

CSTE ENTERPRISES

The New Power Path

How AI Is Rewriting the Infrastructure Between the Grid and the Compute Rack

AIDC INFRASTRUCTURE WHITEPAPER | WP01

Executive market intelligence on grid-to-compute architecture, power constraints, emerging 800 VDC pathways, modularization, value migration and lifecycle control points.

Charles Trammell, MBA | CSTE Enterprises

September 2026

CSTE Enterprises | csteenterprises.com

Executive Summary

- Power has become a front-end constraint on AI infrastructure growth, not merely a facility utility input.
- The industry is moving from a single mature AC distribution paradigm toward a portfolio of conventional AC, hybrid high-voltage-DC, and potentially more DC-native architectures.
- 800 VDC matters strategically because coordinated hyperscale and supplier investment proves the power-delivery path itself is being redesigned for extreme rack density.
- Grid access, on-site generation, BESS, modular power blocks and factory integration are increasingly part of one time-to-power problem.
- Power and thermal architecture are converging; reference designs increasingly validate them together rather than as independent systems.
- Value can migrate toward interfaces that control interconnection, conversion, integration, operating visibility, commissioning and lifecycle service.

CSTE GOVERNING THESIS

Capital and margin tend to move toward the bottleneck. In AI data centers, the path from available power to usable compute is becoming one of those bottlenecks.

Scope and boundary

This whitepaper is strategic market intelligence. It synthesizes public technical sources, regulatory and reliability evidence, supplier roadmaps, CSTE AIDC research and repeated market signals. It is not electrical design guidance and does not represent firsthand data-center electrical engineering, construction, procurement or operator experience.

Executive Thesis

AI is not simply increasing the amount of electricity consumed by data centers. It is changing the architecture through which electricity becomes usable compute power.

The emerging shift is visible across the full path: utility and grid access, substations and medium-voltage distribution, backup and stored energy, power conversion, facility distribution, row and rack power, controls, and the interface to increasingly dense accelerated-compute systems. As these layers become more tightly coupled, decisions that were once treated as separate electrical, IT, thermal, and facilities choices increasingly behave like one infrastructure system.

CSTE's governing thesis is that value will not migrate only toward the companies selling more megawatts of equipment. It can migrate toward the interfaces that reduce time-to-power, remove conversion and distribution constraints, integrate power and thermal infrastructure, standardize deployment, manage transient behavior, and retain the lifecycle relationship after the facility is energized.

The strategic question is therefore larger than whether 800 VDC becomes a major standard. The more important question is: when AI changes the grid-to-compute path, which control points become scarce, who owns them, and where do margin pools move?

1. Power Has Become an AI Infrastructure Constraint

The demand backdrop is large enough to force architectural change. The U.S. Department of Energy, citing Lawrence Berkeley National Laboratory, estimated that U.S. data centers consumed about 176 TWh in 2023, or roughly 4.4% of U.S. electricity, and projected 325-580 TWh by 2028. The International Energy Agency's 2026 update kept its central global outlook near a doubling of data-center electricity consumption, from roughly 485 TWh in 2025 to about 950 TWh in 2030, with AI-focused facilities growing faster than the overall category.

The constraint is not only aggregate generation. Data centers are unusually large, concentrated, time-sensitive loads. FERC's June 2026 action directed all six U.S. regional grid operators under its jurisdiction to justify or reform how they integrate data centers and other large loads. NERC has separately elevated large computational loads as a reliability issue, emphasizing both rapid load growth and the unusual operating behavior of power-electronic loads.

That combination changes the executive problem. The question is no longer simply whether a site can buy electricity. The question is whether power can be secured, interconnected, transformed, protected, distributed, buffered, and delivered to the compute load on the schedule and at the density the business case requires.

2. The Grid-to-Compute Architecture Transition

A conventional data-center power path typically begins with utility or generated AC power, moves through substation and medium-voltage infrastructure, steps down through transformers and switchgear, passes through UPS and low-voltage distribution, and is converted again closer to the IT load. For traditional rack densities, this layered architecture is proven and broadly supported by a mature equipment and service ecosystem.

AI racks change the boundary conditions. Higher rack power concentrates current, copper, connectors, protection, conversion equipment and heat into less physical space. The resulting pressure is not confined to the rack. It propagates upstream into row distribution, switchgear, UPS capacity, medium-voltage architecture, site power blocks, utility interconnection and cooling.

The Open Compute Project's 2026 LVDC work makes this transition unusually visible. Google, Microsoft, NVIDIA and the broader OCP community are coordinating common requirements for 800 VDC architectures. The near-term path uses local power racks or sidecars that convert existing 480 VAC infrastructure to high-voltage DC near the compute racks. The longer-term path moves conversion upstream, potentially from medium-voltage AC directly to 800 VDC using transformer-rectifier or solid-state-transformer approaches.

The significance is not that every data center will become DC-native. OCP explicitly frames 800 VDC as an additional deployment option that can coexist with existing AC systems. The

significance is that the industry is designing a transition architecture rather than assuming today's distribution path can scale indefinitely.

3. CSTE Grid-to-Compute Architecture Transition Map™

CSTE evaluates the transition as a chain of economic and control interfaces rather than a list of electrical components. The map begins at power availability and ends at usable rack-level compute power. Each layer has a different bottleneck, decision owner, failure consequence, and opportunity for margin capture.

This framework is intended for strategic analysis - not electrical design. Its purpose is to identify where architecture changes can alter specification authority, supplier position, integration responsibility, deployment risk and lifecycle value.

Layer	What is constrained / enabled	Primary influence	Strategic implication
Grid / Generation	Available MW, interconnection timing, tariff / reliability conditions	Utilities, developers, grid operators	Time-to-power becomes a project constraint
Substation / MV	Transform, protect and distribute campus-scale power	Utilities, EPC/MEP, equipment OEMs	MV capability moves closer to core data-center architecture
Conversion / UPS	Power quality, continuity, AC/DC conversion, transient management	Power OEMs, design engineers, operators	Conversion topology can shift with rack density
Storage / BESS	Ride-through, resilience, smoothing, grid interaction, optional energy services	Owners, utilities, power OEMs, controls providers	Storage can become architecture, not only backup
Facility Distribution	Switchgear, busway, cabling, protection, modular power blocks	OEMs, integrators, contractors	Copper, footprint and repeatability become material
Row / Rack Power	Final conversion and delivery to accelerated compute	Server/rack OEMs, power-rack suppliers, hyperscalers	High density pressures legacy low-voltage approaches
Compute Interface	Usable power converted into tokens / model throughput	Hyperscalers, server OEMs, silicon platforms	Compute roadmap increasingly dictates facility requirements
Controls / Lifecycle	Power quality, monitoring, commissioning, service, upgradeability	Operators, controls firms, service organizations	Operating data and installed-base access create recurring value

Framework use

Track architecture change by layer, then ask: What bottleneck is forming? Who specifies the interface? Who absorbs integration risk? Who owns operating data? Who remains attached to the installed base?

4. Three Architectures Can Coexist During the Transition

The market should not be analyzed as a binary choice between 'legacy AC' and '800 VDC.' At least three architecture families can coexist for years.

First, conventional AC remains a viable and proven architecture across a large installed base and many new deployments, particularly where rack density, facility constraints and customer risk tolerance do not justify a major architectural shift.

Second, hybrid architectures can preserve upstream AC infrastructure while moving AC-to-DC conversion into dedicated power racks or row-level systems adjacent to the compute. NVIDIA's 2026 roadmap positions this as an on-ramp for existing facilities, while Schneider Electric has similarly argued that rack-level conversion can reduce disruption and isolate early deployment risk.

Third, purpose-built DC-native facilities can move conversion farther upstream. OCP's longer-term architecture contemplates direct medium-voltage AC to 800 VDC conversion feeding a DC backbone, with optional integration of battery energy storage and DC UPS functionality. This architecture remains an evolving ecosystem rather than a universally deployed end state.

For strategy, the coexistence matters. A transition market rewards companies that can bridge architectures, not only those optimized for one technical end state. Retrofit pathways, hybrid integration, controls compatibility, commissioning, serviceability and installed-base support can be as important commercially as the theoretical efficiency of the target architecture.

5. Time-to-Power Is Becoming a Design Variable

Grid access is increasingly entering the architecture decision itself. FERC's recent large-load proceedings and NERC's reliability work show that data-center interconnection is no longer a background utility process. It is becoming a front-end development constraint with implications for site selection, generation strategy, load flexibility and project sequencing.

That helps explain the renewed strategic attention to on-site generation, co-location, microgrids, battery storage and modular power systems. Siemens Energy and Eaton, for example, are marketing repeatable data-center campuses with integrated on-site generation as a way to decouple deployment schedules from slower grid expansion. The commercial signal is broader than any specific vendor claim: customers are willing to reconsider the traditional boundary between 'utility supply' and 'data-center infrastructure' when schedule risk becomes economically material.

CSTE interprets this as a migration in the meaning of power availability. Historically, power could be treated as an input to the facility. In the AI era, the ability to secure and orchestrate power can become part of the product architecture and part of the competitive advantage of the site.

6. BESS Is Moving From Backup Adjacent to Architecture

Battery energy storage is another example of a category whose strategic role can broaden as the power path changes. Conventional UPS batteries are primarily associated with ride-through and backup continuity. A more integrated energy architecture can use storage for additional functions such as power smoothing, transient response, peak management, grid interaction or coordination with on-site generation - depending on market rules and facility design.

OCP's emerging direct-MV-to-800-VDC architecture explicitly identifies a common DC backbone as an enabler for optional BESS integration and DC UPS functionality. That does not prove that BESS economics will be attractive for every site. It does show that energy storage is being designed into the architecture discussion earlier and closer to the power-distribution system.

The strategic implication is that the boundary between 'backup power equipment' and 'energy system' can blur. Suppliers with strength across conversion, protection, controls, storage integration and lifecycle software may therefore gain more influence than a component-level market share view would suggest.

7. Factory Integration Moves Value Upstream

As power density rises, field complexity becomes a commercial constraint. More equipment, interfaces and commissioning steps increase schedule risk, quality risk and the burden on scarce technical labor. That creates a structural incentive to move more integration into controlled factory environments.

The market signal is already visible in reference designs, prefabricated power modules, power skids, e-houses, sidecars, integrated power-and-cooling blocks and repeatable campus architectures. Eaton and Trane's 1 GW DSX reference design, for example, is organized around repeatable 102.4 MW compute modules and explicitly links medium-voltage distribution, UPS-backed cooling continuity, 800 VDC evolution, digital connectivity and lifecycle services. Schneider's 2026 reference designs similarly present power, cooling, IT space and lifecycle software as one validated infrastructure design rather than independent equipment schedules.

CSTE calls this Factory Integration Migration™: when schedule, labor, repeatability and execution risk matter more, value can move from field assembly toward standardized, pre-engineered, pre-validated and factory-integrated infrastructure.

8. Power and Thermal Can No Longer Be Designed as Separate Stories

The electrical transition is inseparable from thermal architecture. Every conversion loss ultimately becomes heat, every increase in compute density changes the cooling requirement,

and every decision to relocate power electronics changes equipment placement, airflow, liquid-cooling interfaces, service access and failure domains.

This is why recent supplier reference designs increasingly present power and cooling together. The commercial consequence is not necessarily that one vendor must own the entire stack. It is that the interface between power and thermal becomes strategically valuable. The companies that can define, validate, integrate or service that interface can influence customer decisions disproportionate to the share of hardware they manufacture.

For CSTE, this is also the bridge between the completed AIDC Thermal Infrastructure Intelligence Series and the broader infrastructure-economics work. The thermal series established that cooling follows compute architecture. WP01 extends that logic upstream: power architecture increasingly follows compute density as well, and the two systems converge in the facility.

9. Where Value Can Migrate

The grid-to-compute transition can redistribute value across the ecosystem. The direction of travel is more important than any single vendor ranking.

Equipment suppliers can gain when higher capacity and new architectures require more switchgear, conversion, busway, protection, storage, monitoring and rack-level power hardware. Integrated infrastructure providers can gain when customers prefer validated reference designs and fewer interface owners. EPCs, MEP firms and systems integrators can gain when architecture choices, commissioning and cross-system coordination become more complex. Controls and software providers can gain when power quality, transient behavior, energy assets and thermal systems require tighter orchestration. Service organizations can gain when the installed base becomes more complex, less tolerant of downtime and more dependent on specialized technical capability.

At the same time, some value pools can compress. Standardization can reduce differentiation in individual components. Open interfaces can reduce proprietary lock-in. Factory integration can shift labor away from site-based activities. New voltage architectures can eliminate or resize certain conversion stages. Companies positioned around a legacy layer must therefore ask whether AI growth expands their category, bypasses it, or changes the location at which the function is performed.

10. The Control Point Matters More Than the Component

CSTE's broader infrastructure thesis is that equipment presence and system-level control are not the same thing. A company can ship a large amount of hardware without owning the architecture, specification, operating data, service relationship or next upgrade decision.

The most defensible positions may emerge where a company controls one or more of five scarce interfaces: power access and interconnection; power conversion and distribution architecture; validated system integration; operating visibility and controls; and lifecycle service / installed-base access.

This leads to a different competitive question. Instead of asking only 'Who sells the UPS, switchgear, transformer or busway?', decision-makers should ask: 'Who determines the architecture, who absorbs integration risk, who owns the operating data, and who remains attached to the asset for the next ten years?'

11. What 800 VDC Does - and Does Not - Prove

800 VDC is the clearest current signal that AI power architecture is changing, but it should not be treated as a settled universal end state.

Direct evidence: OCP released an LVDC whitepaper in March 2026 and has active workstreams spanning system architecture, voltage levels, protection, codes, controls and rack distribution. Google, Microsoft and NVIDIA are coordinating open 800 VDC requirements, and OCP reports more than 80 ecosystem partners developing compatible infrastructure. NVIDIA is introducing hybrid power-rack architectures in 2026 and describes direct medium-voltage-to-800-VDC conversion as a longer-term facility path.

Vendor-reported benefits include fewer conversion stages, lower current, lower conductor burden and more usable rack space at very high power densities. Those benefits are credible in direction, but exact efficiency, copper, capex and reliability outcomes are architecture-specific and should not be generalized from a vendor reference design to every facility.

Open questions remain substantial: protection and grounding practices, standards and certification, fault management, service procedures, retrofit economics, supply-chain maturity, operator familiarity, failure-domain design, long-term reliability and the pace at which hyperscale requirements propagate into colocation and enterprise markets.

The appropriate strategic conclusion is therefore not '800 VDC wins.' It is that a coordinated ecosystem is now investing to make high-voltage DC a practical additional architecture, which is itself evidence that traditional power distribution faces scaling pressure from AI.

12. Commercial and Investor Implications

For OEMs and infrastructure suppliers, the transition favors portfolio thinking. The relevant question is whether the company participates in isolated components or in a broader sequence of power access, conversion, distribution, controls, commissioning and lifecycle service. Build / Partner / Buy™ decisions should be evaluated against time-to-capability, strategic control, interface ownership and scarcity value rather than product adjacency alone.

For engineering, construction and integration firms, the opportunity is less about inventing the architecture than becoming the trusted translator between rapidly changing reference designs and project-specific deployment. Standardization can reduce design variation while simultaneously increasing demand for competent integration, commissioning and retrofit execution.

For investors, category growth alone is an incomplete lens. Stronger questions include: Does the company sit at a bottleneck? Does it gain when rack density rises? Does the architecture require

more of its function, or move that function elsewhere? Does it control a specification interface? Can it attach software, service or lifecycle revenue? Is its manufacturing capacity scarce? Can it support both legacy and emerging architectures during the transition?

For data-center developers and operators, the architecture question increasingly becomes an economics question: which design can secure power fastest, support the required compute density, scale through multiple GPU generations, preserve reliability, fit the site and labor environment, and avoid creating an unserviceable installed base?

13. Evidence Ledger

CSTE uses an explicit evidence discipline to separate what is known from what remains a thesis. The table below summarizes the current maturity of WP01's major claims.

Thesis / Claim	Classification	Current support
U.S. data-center electricity demand is rising rapidly	DIRECT EVIDENCE	DOE/LBNL: 176 TWh in 2023; 325-580 TWh projected for 2028.
Large computational loads are changing grid-planning and reliability requirements	DIRECT EVIDENCE	FERC 2026 large-load actions; NERC large-load reliability program and standards work.
Industry is actively standardizing 800 VDC for next-generation AI infrastructure	DIRECT EVIDENCE	OCP LVDC program; Google/Microsoft/NVIDIA collaboration; 80+ ecosystem companies reported by OCP/NVIDIA.
Hybrid AC-to-800-VDC sidecars are an immediate transition path	DIRECT EVIDENCE	NVIDIA and OCP 2026 roadmap; Schneider 2026 rack-level architecture guidance.
Direct MVAC-to-800-VDC is a longer-term facility architecture	DIRECT EVIDENCE / EMERGING	OCP and NVIDIA roadmap; still an evolving deployment ecosystem.
Power and thermal design are converging into integrated reference architectures	REPEATED SIGNAL	Eaton/Trane, Schneider and NVIDIA ecosystem reference-design activity.
Value can migrate toward integration, controls, commissioning and lifecycle service	INFERENCE	Derived from rising interface complexity, modularization and CSTE thermal-series findings.
Factory integration can become a margin pool as field complexity rises	INFERENCE / REPEATED SIGNAL	Repeated modular power-block, skid, sidecar and repeatable-campus strategies.
800 VDC will become the universal data-center architecture	SPECULATION - NOT ADOPTED	Evidence supports an additional standardized option, not a universal outcome.
BESS will create material non-backup economics at most AI sites	OPEN QUESTION	Architecture support is emerging; site economics, market rules and duty cycles remain variable.

14. Open Questions That Should Drive the Next Research Cycle

- How quickly will 800 VDC move from hyperscale reference architecture into colocation and enterprise deployment?
- Which functions are structurally eliminated, resized or relocated when conversion moves upstream?
- How do capex, conversion efficiency, footprint, copper, failure domains and serviceability compare across conventional AC, hybrid sidecar and direct-MV-to-DC architectures on a full-system basis?
- When does BESS create sufficient economic value beyond backup to justify deeper integration into the facility power architecture?
- Which companies control specification and commissioning responsibility at the MV-to-rack interface?
- How much of AI infrastructure will rely on grid supply, co-location, behind-the-meter generation or hybrid energy systems by geography?
- Where does factory integration reduce total installed cost versus merely relocating margin from field labor to equipment suppliers?
- Which lifecycle capabilities - commissioning, remote monitoring, parts, controls, field service, retrofit - become scarce enough to create durable pricing power?

Conclusion: The New Power Path Is a Value-Chain Shift

AI is forcing the data-center industry to reconsider the path between available electricity and productive compute. The pressure appears first as load growth and rack density, but it propagates through grid access, medium-voltage distribution, conversion, backup, storage, busway, rack power, cooling, controls, commissioning and service.

The most important strategic signal is not one voltage level or one vendor announcement. It is the convergence of utility policy, hyperscale standards work, equipment roadmaps, reference designs and modular infrastructure around the same problem: the existing path to compute is becoming a constraint on how quickly and densely AI infrastructure can scale.

That changes where value can be captured. As the architecture evolves, the strongest positions may belong to the companies that control scarce interfaces - time-to-power, conversion, integration, operating visibility and lifecycle support - rather than to the companies with the largest share of any single component category.

For CSTE, that is the governing infrastructure thesis: capital and margin tend to move toward the bottleneck. In AI data centers, power is no longer only an input. The power path is becoming part of the architecture.

Sources and Research Base

Selected primary and first-party sources used for current-state validation. CSTE analysis distinguishes externally documented facts from market signals and inference; vendor-reported performance claims are treated as architecture-specific unless independently validated.

U.S. Department of Energy / Lawrence Berkeley National Laboratory - 2024 United States Data Center Energy Usage Report and DOE release. [Source](#)

International Energy Agency - Key Questions on Energy and AI (2026), Executive Summary. [Source](#)

Federal Energy Regulatory Commission - FERC Launches Aggressive Targeted Action to Speed Large Load Integration, June 18, 2026. [Source](#)

North American Electric Reliability Corporation - 2025 Long-Term Reliability Assessment and Large Loads Action Plan. [Source](#)

Open Compute Project - Powering the Next Era of AI: Google, Microsoft and NVIDIA standardization work, Aug. 11, 2026. [Source](#)

Open Compute Project - Data Center Facility / Power - Power Distribution Sub-Project. [Source](#)

NVIDIA - Why Scaling AI Compute Performance Requires a New Power Architecture, Aug. 11, 2026. [Source](#)

Schneider Electric - 800 VDC: Powering the Future of AI Data Centers, Apr. 20, 2026. [Source](#)

Eaton - 800 VDC reference architecture for AI factories, Oct. 13, 2025. [Source](#)

Eaton / Trane Technologies - 1 GW DSX reference design overview. [Source](#)

Siemens Energy - Modular data center solutions with integrated onsite power generation. [Source](#)

Schneider Electric - Reference Design 113 - 10.4-11.4 MW Vera Rubin AI clusters, Sept. 1, 2026. [Source](#)

About CSTE Enterprises

CSTE Enterprises is an executive advisory and market-intelligence platform focused on HVAC systems, AIDC thermal infrastructure, product and commercial strategy, lifecycle economics and value-chain intelligence. CSTE translates infrastructure shifts into executive, commercial and investment-relevant decisions.

WP01 extends CSTE beyond the completed AIDC Thermal Infrastructure Intelligence Series into a broader infrastructure-economics lens spanning GRID -> POWER -> COMPUTE -> THERMAL -> WATER -> GEOGRAPHY -> CONTROLS -> SERVICE / LIFECYCLE.

NEXT IN THE SERIES

WP02 - Water, Air and the AI Cooling Constraint: how water availability, cooling architecture, heat rejection, climate, resilience and economics interact.