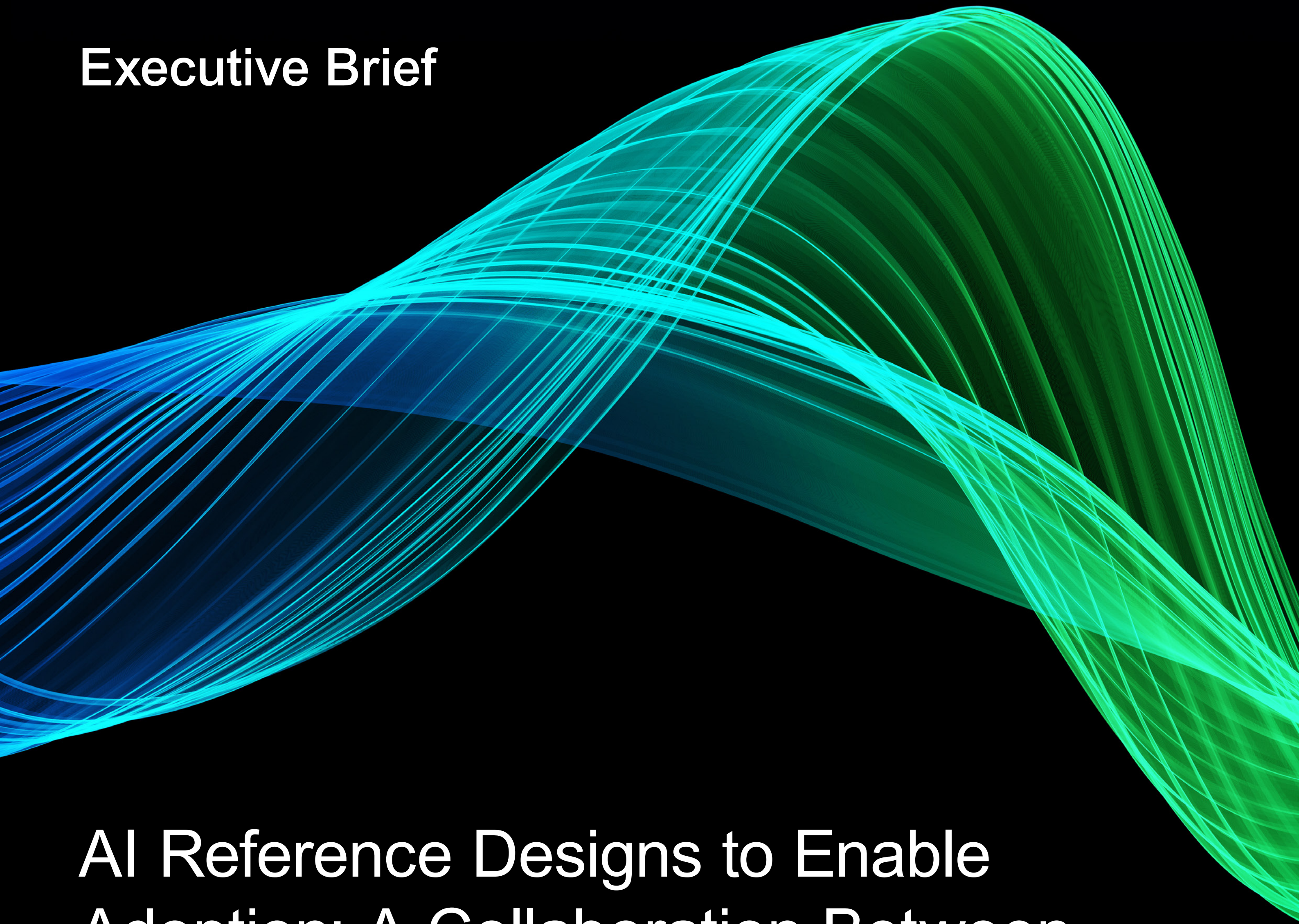
Executive Brief



AI Reference Designs to Enable Adoption: A Collaboration Between Schneider Electric and NVIDIA

Steven Carlini

Vice President,
Innovation and Data Center
Schneider Electric

Wendy Torell

Sr. Research Analyst,
Energy Management Research Center
Schneider Electric

June 27, 2024

Life Is On Schneider Electric

About the Authors

Steven Carlini is the Vice President of Innovation and Data Center for Schneider Electric. Steven is responsible for developing integrated solutions and communicating the value proposition for Schneider Electric's data center segment including enterprise and cloud data centers. A frequent speaker at industry conferences and forums, Steven is an expert on the foundation layer of data centers which include power & power distribution, cooling & technical cooling, rack systems, physical security, DCIM management solutions that improve availability and maximize performance. Steven has been responsible for guiding the direction of many industry changing products and solutions that solve real customer problems or give businesses competitive advantages. Steven holds a BS in Electrical Engineering from the University of Oklahoma, and an MBA in International Business from the CT Bauer School at the University of Houston.

LinkedIn: https://www.linkedin.com/in/steven-carlini
Forbes: https://councils.forbes.com/profile/Steven-Carlini-VP-Innovation-Data-Centers-Schneider-Electric

Steven Carlini
Vice President,
Innovation and Data Center
Schneider Electric

Wendy Torell is a Senior Research Analyst in Schneider Electric's Energy Management Research Center 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.

LinkedIn: https://www.linkedin.com/in/wendy-torell-008a4a4

Wendy Torell
Sr. Research Analyst,
Energy Management Research Center
Schneider Electric

Key takeaways

AI's impact is rapidly spreading across industries, with new applications constantly emerging. This will fundamentally change how we work, play, and learn. But before AI can truly become a powerful tool for society, foundational models need to be trained. These models require fast specialized chips called accelerators, with NVIDIA GPU-based compute as the dominant accelerator type. NVIDIA itself is evolving from a GPU component supplier to a supercomputer and software stack vendor. NVIDIA now offers server boards, servers, and SuperPODs (AI clusters) to fast-track AI deployment with lower risk.

Traditional data center power, cooling, and racks aren't sufficient for GPU-based servers arranged in high-density AI clusters. Schneider Electric collaborated with NVIDIA to develop reference designs of the physical infrastructure for both retrofit and purpose-built scenarios. These designs provide data center operators the guidance and technical detail to streamline and accelerate deployment of these high-density AI clusters.

For data center operators looking to add AI clusters into an existing facility (retrofit) we have developed three designs. Each has specific benefits and demonstrates the addition of a ~1 MW AI cluster to a data center with ~1 MW of traditional workloads, to support a total of ~2 MW of IT load:

Retrofit 1 – For the data center operator looking to add a ~1 MW AI cluster(s) without implementing liquid-cooled servers. This design uses air-cooled servers and can accommodate 40 kW per IT rack. However, the cluster footprint is larger than the liquid-cooled versions.

Retrofit 2 – For data center operators looking to add ~1 MW liquid-cooled AI clusters and remove heat using their existing (air-cooled) cooling system. This design can accommodate 73 kW per IT rack; it balances floor space savings with density.

Retrofit 3 – For data center operators looking to add ~1 MW liquid-cooled AI clusters and remove heat by tapping into pipes of an existing chilled water-cooling system. This design can accommodate 73 kW per IT rack while minimizing floor space.

In addition to the 3 retrofit scenarios, we have also developed a new-build design that can be scaled for data center operators looking to build an IT space specifically for AI clusters.

New data center – For data center operators looking to deploy one or many 1.8 MW high-density AI clusters with a maximum rack density of 73 kW, using purpose-built heat rejection optimized for efficiency.

Exponential growth carries challenges

AI is experiencing exponential growth. Advancements in machine learning and data analysis are fueling more and more sophisticated algorithms, capable of complex tasks once thought to be the exclusive domain of humans. We anticipate that by 2028, AI power consumption will be 316% higher than it was in 2023.[1] This growth is having a ripple effect across all industries.

In healthcare, AI is assisting doctors in diagnosing diseases, analyzing medical images, and even developing personalized treatment plans. Manufacturing is witnessing the rise of intelligent robots for precision assembly and predictive

1 Schneider Electric White Paper 110, The AI Disruption: Challenges and Guidance for Data Center Design, Table 1.

maintenance. AI is personalizing education by tailoring learning programs to individual student needs. In the entertainment industry, it's powering realistic special effects and even creating new forms of interactive content. These are just a few examples where the potential applications for AI's value seems limitless.

Foundational AI models are needed for widespread adoption. Foundational AI models, also known as base models, are crucial to the AI revolution. These powerful models are "AI neural networks trained on massive unlabeled datasets to handle a wide variety of jobs."[2] They enable specific spinoff models fine-tuned to accomplish a narrower, more specialized set of tasks, from language translation to medical diagnosis. The GPT1 – the Generative Pretrained Transformer – version 1 came out in 2018. Since then, GPT-n series of large language models (LLMs) provide the most recognizable example of a foundational model. Their versatility has made them a de facto standard for widespread AI adoption. Fine-tuning a foundational model accelerates the development time and reduces the cost for a specific application.

However, training these massive models is no easy feat. It requires immense computing power, pushing the boundaries of GPUs. NVIDIA's GPUs are currently the frontrunners, offering unparalleled speed and adoption within the AI community. The deployment of foundational AI training models utilizes a mix of colocation companies and large, cloud-based infrastructure providers. Colocation provides a flexible option for companies who need a single cluster or a few. Large-scale operations are typically being deployed by major cloud providers.

Training complex models poses a significant challenge to traditional data center infrastructure. In the Energy Management Research Center's white paper, The AI Disruption: Challenges and Guidance for Data Center Design, we highlight the limitations of traditional physical infrastructure. We then offer insights on adapting data centers for AI clusters. Traditional designs aren't equipped to handle the sheer power capacity and rack density demands of AI training clusters. They require innovative solutions for power delivery, cooling strategies like liquid cooling, and fortified IT cabinets.

Collaboration for innovation, leveraging strengths

Schneider Electric and NVIDIA are collaborating. We address the data center challenges posed by the growing complexity of AI by assembling experts from both organizations. We pulled together the IT and physical infrastructure knowledge necessary to simplify and optimize data center infrastructure for AI workloads, enabling and accelerating AI adoption across industries. And we kept energy efficiency as a foundational principle in our designs, a major component of environmental sustainability (scope 2 and scope 3).[3]

Reference designs provide a blueprint for a reliable, validated AI deployment. "A reference design is a system blueprint, list of attributes including system level performance specifications, and (ideally) includes a detailed list of materials or components that comprise the system. While a reference design can be directly implemented, in most cases it acts as a baseline design that is adapted to meet specific user preferences or constraints."[4] Reference designs aren't new. For many years, they've been applied to help data center owners shorten their planning cycle and reduce downtime risks once operational. But these designs become **crucial** when the physical infrastructure (i.e., power and cooling) demands different and more complex solutions than owners and operators have deployed in the past.

Through this collaboration, we've produced an industry-first comprehensive, validated reference design for high-density AI clusters like NVIDIA's DGX SuperPODs, with rack densities up to 73 kW. These designs provide a necessary head start to successful,

2 NVIDIA blog, What Are Foundation Models?, accessed on May 13, 2024
3 Schneider Electric White Paper 99, Quantifying Data Center Scope 3 GHG Emissions to Prioritize Reduction Efforts
4 Schneider Electric White Paper 147, Data Center Projects: Advantages of Using a Reference Design

reliable deployment. They're crucial for data center designers who lack experience architecting solutions at extreme densities, and/or have not implemented liquid cooling (which is evolving as the industry matures).

Alternative designs for high-density AI clusters

The benefits of the resulting AI reference designs are summarized below:

- **Simplify implementation feasibility analysis.** The extreme rack densities of AI workloads make many traditional designs impractical. In addition, when retrofitting a high-density AI cluster into an existing data center, certain constraints impact design choices. A feasibility study is simplified with these designs as a starting point.
- **Reduce planning cycle time.** The designs include an equipment selection list with electrical and mechanical systems available on the market today. They also provide documented physical layouts. This information cuts significant time out of the project.
- **Leverage our learning on key design considerations and obstacles.** The designs provide best practices to overcome obstacles arising from the complexity of high density. Extreme densities and capacity in small footprints bring challenges in electrical attributes such as short circuit current, breaker coordination/selectivity, and arc flash. We've covered how to handle these. New approaches to cooling are addressed: handling set points, airflow challenges, and key concerns for liquid cooling, CDUs, etc.
- **Increase confidence that your AI cluster can be deployed reliably.** When new technologies are deployed, there is a greater risk that something will go wrong. We have iterated these designs using state-of-the-art electrical design and mechanical design software tools to ensure they are validated to operate reliably.
- **Leverage expertise in efficient design.** Reliably operating high-power, high-density AI loads is first and foremost, but the designs also follow best practices for energy efficiency to align with industry sustainability goals.

Four design scenarios to support NVIDIA's DGX SuperPODs were engineered and documented, including three retrofit scenarios and one purpose built / greenfield scenario. In the retrofit scenarios, each design supports roughly a 1 MW AI cluster; and in the purpose-built room, it supports a roughly 2 MW cluster. Below is a high-level description of the AI clusters in each scenario as well as the physical infrastructure to support them. **Table 1** then provides further details and side-by-side comparison of the scenarios.

Retrofit scenario 1, high-density air-cooled AI cluster. 1,050 kW AI cluster with maximum rack density of 40 kW, using traditional air cooling, with wider, contained hot aisles, and 630A busway powers each high density (40 kW) rack via 63A busway feeds.

Retrofit scenario 2, high-density liquid-cooled AI cluster with liquid-to-air CDUs. 904 kW AI cluster with maximum rack density of 73 kW, using liquid-to-air coolant distribution units (CDUs) for heat rejection. This is ideal for scenarios when you cannot connect to facility water systems. The busway feeds multiple 63A circuits to OCP racks each with 3 power shelves.

Retrofit scenario 3: high-density liquid-cooled AI cluster with liquid-to-liquid CDUs. 904 kW AI cluster with maximum rack density of 73 kW, using liquid-to-liquid CDUs for heat rejection. This is ideal for scenarios where you can tap into facility water systems.

Purpose-built data center optimized for a liquid-cooled cluster. 1,808 kW AI cluster with maximum rack density of 73 kW, using liquid-to-liquid CDUs for heat rejection, optimized with high temperature chillers for improved efficiency.

Table 1 – Summary of AI training cluster reference designs

Attribute	Retrofit design 1	Retrofit design 2	Retrofit design 3	Purpose-built design
AI cluster IT capacity	1,050 kW	904 kW	904 kW	1,808 kW
AI server racks	24 x 40 kW	8 x 73 kW	8 x 73 kW	16 x 73 kW
AI networking racks	6 x 15 kW	8 x 40 kW	8 x 40 kW	16 x 40 kW
Existing IT capacity in room	960 kW	960 kW	960 kW	0 kW
Traditional IT racks	80 x 12 kW	80 x 12 kW	80 x 12 kW	0 kW
Total IT capacity	2,010 kW	1,864 kW	1,864 kW	1,808 kW
AI server rack type	Standard on hard floor	OCP with 3 power shelves on hard floor	OCP with 3 power shelves on hard floor	OCP with 3 power shelves on hard floor
AI server GPU cooling method	Air-cooled	Direct-to-chip	Direct-to-chip	Direct-to-chip
AI cluster heat rejection design	Wide hot aisles with ducted containment, chilled water fan walls, 30C chiller return	Wide hot aisles with ducted containment, liquid-to-air CDUs, chilled water fan walls, 30C chiller return	Wide hot aisles with ducted containment, liquid-to-liquid CDUs, chilled water fan walls, 30C chiller return	Wide hot aisles with ducted containment, liquid-to-liquid CDUs, 2 chiller loops, chilled water fan walls, 40C high temp chiller return
AI server racks power design (per row)	630A busway to redundant tap-offs with redundant 63A 400V rPDUs	800A busway to redundant tap-offs with 3+3 63A 400V power feeds	800A busway to redundant tap-offs with 3+3 63A 400V power feeds	Split 800A busway to redundant tap-offs with 3+3 63A 400V power feeds

Partners and engineers in early planning of AI clusters can obtain detailed documentation of the reference designs, including specifications, layouts and system drawings, equipment lists, and computational fluid dynamics (CFD) analysis.

While the detailed designs were based on specific NVIDIA AI clusters, they may be applied to other high-density AI clusters. In doing so, they should be tailored based on the IT design temperature set points, server configuration, etc.

Impact of future generations of AI GPUs and high-density AI clusters

NVIDIA is dedicated to staying ahead of the market. They're continuing to advance their GPUs and AI "server boards" (GPUs and CPUs on a circuit board connected through NVLink), and providing blueprints to package them into supercomputers (SuperPODs). The most recent evolution is from the DGX GH200, the "Hopper," to the DGX GB200, the "Blackwell" (named after famous scientist/mathematician). The Blackwell platform "enables organizations everywhere to build and run real-time generative AI on trillion-parameter large language models at up to 25x less cost and energy consumption than its predecessor."[5]

Servers based on NVIDIA's design include the GB200 NVL36 with 36 GB200 Superchips (18 Grace CPUs and 36 enhanced B200 GPUs, 73 kW per rack[6]) and the NVIDIA GB200 NVL72 with 72 GB200 Superchips (36 CPUs and 72 GPUs, 132 kW per rack[6]) which, at the time of this writing, have multi-million-dollar price points. Many organizations are expected to adopt Blackwell, including internet giants such as, Amazon Web Services, Dell Technologies, Google, Meta, Microsoft, OpenAI, Oracle, Tesla and xAI.

NVIDIA's goal with DGX is to offer their IT reference architectures broadly to all of its

5 NVIDIA, NVIDIA Blackwell platform arrives to power a new era of computing, April 4, 2024
6 NVIDIA, 8th Annual DGX User Group Meeting, March 19, 2024, Slide 32

partners. The collaboration between NVIDIA and Schneider Electric on these reference designs presents the perfect companion piece because it saves them significant time and money in R&D. As NVIDIA develops the next GPU's server boards and clusters, their ambition is to create designs that operate within the power and cooling constraints of previous generations. However, as new designs launch, Schneider will develop accompanying solutions and reference designs to support them, with the possibility of seeing architectural changes such as higher voltage DC (e.g., 400 VDC) powering a single rack.

Conclusion

Schneider Electric and NVIDIA's collaborative reference designs add significant value for data center operators looking to add AI clusters to an existing facility or build a new one. By producing the first publicly available, comprehensive blueprints for AI data centers, this partnership goes beyond optimizing infrastructure – it empowers companies to deploy AI with greater efficiency and ease.

To get the most out of the reference designs, we recommend you:

1. Download/share the design summary with your data center design team, and then request the engineering package which includes technical documentation and details from referencedesigns@se.com.

2. Get specific about infrastructure and material specification in your design or redesign. Planning now increases confidence in later design choices.

3. Assess the pre-selected equipment and documented layouts to save significant time in the planning cycle.

4. Ask your design team to implement and or document the best practices as applied to your design. This translates to reliable deployment and reduced risk of unforeseen issues.

Looking forward, this collaboration positions data center design teams and the companies they work with to deliver reliable, sustainable AI infrastructure, ensuring that future advancements can build upon a strong foundation.