Life Is On Schneider Electric

How 6 AI Attributes Change Data Center Design

White Paper 110
Version 3

Data Center Research & Strategy

by Victor Avelar
Patrick Donovan
Wendy Torell
Maria A. Torres Arango

Executive summary

From large training clusters to small edge inference servers, AI is becoming a larger percentage of data center workloads. This represents a shift to higher rack power densities. AI start-ups, enterprises, colocation providers, and internet giants must consider the impact of these densities on the design and management of the data center physical infrastructure. This paper explains the attributes and trends of AI applications that impact data center physical infrastructure. We then present key design considerations for power, cooling, and rack systems. Finally, we explain how good supply chain practices, software management tools, and services mitigate risks inherent in AI deployments and highlight future physical infrastructure.

RATE THIS PAPER ★★★★★

Key takeaways

1. **Six key AI attributes and trends impact physical infrastructure**: These are driving rack densities past 100 kW. Rack density, among other impacts, force data center stakeholders to make changes to power, cooling, and rack infrastructure. This also necessitates changes to operations.

2. **Energy procurement is challenging due to the immense power demands of AI data centers:** Data center developers must secure hundreds of megawatts, which can result in multi-year delays for grid connections. To mitigate this risk, developers must plan their power sourcing, including the option of building on-site prime power generation.

3. **AI workloads require that you upgrade data center electrical infrastructure**: The confluence of rack density with peaky, synchronous power consumption, and increased arc flash hazards, challenges traditional electrical infrastructure. Existing data centers must upgrade internal distribution systems (e.g., higher voltages, larger power blocks) to support these workloads, especially in older facilities.

4. **AI's escalating thermal demands are driving the adoption of liquid cooling solutions**: This shift introduces significant complexities, making it more important to leverage partner expertise. The absence of industry-wide standards for coolant properties, interfaces, and integration drives up costs and poses retrofitting challenges. Data center operators should conduct thorough design assessments and leverage partner expertise with advanced tools to overcome these hurdles and ensure effective, compatible cooling solutions for high-density AI.

5. **High-density AI requires stronger racks**: AI workloads are pushing rack densities and weights to over 100 kW and over 1300 kg, demanding wider, deeper, and taller racks with higher weight capacities. Data center floors must also support these extreme loads. You can help future-proof your infrastructure by proactively specifying and independently validating it against evolving AI deployments.

6. **Leveraging an ecosystem of partners and tools reduces project risks**: Designers should integrate IT and physical infrastructure design early on in the project due to increasing lead times and complexity. You can help mitigate project delays through effective supply chain management, leveraging standardization, and global partners. Advanced software tools like EPMS and DCIM, including digital twins, are vital for accurate capacity assessment, dynamic load management, and reduction of operational risks. Outsourcing these functions through service contracts offers five benefits.

Introduction

This paper focuses on AI applications that process large amounts of data in a data center, for purposes such as drug discovery, climate modeling, training large language models (LLMs), and large-scale video analytics. The impact certain AI attributes and trends have on power, cooling, and rack infrastructure, as well as operations and maintenance, creates challenges for data center design. This paper examines these attributes and trends and presents related physical infrastructure design considerations. We then show how supply chain management, data center design and management software tools, and related services, become more vital to reduce risk and complexity in AI deployments.

Finally, we provide a forward-looking view of what's to come in data center design. Note, this paper is not about applying AI to physical infrastructure systems. While

next-generation physical infrastructure systems will continue to advance and leverage more AI, this paper focuses on supporting AI workloads with *existing* power, cooling, and rack systems available today.

AI attributes and trends

Six key AI attributes and trends impact physical infrastructure (**Figure 1**):

- AI workloads
- Accelerator network communication
- Thermal design power (TDP) of accelerators
- Peak power of accelerators
- Synchronous computation
- AI cluster size

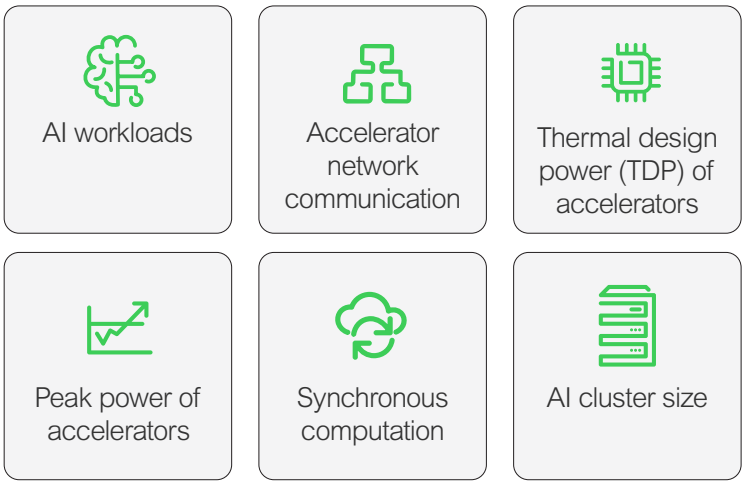
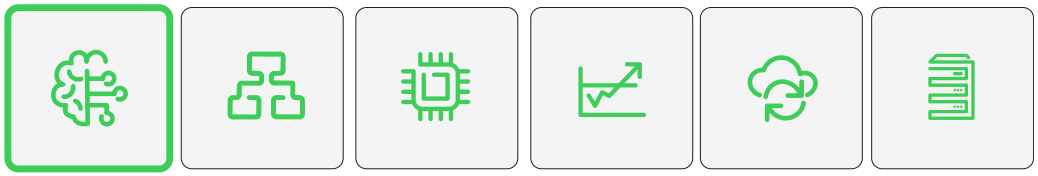


Figure 1

Six key AI attributes and trends impact physical infrastructure



AI workloads

AI workloads fall into two general categories: training and inference. **Training** workloads encompass pre-training and post-training (e.g., fine-tuning) of AI models like large LLMs. When we talk about training workloads here, we are talking about distributed training - or many machines working in parallel toward a single purpose.[1] To make this possible, those machines need to process massive amounts of data - and that's where accelerators come in. The GPU is a common example - a graphics processing unit.[2] GPUs and other accelerators, are built to do many operations at once, making them perfect for the kind of parallel processing enabling LLM training.

[1] The large number of parameters and tokens in a model require that the processing workload be split up across many GPUs to decrease the time it takes to train the model.

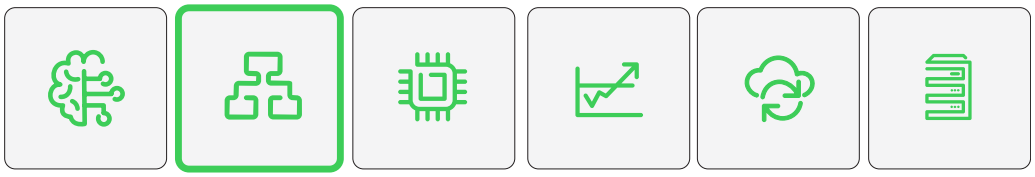
[2] Other examples of accelerators are tensor processing units (TPUs), field programmable gate arrays (FPGAs), and application-specific integrated circuits (ASICs).

In addition to servers, training also requires data storage and a network to connect it all together. These elements are assembled into an array of racks known as an AI cluster which trains a model as a single computer. The accelerators in a *well-designed* AI cluster operate at high utilization for most of its training duration, which ranges from hours to months. The accelerator model and quantity impact rack densities in AI clusters, with some exceeding 100 kW. Clusters can range from a few racks to thousands of racks that consume hundreds of megawatts of power. Training workloads save their models as "checkpoints." If a cluster fails or loses power, it can continue from where it left off. This feature makes redundancy and UPS backup less critical.

Inference workloads process a user's input (i.e., query) and provide a response. Placing a trained model into "production" typically requires compressing it to reduce the server's memory requirements.[3] If the original pre-trained model was used for inference, it would require an enormous compute cluster which is economically impractical. From the user's perspective, there may be a trade-off between an output's accuracy and the inference time (i.e., latency). If I'm a scientist, I may be willing to wait longer in between queries to receive more accurate outputs. Long-thinking *inferencing* is an example of this. On the other hand, if I'm a consumer looking for a great hotel, I may want a chatbot with instant answers. In short, the user's need determines the inference model, but very rarely is the original pre-trained model used.

Inference workloads tend to use accelerators for large models, and may also depend heavily on CPUs, depending on the application. Applications including autonomous vehicles, recommendation engines, and ChatGPT leverage IT stacks, "tuned" to their requirements. Depending on the size of the model, hardware requirements per instance[4] can range from a single server to several racks of servers.

This means that the rack densities can range from a few hundred watts to over 100 kW. Unlike training, the number of inference servers increase with the number of users or queries. In fact, a popular model (e.g., ChatGPT) often requires many more times the quantity of racks for inference as it did for training, given it handles over one billion queries per day. Finally, inference workloads are often business-critical and that requires resiliency (e.g., delivered by a UPS and/or geographic redundancy). For more information on inference see Executive Report 5, *Generative AI Inferencing Ramp-up: A CIOs Guide to Physical Infrastructure Considerations*.



Accelerator network communication

With AI workloads, every accelerator requires a network port to establish the compute network fabric. For example, if an AI server has eight GPUs, that server needs eight compute network ports. This compute fabric allows all accelerators in a large AI cluster to communicate in concert at high speeds (e.g., 900 gigabytes/second). Accelerators can process data and generate tokens (see sidebar) at a much faster rate than the communication speed between them. This makes accelerator network

What are tokens?

"Tokens are tiny units of data that come from breaking down bigger chunks of information. AI models process tokens to learn the relationships between them and unlock capabilities including prediction, generation and reasoning. The faster tokens can be processed, the faster models can learn and respond." – source *Nvidia*

[3] Compression techniques decrease the model's accuracy and the required amount of memory.

[4] A single query submitted to a model for inference.

communication latency a bottleneck that determines how fast you can complete a given workload.

To reduce the time and cost of training models, accelerator processing speeds need to increase in tandem with network speeds. For example, using GPUs that process data from memory at 900 GB/s with a 100 GB/s compute fabric decreases the average GPU utilization because it's waiting on the network to orchestrate what the GPUs do next. This is like buying a 500-horsepower autonomous vehicle with an array of fast sensors communicating over a slow network; the network speed will limit the car's speed, and therefore won't fully use the engine's potential power.

High-speed network cables are expensive. For example, InfiniBand optical connections run 10 times the cost of copper and they consume more power. A cost-effective and power-efficient approach to decrease this latency within the rack, is to use copper network cables across the short distances. However, if we used copper to connect the racks to each other (*inter-rack* communication), the longer distances would introduce untenable latency.

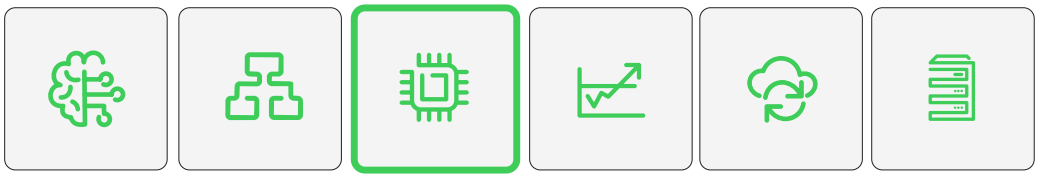
This means we need faster, more expensive, and energy-intensive, *inter-rack* communication (e.g. fiber). Hence the incentive to maximize the number of accelerators in a rack, leading to higher rack densities in AI clusters. This attribute applies to training as well as inference.

With inference, the speed of each query relies on far fewer accelerators because the models are compressed and require much less memory to store. The inference model size (e.g., terabytes) and the server's memory capacity tend to determine the server quantity. Table 1 provides simplistic examples of inference servers required as a function of model size. This assumes eight 80 GB GPUs per server and 10 kW per server. Note that spreading the servers out to multiple racks would increase the latency and energy per query.

Table 1

Number of servers needed as a function of model size (gigabytes of required memory)

Model size	# of servers	kW/rack
600 Gigabytes	1	10
1.2 Terabytes	2	20
2.4 Terabytes	4	40



Thermal design power (TDP) of accelerators

While training or inference is impossible without storage and network, the accelerators represent nearly half of an AI cluster's power consumption.[5] Accelerator power is trending higher with every new generation. A chip's power consumption, measured in watts, is commonly specified with TDP. The increasing TDP trend is a consequence of designing accelerators for an increased number of operations, to train models and infer answers in less time and with lower cost. Table 2 illustrates this

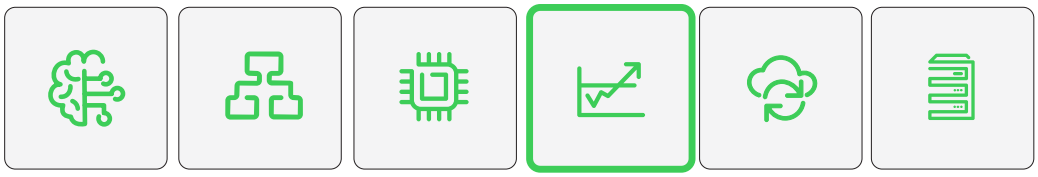
[5] E.g., 8 x B200 GPUs (8 kW) represent ~60% of a server's 1N power capacity (13.2kW)

trend by comparing five generations of NVIDIA GPUs in terms of TDP and performance.[6]

Table 2

TDP and performance across different generations of GPUs

GPU	TDP (W)[7]	TFLOPS[8] (Training)	Performance over V100	TOPS[9] (Inference)	Performance over V100
V100 SXM2 32GB	300	15.7	1X	62	1X
A100 SXM 80GB	400	156	10X	624	10X
H100 SXM 80GB	700	500	32X	2,000	32X
B200 SXM 180GB	1,000	1,125	72X	4,500	73X
B300 SXM 288GB	1,400	1,880	120X	7,500	121X

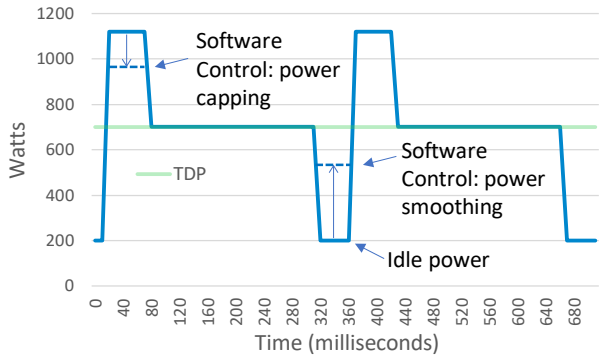


Peak power of accelerators

While all IT workloads exhibit peak power consumption, accelerators used in AI workloads may exceed their TDP multiple times per second and then fall to an idle state. This higher threshold is referred to as the electrical design point (EDP) and these high frequency changes can be described as transients. These transients may exceed the steady state TDP (e.g., 50%) for tens of milliseconds while not violating the average long-term TDP thermal limits. Figure 2 illustrates an example of power peaks that exceed TDP and later fall to an idle power state. The profile, duration, and frequency of these peaks and valleys will vary depending on some key variables. These include IT hardware (i.e., GPUs, power supplies, storage, and network), AI workload, and software limits imposed on loads (i.e., power limits). The magnitude of these peaks and valleys will be lower when measured at the data center level (due to other IT hardware used in an AI cluster such as network switches and storage). This other equipment dilutes the overall variance of the accelerators because their power profile doesn't vary as much.

Figure 2

Example of accelerator peaks and idle states shown in millisecond timescale (green line represents TDP)

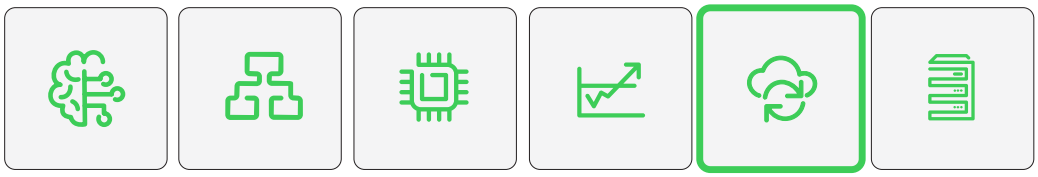


[6] While the GPU is key to these performance gains, other system improvements were made to take advantage of improved GPUs such as increasing memory and inter-GPU communication.

[7] V100, A100, H100, B200, B300

[8] TFLOPS - tera (trillion) floating-point operations per second - measure of matrix multiplication throughput at tensor float 32 (TF32) precision, generally used with training workloads. V100, A100, H100, B200, B300

[9] TOPS is tera (trillion) operations per second is a measure of integer math throughput at 8-bit integer (INT8) precision, generally used with inference workloads. V100, A100, H100, B200, B300



Synchronous computation

The power consumption pattern for virtually all IT equipment and workloads resembles a series of peaks and valleys. These peaks and valleys can either be random or synchronized. The power consumption pattern for traditional IT workloads is asynchronous, meaning that the power consumption peaks occur at different times and don't coincide (i.e., loads are diversified).

For example, while individual servers may have a power variance of 60% between idle and full load, the aggregate load from all servers (as seen by a UPS) will have a lower variance. The probability that all these peaks occur at the same time is low. This asynchronous pattern allows data center designers to "oversubscribe" power and cooling systems like UPSs and chillers.

In contrast, certain AI workloads may result in a synchronized power consumption pattern for all servers in an AI training cluster. This means that peak power draw occurs at the same time many times per second, acting like quick step loads. Figure 3 provides a hypothetical illustration of how the sum (blue line) of all the workloads varies only slightly for asynchronous workloads but for synchronous workloads, the sum is highly variable.

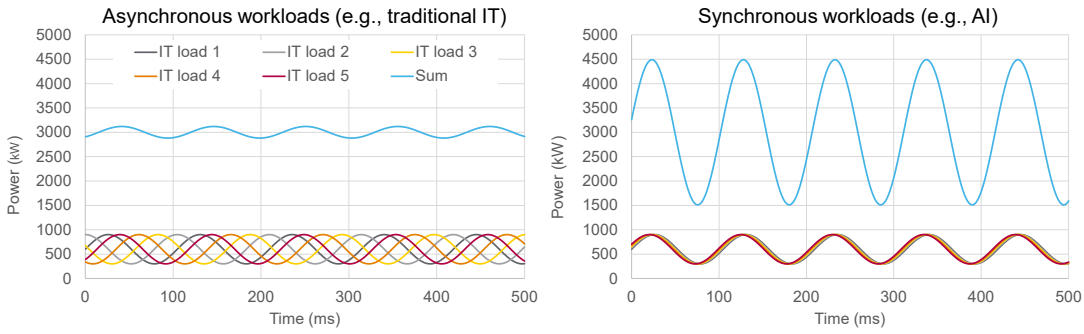
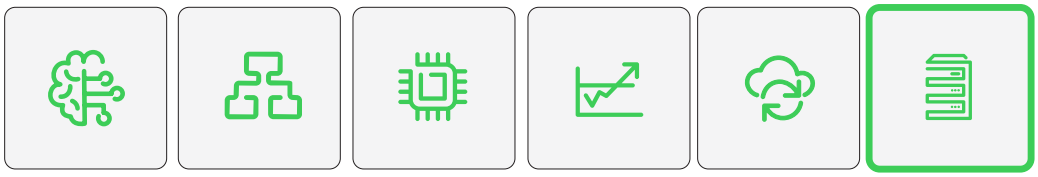


Figure 3

Hypothetical comparison between asynchronous and synchronous workloads



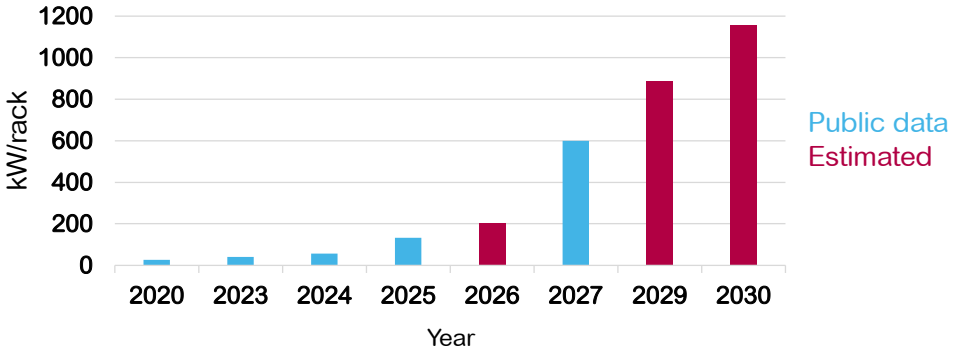
AI cluster size

As discussed above, certain AI workloads can be large, parallel processes that extend beyond single servers, potentially utilizing thousands of accelerators. For example, pre-training LLMs can require a dedicated data center loaded to nearly 100%. Depending on the data center size, even a modest AI training cluster represents a significant percentage of data center load. At higher rack densities this presents some challenges.

As shown in **Figure 4**, AI compute rack densities have continued to increase.[10] Both academia and industry are pursuing solutions to increase the power efficiency of AI workloads, especially training.[11,12,13,14] However, even with increased efficiency, two key drivers will likely perpetuate the upward rack density trend: increasing TDPs and the incentive to reduce accelerator network latency.

Figure 4

kW/rack trend for AI compute racks



Most traditional data centers today can support peak rack power densities of about 10 to 20 kW.[15] However, deploying tens or hundreds of racks all greater than 40 kW in an AI cluster presents physical infrastructure challenges to data center operators.

These challenges are not insurmountable, but operators will benefit from a deeper understanding of AI cluster requirements, not only with respect to IT, but to physical infrastructure, especially for existing data center facilities.

Ultimately, the older the facility, the harder it will be to support large AI clusters.

Design considerations

With an understanding of these AI attributes and trends, we can now review the related physical infrastructure design considerations. The following sections are broken up by power, cooling, and rack infrastructure. A common theme throughout this paper is to leverage an ecosystem of partners throughout the design process.

As experts in their respective fields and with visibility into their own product roadmaps, they can help you anticipate potential pitfalls. It's especially beneficial to get advice from vendors who collaborate with each other. These relationships often result in insights that save time, energy, reduce risk of failure, etc.

Power design considerations

The first three trends and attributes (**AI workloads**, **accelerator network communication**, **TDP of accelerators)** contribute to rack densities over 100 kW. The last three (**peak power of accelerators**, **synchronous computation**, **AI cluster size**) contribute to power anomalies. These trends and attributes have implications at

[10] Estimated from a combination of stated rack densities and AI training performance (floating point operations per second) across various GPU generations.

[11] W.D. Heavenarchive, *Four reasons to be optimistic about AI's energy usage*, MIT Technology Review, 5/2025

[12] *Training neural networks more efficiently*, Technical University of Munich, 3/2025

[13] D. Smith, *Up to 30% of the power used to train AI is wasted. Here's how to fix it*, Univ. of Michigan, 11/2024

[14] *New approach to training AI could significantly reduce time and energy involved*, King's College London, 1/2025

[15] Uptime Institute, *Rack Density is Rising*, 12/2022

virtually all parts of a data center's power system, including the procurement of electricity and the power systems behind the utility meter.

Electricity procurement

If AI training clusters could be split across different geographies (i.e., no accelerator network latency penalty), we wouldn't need massive data center campuses for AI training. The need for low latency is a key reason why AI data center developers must procure hundreds of megawatts of power from electric utility providers.

The unprecedented demand for powering these data center campuses has contributed to grid connection queues of 3-5 years in energy-constrained regions[16]. This presents data center developers with a dual challenge: deploying new facilities quickly and procuring electricity in a way that minimizes carbon emissions. The larger the power requirements, the more challenging it is to acquire energy of any kind. Here are 5 resources that detail power procurement strategies to help address this, with the last two being deeper technical documents.

- Executive Report 1 - *The Looming Power Crunch: Solutions for Data Center Expansion in an Energy-Constrained World*
- Executive Report 2 - *Assessing the U.S. Power System's Ability to Support Data Center Growth*
- Executive Report 10 - *Navigating Data Center Energy Constraints: Considerations for On-Site Prime Power* - guidance on forming and selecting an alternative prime power technologies and fuels
- White Paper 212 - *Bending the Energy Curve: Decoupling Digitalization Trends from Data Center Energy Growth*
- White Paper 202 - *How to Choose IT Rack Power Distribution*

Power system

The power system inside the data center must comprehend things like peaky loads, step loads, and increased arc flash hazards. Traditional data centers weren't designed to support these types of anomalies. For example, some existing North American data centers still distribute power at 120/208V and will need to increase the distribution voltage. The power distribution "block" sizes also need to increase to support an increased number of 100 kW racks. Even new data center designs must be updated to accommodate the requirements of AI workloads. These and other considerations are detailed in the following white papers.

- White Paper 128 - *High-Efficiency AC Power Distribution for Data Centers*
- White Paper 126 - *Retrofitting Existing Power Systems for AI Clusters*

Cooling design considerations

The cooling system is primarily impacted by two AI trends and attributes: **accelerator network communication,** and **TDP of accelerators**. Together they are driving the need for direct-to-chip liquid-cooled AI servers, presenting a major change to traditional data center cooling. It's important to understand why air-cooled servers don't meet this need.

Is it impossible to cool higher TDPs with air? Not at all. Given a tall-enough heat sink and enough airflow, air *could* cool higher TDPs. The key issue is that taller heat sinks translate into taller servers. Taller servers are incompatible with the

[16] P. Donovan, *The Looming Power Crunch*, Schneider Electric, p. 4

"accelerator network communication" attribute. This attribute aims to maximize the number of accelerators in a rack. Given the taller heat sinks, you can fit less servers, and therefore less accelerators, in a rack. The following example articulates this point. Assume an air-cooled server height is 10U (44 cm) and it has 8 GPUs. You could stack four of these servers in a 42U rack, allowing a total of 32 GPUs.

With liquid-cooling, the heat sinks, known as cold plates, are significantly smaller (less than an inch tall). With the cold plate's low profile, assume a liquid-cooled server height is 2U (89 mm) and has 8 GPUs. You could stack 21 of these servers in a 42U rack, allowing a total of 168 GPUs. This hypothetical example highlights the value of liquid-cooled IT equipment, but it has some unique requirements.

Compared to traditional chilled water systems, direct-to-chip liquid-cooled servers have more stringent requirements in coolant temperature, flow, and chemistry. This means that operators cannot run water directly from a chiller system through a chip's cold plate.[17] The coolant distribution unit (CDU) isolates the two coolant streams to prevent contamination.

The lack of liquid cooling standards also drives integration complexity and cost (both in retrofit *and* new data centers). There are many different stakeholders in the liquid-cooling ecosystem including chip manufacturers, server OEMs, liquid-cooling hardware vendors, consulting engineering firms, and data center operators. Reaching consensus on key design and operation attributes takes time and is further complicated with the fast pace of chip development. Some examples of this complexity include:

- Some materials and fluid chemistries can react leading to galvanic corrosion.
- Costs can rise quickly because systems are highly customized. This adds costs in design, installation, and testing, increasing capital costs but also potentially support and service costs over time.
- In a mixed server environment, there may be a mix of liquid flow and temperature requirements, that may be more challenging to support.
- Some IT equipment still requires air-cooling and must coexist with the liquid cooling system.

All these complexities and uncertainties are eventually baked into the cost of hardware, software, and services. We recommend data center operators perform a design assessment of the proposed liquid-cooled loads and the facility's existing conditions in advance of deploying liquid cooling. Expert review is essential to evaluate possible designs and avoid the cost implications of unforeseen constraints.

Like many new technologies, it takes time to develop liquid cooling standards for attributes such as thermal ratings, physical dimensions, interfaces, and liquid properties (e.g., temperatures, flowrates, pH levels). As we await a critical mass of standards, data center stakeholders should leverage the experience of their partner ecosystem and their associated tools. These collaborative networks offer crucial support while formal standards are still evolving, partly through the expertise from past projects and partly through their tools.

Thermal modeling and fluid dynamics are examples of these tools. The following documents provide deeper discussion on liquid cooling and related challenges.

[17] Running untreated water through a server's cold plate can cause corrosion, biological growth, and fouling. All of these compromise heat transfer from the GPUs, eventually leading the GPUs to throttle or shutdown to prevent damage.

- White Paper 210 - *Direct Liquid Cooling System Challenges in Data Centers*
- White Paper 133 - *Data Center Design Practices for Integrating Liquid-cooled AI Workloads*
- Executive Report 3 - *Optimizing AI Infrastructure: The Critical Role of Liquid Cooling*
- Executive Report 5 - *Generative AI Inferencing Ramp-up: A CIOs Guide to Physical Infrastructure Consideration*

Rack design considerations

IT racks are impacted by three AI trends and attributes: **AI workloads**, **accelerator network communication** and rising **TDPs of accelerators**. Together these three are driving rack densities and weights over 100 kW and over 1,800 kg (4,000 lb). As discussed in the cooling section, the trend is toward more accelerators/equipment in the rack, thereby pushing beyond the limits of traditional IT racks.

Specifically, this places limits on rack width, depth, height, and weight load. For example, there's less space in the back of the rack to mount the required rack PDUs and liquid cooling manifolds. As server power ratings continue to increase, it will become very difficult if not impossible to accommodate the necessary power and cooling distribution in a standard rack. Furthermore, for hybrid-cooled servers (air & liquid), narrow racks are likely to congest the exhaust airflow behind the rack due to power and network cables.

We recommend racks at least 750 mm (29.5 in) wide, 1,200 mm (47.2 in) deep, and 48U[18] high. Consider racks with maximum mounting depths greater than 1,000 mm (40 in). Although these racks will not line up with 600 mm wide raised floor tiles like standard 600 mm racks do, this is no longer a relevant constraint. This is because air-cooled AI servers require high airflow rates and raised floors are not typically used for air distribution but rather for piping and cabling.

A high-density AI rack can weigh over 1,800 kg (4,000 lb) and places a significant load on IT racks and raised floors. Racks not rated for these weights may deform around the frame, leveling feet, and/or casters. IT rack weight-bearing capacities are specified as static and dynamic. Static refers to the weight a rack can support while stationary. Dynamic refers to the weight a rack can support while moving.

We recommend specifying AI compute racks with a static weight capacity greater than 2,270 kg (5,000 lb) and a dynamic weight capacity greater than 1,600 kg (3,500 lb). Validate that concrete slab floors are rated for rack weights over 3,000 kg (6,600 lb).[19] Structured cabling trays will also need to be designed to accommodate the changing network cabling layouts and densities.

Rack capacities should be validated by an independent third party.[20] Even if your current AI deployment is small and doesn't yet require larger capacities, racks tend to have a longer service life than IT equipment. It's likely that the next generation of your AI deployment will require some or all of these rack recommendations.

In some cases, IT racks are preconfigured offsite and then transported to the data center. These racks must sustain the dynamic forces generated during transportation and the associated packaging must also protect the racks and the valuable IT gear they support. Data center floors, and raised floors in particular, should be

[18] 1U is equal to 44.45 mm (1.75 in). For example, 48U = 2.13 m (84 in) of interior vertical space.

[19] Rough estimate by doubling the current weight of an NVIDIA NVL72 rack.

[20] We recommend Underwriters Laboratory (UL) and International Safe Transit Association (ISTA). For more information see White Paper 201, "*How to Choose an IT Rack*".

assessed to be certain they can support the weight of an AI cluster. This is especially important to raised floor dynamic capacity when moving heavy racks around the data center.

Finally, increasing rack densities and higher operating temperatures (for improved cooling system efficiency) are subjecting power cords to high temperatures. This may lead to safety hazards like melting cords. Validate that IT equipment in AI racks use high-temperature power cords. IEC 60320 is the recognized international standard used by most of the globe for the connection of power supply cords. Table 3 compares the standard C19/C20 connectors to the high temperature C21/C22 connectors. Placing temperature sensors in the back of the rack, with data center infrastructure management (DCIM) monitoring, is recommended to validate the operating conditions are as expected.

Table 3

Comparison of IEC 60320 standard and high temperature connectors for 250 V and 16/20 A

	Female	Male	Limit	Notes
Standard	 C19	 C20	65˚C	**C20 is commonly used as a jumper cable, providing power from a rack PDU to high powered IT devices.**
High temperature	 C21	 C22	155˚C	**C21 mates with either C22 or C20 connectors and used when temperatures exceed the C19 rating.**

Supply chain considerations

The demand for data centers to support AI workloads combined with racks densities over 100 kW, requires that you work with your design partners early in the design process and specify the physical infrastructure in parallel with the IT stack. The high demand for physical infrastructure is driving longer delivery lead times: it's wise to procure infrastructure as soon as possible. With AI applications, this becomes even more critical to a data center's project schedule. The following considerations help reduce the risk of project delays related to your ecosystem of partners.

- Leverage standardization and prefabrication – Standardization is a challenge when it comes to newer technology and architectures like direct-to-chip liquid cooling. However, even with newer technology, leverage partners using modularity in their manufacturing. For example, you can assemble a "custom" CDU per specifications and still use a series of standard sub-assemblies. This improves lead times while reducing the risk of unique failures from bespoke "one-off" designs.
  Standard prefabricated modules (e.g., power skids) are another strategy for leveraging supply chain for faster delivery. The more standardized prefab modules are, the shorter the lead time compared to ordering the parts and assembling yourself. For more information, see White Paper 163, *Benefits and Drawbacks of Prefabricated Modules for Data Centers*.
- Use supplier partners with a global footprint – This is more important for data center projects with multiple international locations. By choosing partners with manufacturing and distribution centers located throughout the world you increase your chances of on-time delivery. Delivering a product out of a local warehouse avoids delays due to port congestion, customs, unpaid duties, container shortages, and cargo damage.
- Use fewer partners – Leveraging a partner that supplies a large portion of required components brings two key benefits. 1. Buying more components from the same supplier can lead to lower pricing. 2. When a manufacturer has a

sufficiently diverse portfolio, it often has experience dealing with complex supply chains and managing the availability of stocked items. This means that you reduce your risk of project delays.

- Seek partners recognized for their supply chain – Use industry rankings and reports to assess partners (e.g., Gartner Supply Chain Top 25). Oftentimes these rankings weigh the sustainability actions a supplier takes, in addition to supply chain efficiency. Think of this as a pointer to good candidates, not as an absolute confirmation for all supply chain best practices.
- Engage supply chain partners early in the design phase – Meeting with your supply chain specialists during initial design helps optimize not just product specifications, but also ensures the design considers material availability and lead times. This collaboration can prevent costly delays and rework by identifying components that are readily accessible when you need them. This may also uncover opportunities for pre-fabricating certain systems versus on-site construction.

Software tools and services

Physical infrastructure software tools and services support the design and operation of the data center. Software tools include DCIM, EPMS, BMS, and digital electrical design tools.[21] Some vendors offer pre-validated reference architectures/designs to accelerate data center projects. Some of these reference designs, like Schneider Electric's, are co-developed with NVIDIA.

Power and cooling configuration tools also save time and increase accuracy for build-to-order systems. Related services available in the market today that take advantage of these software tools include, for example, electrical design consults, system studies services, SLD digitalization, condition-based maintenance, and remote monitoring.

Having clusters of high-power density and liquid-cooled IT alongside traditional air-cooled IT means that certain software functions become more critical. Note that the critical software functions described below may be outsourced through a service contract. Even though some AI training workloads may not require high availability, poor design and poor monitoring can lead to downtime risks for adjacent non-AI racks. For example, colocation tenants may have business-critical racks adjacent to high-density AI racks that could be adversely affected by their hot exhaust air.

The following two considerations highlight important management software functions that become particularly relevant in the context of high-density AI training workloads:

- AI clusters are pushing power use and density to new levels, creating design uncertainty
- With less room for error, the risk of operational issues increases in an already fast-changing environment

AI clusters are pushing power use and density to new levels, creating design uncertainty

Before retrofitting an existing site to accommodate new AI clusters, a feasibility study should confirm there is enough power and cooling capacity. The study should also cover the infrastructure needed to distribute that capacity to the new loads. When rack power densities are below 10 kW and there is excess bulk power and

[21] DCIM (data center infrastructure management), EPMS (energy and power monitoring system), BMS (building management system)

cooling capacity, adding standard IT might be straightforward, without as much scrutiny and verification. Point-in-time power and cooling measurements might be used along with common power distribution components and existing cooling units that you're familiar with.

This more manual, "eye-ball" retrofit design approach will not suffice for large high density AI training clusters. An AI cluster drawing hundreds of kilowatts presents more serious consequences if you make a design mistake (i.e., not knowing actual peak to average power draws, being unsure of what loads are on which circuits, etc.). You can't afford to have unknowns and uncertainties with the design. Also, because AI cluster designs are so unique (e.g., non-standard high amperage rPDUs/busway, use of liquid cooling, etc.) there is greater uncertainty about how the cluster will perform on startup.

We recommend using **EPMS** and **DCIM** to provide an accurate view of the current power capacity and its trends, both at the bulk power level and the distribution level within the IT space. These tools will show what the actual peak power draw is over a long period of time. This understanding will prevent inadvertently tripping a breaker.

The capacity assessment can help determine the capability of hosting AI loads. This assumes that the necessary power meters are in place. Next, prior to any changes, we recommend performing safety and technical studies including:

- capacity analysis
- protection coordination
- arc flash study
- short-circuit & device evaluation[22].

Using **electrical design (a.k.a., power system engineering) software tools** can simplify the data collection and calculations.

After the assessment, changes to the electric network will likely be required to add the AI clusters. In this case, electrical design software tools help you:

- select optimal electrical equipment
- prevent electrical faults
- develop effective methods of procedure
- implement proper safety protocols when working on and servicing the electrical network in the IT space

Existing data centers with digitalized single-line diagrams (iSLDs)[23] will be able to simplify the assessment process described above. Accurate, intelligent, iSLDs reduce the time and expertise needed to collect data and perform the calculations. An iSLD is a more advanced single-line diagram stored and managed in specialized software that includes advanced functionality, awareness of the devices' characteristics and their operating behavior. It creates a digital twin of the physical electrical network. In essence, this one software platform can be used to design the electrical network, create and maintain the SLD, and perform all technical studies and safety assessments. See White Paper 281, *How Modern DCIM Addresses CIO Management Challenges within Distributed, Hybrid IT Environments*.

[22] i.e., evaluating capacity, kA ratings, and other specs for suitability for the given design
[23] Some vendors offer iSLD creation and maintenance as a service.

For those designing large AI training clusters or "AI factories," emerging electrical design tools offer a digital twin platform to design and simulate AI cluster power requirements.[24] Once operational, this virtual replica can then be combined with real-time power system data and analytics to provide performance tracking, energy optimization, "what-if" scenario planning, and so on.

With less room for error, the risk of operational issues increases in an already fast-changing environment.

Compared to other types of facilities, data centers are dynamic environments where frequent moves, adds, and changes of IT equipment take place. With the addition of an AI cluster, capacity safety margins shrink. This increases the risk of tripping a breaker, creating a hot spot, or stranding resources. The underlying reasons for the increased risk are the high rack densities and power variation of AI clusters. Given a tight margin for error, operators need increased situational awareness to prevent downtime and manage operational risk.

We recommend creating a digital twin of the entire IT space (including the equipment and VMs in the racks) which minimizes or prevents the challenges we've indicated. DCIM planning and modeling functions allow you to design effective IT space floor layouts using a rules-based tool. However, this requires that you update the layout after any changes are made. By digitally adding or moving IT loads, you can validate that there is sufficient power, cooling, and floor weight capacities to support them. DCIM creates a digital twin of the IT space and documents all equipment dependencies on resources. This avoids stranding resources and minimizes human error that might lead to downtime.

EPMS and DCIM together allow you to monitor power capacities across all PDUs, UPSs, rPDUs, etc. They create an early warning system to avoid exceeding power thresholds, averting downtime. DCIM software will advise the best place to locate new equipment based on power, cooling, redundancy level requirements, as well as available U-space, network port, and weight capacity. This applies more to non-AI equipment and to AI inference servers.

Unlike inference loads, AI training loads require a pre-designed configuration that seldom changes.

Many DCIM planning and modeling software tools include a computational fluid dynamics (CFD) tool to deliver adequate air flow given the physical layout of the equipment and heat load. DCIM can be used to optimize cooling capacity by releasing stranded cooling capacity through the optimal placement and configuration of infrastructure and loads. The CFD tool applies more to AI inference loads since servers are added to meet user demand (i.e., queries). In contrast, AI training layouts don't experience as many moves, adds, and changes.

Sometimes, the AI training or inference cluster is isolated on its own power segment and cooling architecture. In these cases, non-AI loads are less susceptible to the effects of the AI cluster. However, in both cases, establishing a digital twin of these spaces is beneficial.

Very large AI factories may opt for industrial-scale operations management software platforms like AVEVA's Unified Operations Center, in addition to DCIM. Whereas DCIM is focused on the IT space, these industrial, "system of system" platforms create a digital twin of the entire facility (i.e., all building systems including

[24] ETAP and Schneider Electric Unveil World's First Digital Twin to Simulate AI Factory Power Requirements from Grid to Chip Level Using NVIDIA Omniverse

mechanical and electrical) from the grid connection to the rack PDUs in the IT space and from the chiller to the heat rejection systems. This higher-level software provides the benefit of managing and monitoring *all* systems across a facility or even a data center campus.

Services

Power and cooling infrastructure vendors offer a variety of services that span across the data center and its lifecycle. The software functions above can be outsourced through service contracts with vendors. Employing service contracts versus performing these software functions yourself:

- reduces workload on operations teams, enabling them to focus on more strategic initiatives.
- addresses labor shortages in experienced data center facilities personnel
- allows you to benefit from the expertise of the vendor and their knowledge of their own infrastructure systems, as well as their software design and operations management tools.
- speeds time to market with less risk of re-design and human error.
- improves reliability and operational efficiency beyond doing it yourself.

For more information see Executive Report 6, *Transforming Data Center Services: AI-Driven Condition-Based Maintenance*.

Future outlook of physical infrastructure to support AI

Technologies, including the physical infrastructure, are evolving fast. In this section, we provide an overview of some *future* technologies and design approaches that will change how AI workloads are supported.

- Medium voltage transformers in the technical / IT space – Distributing power at medium voltage (e.g., 13 kV) reduces copper, requires fewer conductors, and reduces installation time for the upstream electrical system. This distribution architecture also eliminates the traditional 13 kV to 480/277 V transformers and switchgear upstream of the IT distribution.
- Solid state transformers – These are an emerging form of power electronics converters. They use semiconductor components to change the primary voltage to a secondary voltage. They use a medium frequency transformer (MFT) to galvanically isolate the primary and secondary sides. While traditional transformers are heavy and work with only alternating current (AC), solid state transformers are small, light and convert between AC and DC voltage.
- Solid state circuit breakers – A circuit breaker acts like a switch that opens a circuit in the case of a fault or overload. A standard breaker uses a mechanical switch that stops the flow of current. Solid state breakers use semiconductors to do this, like a computer chip's transistors turn on and off. Compared to mechanical switches, semiconductor switches open extremely fast. The faster it opens, the less fault current gets through, reducing arc flash energy at high-density AI racks. To be considered a circuit breaker, solid state breakers must also use a mechanical switch in series with the semiconductors to provide galvanic isolation.
  There are two types: solid state and hybrid, distinguished by the current flow under normal operation. The current flow in the solid state breaker is through the semiconductors. The current flow in the hybrid breaker is through the mechanical switch which generates lower losses compared to the solid state.

- **Sustainable dielectric fluids** – These may replace the water glycol mix as today's choice for direct-to-chip cooling if they can increase the heat transfer efficiency, allow for higher chip TDPs, and address PFAS and GWP concerns.[25]
- **Increased interaction/optimization with the grid –** Scheduling workloads based on utility and micro grid conditions may help with balancing the grid and saving on electricity. Migrating loads to different redundancy zones or placing a UPS on battery operation are examples of workload management.
- **Higher-voltage racks** – As AI rack densities trend from 100 kW today to 1 MW and higher, the distribution voltage must also increase to accommodate the required copper conductors. Distribution architectures are being proposed with increased voltages and side cabinets for power and cooling.
- **Sustainable design –** Environmental sustainability is important and will become more critical, especially as more regulations are enacted. Water-free or water-limited cooling plants will become the norm, especially in water-scarce regions. Cooling designs with dry coolers are an example of this.

Next steps

The rapid growth and application of AI is changing the design and operation of data centers. **Inference** workloads, although collectively expected to consume much more power than training clusters, operate at a wide range of rack densities. Recent trends indicate that some inference use cases operate at high power densities that result in some of the challenges discussed in this paper.

AI training workloads, on the other hand, consistently operate at very high densities, reaching over 100 kW per rack. Networking demands and cost drive these training racks to be clustered together. These clusters of high power density are fundamentally what challenges the physical infrastructure. We recommend you take the following next steps:

POWER: Prioritize early engagement with design partners to provision electricity and electrical system design.

- For existing facilities, conduct a thorough feasibility study using EPMS and DCIM to accurately assess current power capacity, identify peak draws, and understand circuit loads.
- Perform essential safety and technical studies like capacity analysis and arc flash studies, utilizing electrical design software for precise calculations and optimal equipment selection.
- Consider digitalizing single-line diagrams (iSLDs) to streamline assessments.
- Leverage emerging digital twin platforms for AI cluster power design and simulation, ensuring a robust and predictable power supply.

COOLING: As AI's escalating thermal demands necessitate liquid cooling, begin by conducting a comprehensive design assessment of the proposed liquid-cooled loads and your facility's existing conditions. Seek expert review to evaluate designs and avoid unforeseen building constraints, especially given the disruptive nature of retrofitting. Note the current lack of liquid cooling standards, which drives integration complexity and cost. Actively leverage your partner ecosystem's expertise and their thermal modeling and fluid dynamics tools to navigate these challenges and ensure effective, compatible cooling solutions.

[25] PFAS (per- and polyfluoroalkyl substances), GWP (global warming potential)

RACKS: To accommodate AI's extreme rack densities and weights, specify racks wider, deeper, and taller racks with higher weight capacities. Validate concrete slab floor loading capacities for loads over 3000 kg and ensure structured cabling trays can handle changing network layouts. Independently validate these capacities. Proactively procure these robust racks early due to increasing lead times.

SOFTWARE TOOLS AND SERVICES: Leverage your ecosystem of partners and their tools and servers. These offer assistance in mitigating the design and operational risks inherent in AI deployments.

- When designing and managing AI clusters, use software tools such as, DCIM, EPMS, BMS, digital electrical design tools, and reference architectures. They decrease the risk of unexpected behavior with complex electrical networks.
- Create a digital twin of the data center to identify constrained power and cooling resources and inform your layout decisions.
- If you lack the bandwidth and/or expertise, leverage services available from the software vendors, to do this work for you.

About the authors

Victor Avelar is a seasoned expert in data center energy efficiency and design, serving as the Chief Research Analyst in Schneider Electric's Data Center Research & Strategy group. With over 25 years of experience, Victor leads cutting-edge research and best practice development for sustainability, risk management, and next-generation data center technologies. He's a trusted advisor to clients globally, providing actionable insights on enhancing infrastructure performance through innovative solutions such as liquid cooling and energy modeling. Known for his clear, practical guidance, Victor helps organizations tackle the evolving challenges of sustainable and efficient data center operations. He is central to the development of technology adoption forecasts for data centers. He also leads the peer review process for all DCRS content. Victor holds a bachelor's degree in mechanical engineering from Rensselaer Polytechnic Institute and an MBA from Babson College. He is a member of AFCOM and a sought-after speaker on AI infrastructure.

Patrick Donovan is a Senior Research Analyst for the Data Center Research & Strategy group at Schneider Electric. He has over 28 years of experience developing and supporting critical power and cooling systems for Schneider Electric's Secure Power Business unit including several award-winning power protection, efficiency, and availability solutions. An author of numerous white papers, industry articles, and technology assessments, Patrick's research on data center physical infrastructure technologies and markets offers guidance and advice on best practices for planning, designing, and operation of data center facilities.

Wendy Torell is a Senior Research Analyst in Schneider Electric's Data Center Research & Strategy group bringing 30 years of data center experience. Her focus is analyzing and measuring the value of emerging technologies and trends: providing practical, best practice guidance in data center design and operation. Beyond traditional thought leadership, she championed and leads development of interactive, web-based TradeOff Tools. These calculators help clients quantify business decisions, while optimizing their availability, sustainability, and cost of their data center environments. Her deep background in availability science approaches and design practices helps clients meet their current and future data center performance objectives. She brings a wealth of experience across Schneider Electric's broad portfolio and with the market at large. She holds a BS in Mechanical Engineering from Union College and an MBA from University of Rhode Island. Wendy is an ASQ Certified Reliability Engineer.

Dr. Maria A. Torres Arango is a Research Analyst in Schneider Electric's Data Center Research & Strategy group. She investigates technology, materials and infrastructure to guide data center strategies. Her current work focuses on energy storage and liquid cooling technologies. She also examines market forces driving technology advancement and participates in developing tools that highlight tradeoffs or selection in critical data center decisions. Maria's former expertise involves materials design and optimization; and fundamental studies on materials synthesis processes using X-ray characterization at the National Synchrotron Light Source II, Brookhaven National Laboratory. A lifelong learner, Maria holds a BS in aeronautical engineering from Universidad Pontificia Bolivariana, Colombia; and a MSc in aerospace engineering and a PhD in materials science and engineering from West Virginia University.

RATE THIS PAPER ★★★★★

Resources

The Looming Power Crunch: Solutions for Data Center Expansion in an Energy-Constrained World
Executive Report 1

Generative AI Inferencing Ramp-up: A CIOs Guide to Physical Infrastructure Consideration
Executive Report 5

Transforming Data Center Services: AI-Driven Condition-Based Maintenance
Executive Report 6

Retrofitting Existing Power Systems for AI Clusters
White Paper 126

Navigating Liquid Cooling Architectures for Data Centers with AI Workloads
White Paper 133

Direct Liquid Cooling System Challenges in Data Centers
White Paper 210

Bending the Energy Curve: Decoupling Digitalization Trends from Data Center Energy Growth
White Paper 212

Navigating Energy Constraints: Selecting Alternative Power Strategies for Data Center Expansion
White Paper 275

How Modern DCIM Addresses CIO Management Challenges within Distributed, Hybrid IT Environments
White Paper 281

Browse all white papers
whitepapers.apc.com

Browse all TradeOff Tools™
tools.apc.com

Note: Internet links can become obsolete over time. The referenced links were available at the time this paper was written but may no longer be available now.

Contact us

For feedback and comments about the content of this white paper:

Schneider Electric Data Center Research & Strategy
dcsc@se.com

If you are a customer and have questions specific to your data center project:

Contact your Schneider Electric representative at
www.apc.com/support/contact/index.cfm