimize its cooling systems. So far, Google reports that employing AI led to a further reduction in energy usage of nearly 30%.

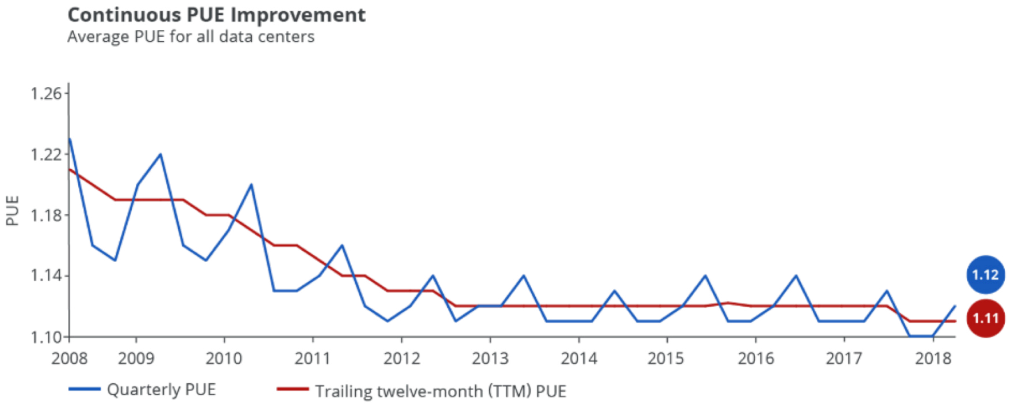


Figure 3 - Google

Growing Demands of the Data Center

While Moore's law and consolidation efforts have worked to 'densify' the data center, new applications and new workloads are working to oppose the trend and expand the size of the data center or redistribute it out to the edges.

- Big data
- IoT
- Artifcial Intelligence
- Augmented Reality
- Virtual Reality
- Video going from HD -> 4K -> 8K
- Hyperconvergence
- Disaggregation
- Silicon photonics
- In rack UPS
- Growing digital media creation and consumption
- Smart cities
- Autonomous vehicles

Corporate Responsibility and IT Infrastructure Reuse

At some point, every piece of IT hardware reaches the end of its useful life. Greenpeace offers their 10 Principles for Corporate Responsibility as a means of communicating their environmental concerns and driving change in corporate behavior. The fifth principle calls for states to require due diligence in reporting, along with cradle to grave responsibility for corporate products and services. The implications for manufacturers in the data center equipment space are clear. Likewise, for the data centers themselves. [5]IT hardware manufacturers have adopted the tenets of "Reduce, Reuse, Recycle," and so have many data center owners. In the EU, the directive on waste electrical and electronic equipment (WEEE Directive) entered into force in February of 2003. The directive was later modified in 2012, and imposed further restrictions on what materials and products could enter the landfills of Europe. As most IT products are designed and certified to operate globally, the WEEE requirements have generally been picked up and enforced in other countries outside of the EU.

While server refreshes have been good for energy efficiency, other parts of the IT infrastructure within the data center have not needed to be replaced as often. Generators, UPS, battery banks, and IT racks have a longer service life than the servers.

Taking the idea of WEEE beyond the electronics of a data center, cities and states across the USA have already taken on the challenge presented by Greenpeace and are offering incentives for data centers to reuse existing buildings such as old factories. In Chicago, [6]"buildings that once processed checks, baked bread, and printed Sears catalogs now stream Netflix and host servers engaged in financial trading."

Retrofitting these buildings as data centers is a complex task, requiring much forward thinking, and planning for a high level of flexibility and reusability of the various devices and systems therein.

5 https://www.greenpeace.org/international/story/14185/how-do-we-make-corporations-more-accountable/

6 https://theconversation.com/the-factories-of-the-past-are-turning-into-the-data-centers-of-the-future-70033

Where Will Your Data Center Be in 20 Years?

John Hawkins, VP of Corporate Marketing for vXchange, wrote "Where Will Your Data Center Be in 20 Years?" an article appearing on DataCenterKnowledge.com in 2015. In it, Hawkins cites the following about data center flexibility:

According to Pitt Turner, executive director with the Uptime Institute, there is no set lifespan to a data center. "A data center that is designed with flexibility really doesn't have a life expectancy," says Turner. "Over the life of the data center, you need to replace the capacity components just like you replace the tires on your car." Turner further states, "Chillers, UPS, that sort of stuff, needs to be replaced and you have to have a data center infrastructure that will allow you to do that." [7]

Reinforcing Hawkins' statement, "A hyperscale facility could last 15 to 20 years," says Malcolm Howe, critical systems partner at engineering consultancy Cundall. "The steel frame and paneling may last 60 years, but the IT will be updated every three to four years, recycling the servers and crushing the drives." [8]

Robert McFarlane wrote on TechTarget "When your (growing) data center runs out of space, power or cooling -- or all three -- you have some difficult decisions to make. Those deliberations become more challenging if your business is likely to move within the next several years, or if there are discussions about eventually transferring some computing to the cloud or to a hosting site. These decisions are important, and not ones you want to rush. The choices an organization makes, after all, could be costly -- in both capital outlay and operational effectiveness." [9]

The design and supplier choices made by a data center team have a direct impact on the longevity of the data center.

7 https://www.datacenterknowledge.com/archives/2015/12/09/will-data-center-20-years

8 http://www.datacenterdynamics.com/content-tracks/design-build/the-data-center-life-story/98671.fullarticle

9 https://searchdatacenter.techtarget.com/feature/Data-center-strategies-to-extend-your-facilitys-life-span

DC Design for Flexibility

There are several design elements to look for when seeking a building suitable for housing a data center. A few of them are:

- High ceilings
- Concrete floors
- Tall, wide doorways
- Minimal number of interior columns
- Large power feed from the utilities, preferably from a renewable source
- External space for generators, fuel cells, or solar
- Ambient air temps that are low

On the interior infrastructure side, a flexible data center is likely to have:

- Overhead power busway
- Overhead cable raceways
- Standardized, rectangular layouts suitable for standard IT racks

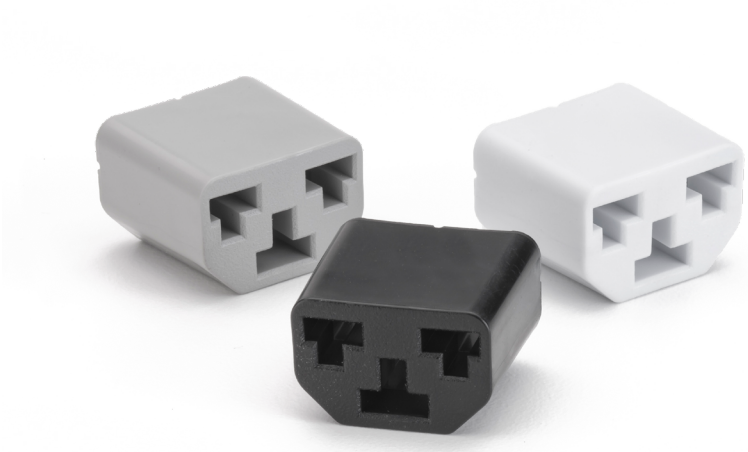
"Innovation is the change that unlocks new value."

- Jamie Notter

The Role of Flexibility in Extending the Life of the Data Center

Forward looking data center design means choosing components that are easily serviced, easily reused, and/or easily recycled. Utilizing standard 19" IT racks means that they can be sold off to another data center for re-use at the end of a data center's life, or that the metal can be sent out for recycling. Rather than using rack power distribution units (PDUs) that are custom for a given rack configuration, choosing a PDU that offers flexibility down to the outlet level can enable different rack elevations to be built over time and help avoid having to rip and replace the PDU to accommodate an equipment change in the IT cabinet. Colocation facilities with multiple tenants per rack and IT lab environments are both in need of flexible PDUs.

The innovative outlet shown below works as both a C13 or a C19 outlet, meaning the PDU doesn't have to be replaced just because two single U servers using C13 outlets were pulled out to accommodate a 2U AI accelerator box that needs a pair of C19 outlets. Instead, the innovative Cx outlets of an HDOT Cx PDU can support the new configuration without requiring adapter cords or replacing the PDU. This flexibility works well for the dynamic data center that is required to make rapid changes driven by data growth or changing workloads, and it helps avoid sending another PDU into the E-waste stream by making a single power strip that can be used in a broader variety of applications.

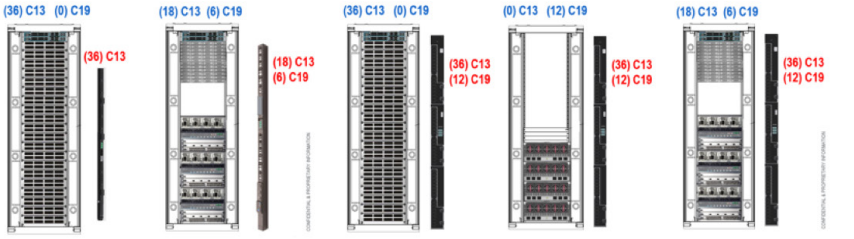


Server Technology®
A brand of legrand®

stay powered. be supported. get ahead.

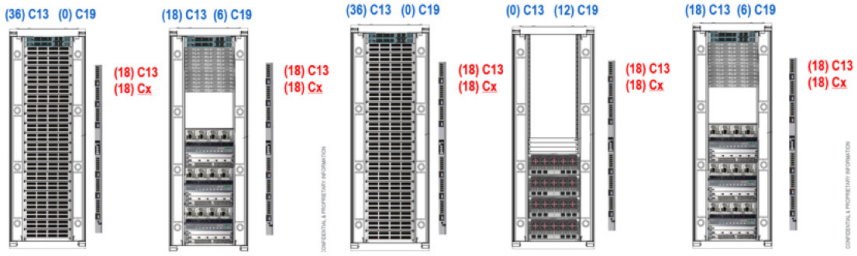
One Incredible PDU with Limitless Possibilities

Multiple Rack Configurations,
Multiple Conventional PDUS



HDOT Cx

Multiple Rack Configurations,
One HDOT Cx PDU



An HDOT Cx PDU is well-suited for:

- Dynamic data centers where frequent hardware change is occurring
- Lab environments where gear is brought for proof-of-concept or development
- Wherever outlet density and power density are both an issue
- The rack level plan is not known until the last moment

REDUCE
REUSE
RECYCLE

The Incentive for Investing Today for Reusability in the Future

Your data center is not a static environment. It is dynamic, with changing workloads, changing amounts of data going east-west within the data center as well as north-south, in and out of the data center. Not all workload expansion is going to the cloud, however, and in those cases data center operators find themselves accommodating newer, faster, denser hardware in the form of new drives, new CPUs, new memory technologies, AI accelerators such as ASICs and FPGAs[10] , and new networking topologies. Flexible infrastructure is crucial to the long-term reusability of your data center.

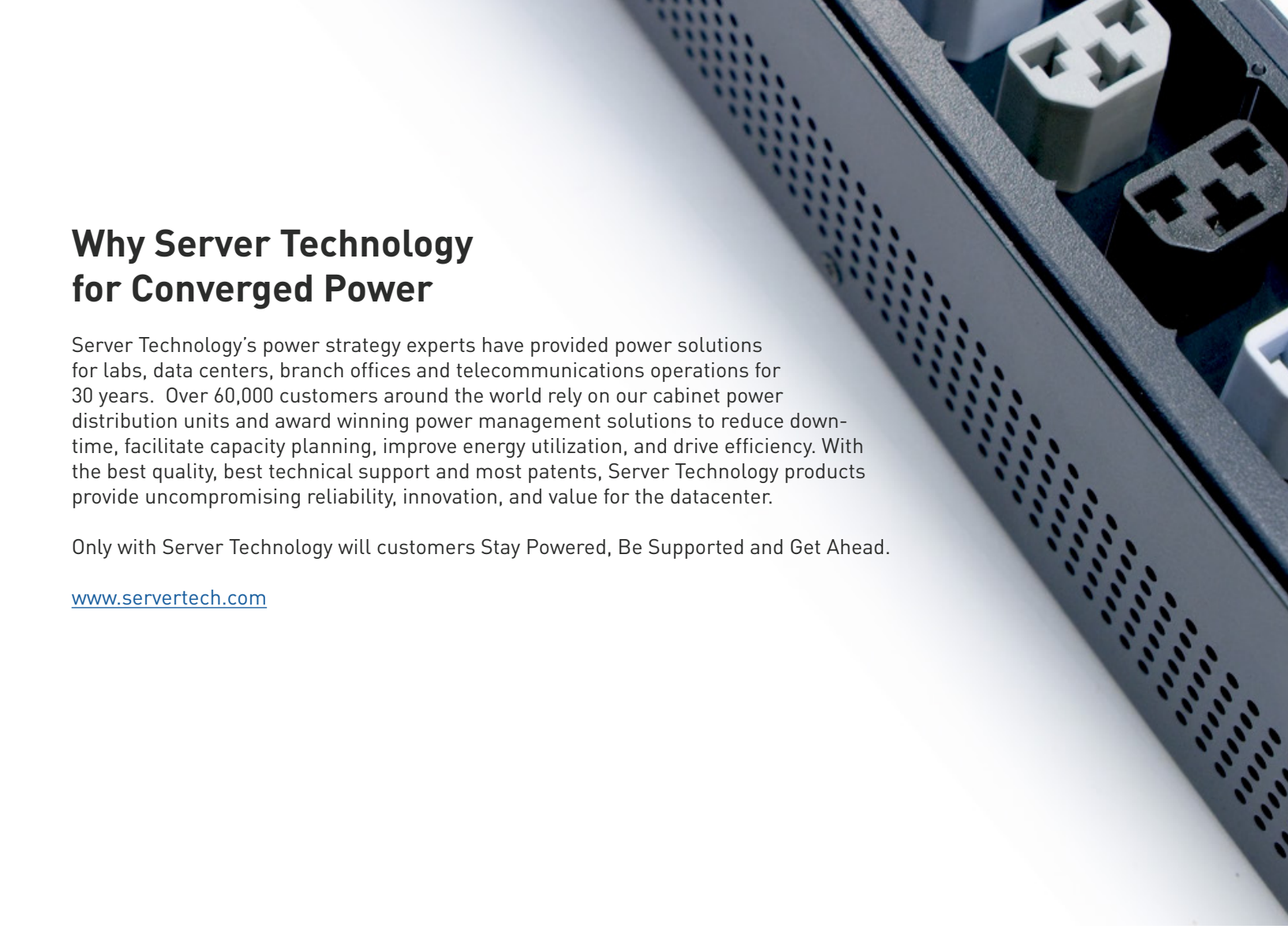
In 2018, a recent update to US tax law known as Section 179 has increased the deduction limit on equipment purchases to $2.5M. For the medium-sized data center operator, this potentially translates into thousands of new unpopulated racks, or thousands of highly functional intelligent PDUs for your data center.

In addition to the expanded Section 179 deduction, the IRS increased bonus depreciation from 50% to 100% for assets placed in service prior to 1/1/2023. The benefit of the bonus depreciation, as opposed to the Section 179 is that there is no limit on the purchases you make. This represents a great opportunity for the small to medium businesses to update their infrastructure today.

If you've not already done so, talk to your accounting and finance teams today to see what the tax law changes mean for your enterprise. At Legrand, we believe that now is the time to invest in powering your data center with HDOT Cx. The future of your data center depends on it!

Why Legrand

10 https://itpeernetwork.intel.com/inteldcisummit-fpga/?cid=em-elq-36376&utm_source=elq&utm_medium=email&utm_campaign=36376&elq_cid=3627084



Why Server Technology for Converged Power

Server Technology's power strategy experts have provided power solutions for labs, data centers, branch offices and telecommunications operations for 30 years. Over 60,000 customers around the world rely on our cabinet power distribution units and award winning power management solutions to reduce down-time, facilitate capacity planning, improve energy utilization, and drive efficiency. With the best quality, best technical support and most patents, Server Technology products provide uncompromising reliability, innovation, and value for the datacenter.

Only with Server Technology will customers Stay Powered, Be Supported and Get Ahead.

www.servertech.com

Interested in learning more about how Server Technology can help you manage and distribute power in your datacenter?
Visit us online at: www.servertech.com/products

North America Headquarters
1040 Sandhill Road
Reno, Nevada 89521
1-775-284-2000 Tel
1-800-835-1515 Toll Free
1-775-284-2065 Fax
sales@servertech.com
www.servertech.com
www.servertechblog.com

U.K., Western Europe, Israel, & Africa
Fountain Court
2 Victoria Square
Victoria Street
St. Albans
AL1 3TF
United Kingdom
+44 (0) 1727 884676 Tel
+44 (0) 1727 220815 Fax
salesint@servertech.com

Server Technology®
A brand of legrand®

stay powered. be supported. get ahead.