800 VDC Architecture: Industry Alignment & Execution

White Paper

Jared Huntington & Mike Tu

Table of Contents

Executive Summary – Shift to Execution

Over the past several years, data center design has undergone a fundamental shift. The rapid adoption of accelerated computing has transformed traditional facilities into high-density AI factories. Power delivery is no longer a supporting function but a primary constraint to scalability. As LLMs expand with more parameters and experts, increasing the number of GPUs within an NVLink domain is fundamental to reducing the cost and energy per token. This increases the rack density required to maximize AI factory performance from hundreds of kilowatts toward megawatt-class rack deployments.

The initial 800 VDC white paper, published in October 2025, established the technical rationale and feasibility of transitioning to 800 VDC distribution to address these challenges. Since its publication, the industry has made measurable progress. Stakeholders across the ecosystem, including cloud service providers (CSPs), data center operators, equipment manufacturers, and standards organizations, have aligned on 800 VDC as a practical and scalable approach for next-generation AI infrastructure.

Compute platforms are evolving to support 800 VDC integration. Current and near-term systems are designed to accept 800 VDC input, utilizing 800 VDC to 54 VDC power shelves that are compatible with existing AC/DC implementations. Looking forward, future platforms are targeting fully native 800 VDC architectures, allowing compatibility with both 800 VDC and AC-based AI Factories through the use of power racks or data hall-level rectifiers. This alignment between compute and facility infrastructure establishes a clear path for end-to-end system integration and optimization.

It is important to emphasize that 800 VDC is not intended to replace existing 415/480 VAC systems, but rather to complement and coexist alongside them as an additional deployment option. This approach enables the industry to:

- Leverage existing AC infrastructure and investments
- Avoid disruption to established supply chains and operational practices
- Adopt 800 VDC incrementally, based on application needs and deployment timelines while aligning future infrastructure with next-generation AI compute requirements.

This coexistence model provides a flexible deployment framework, enabling targeted adoption of 800 VDC for high-density compute environments while maintaining compatibility with established AC-based infrastructure. To support this approach, the NVIDIA DSX reference design will be updated to incorporate 800 VDC architectures with multiple deployment options, allowing operators to tailor implementation based on business needs, existing facility conditions, and deployment timelines. This flexibility enables a practical transition strategy while aligning future infrastructure with next-generation AI compute requirements.

At the system level, architectural approaches have begun to converge into a limited set of deployable configurations. These include rack-level adaptation solutions, row-level rectification, and data hall-level high-density DC distribution architectures. In parallel, equipment manufacturers are developing DC-native solutions, and regulatory bodies are engaging certification pathways and safety considerations.

In addition, the dynamic and synchronized nature of AI workloads introduces new requirements for compute rack power inputs. The design approach prioritizes localized peak shaving, power smoothing, and slew rate control using minimal energy storage (e.g., electrolytic capacitors) deployed within or near the compute racks. This establishes a stable and predictable rack workload power characteristic at the rack interface, enabling the upstream power system to be designed for system stability, operability, grid performance requirements, and long-term reliability. Beyond rack-level mitigation, additional energy storage and power conditioning solutions may be implemented at the facility or grid interface to meet grid interconnection requirements and support behind-the-meter (BTM) generation strategies.

This document reflects a transition from conceptual evaluation to coordinated execution. The focus is on:

- Establishing a unified system architecture framework
- Defining a bounded and interoperable equipment ecosystem
- Advancing engineering design practices, including grounding, protection, and fault management
- Developing a system-level simulation and validation framework to ensure interoperability across vendors
- Supporting regulatory alignment, including ongoing work related to DC arc flash methodologies and safety standards

The next phase of industry activity shifts from alignment to execution. Over the next 12 months, the focus will be on accelerating equipment development, validating system-level performance, conducting pilot deployments, and advancing standards and regulatory efforts. These activities are intended to deliver production-ready 800 VDC solutions that support future compute platforms and enable scalable AI factory deployment.

The adoption of 800 VDC is now actively underway across the industry. Through a coexistence strategy with existing AC systems, the industry can enable a practical, low-risk path toward higher efficiency and scalability while fully utilizing existing AC infrastructure and designs and ensuring continuity of operations. Continued coordination between infrastructure, compute, equipment, and regulatory stakeholders will be essential to support safe, scalable, and efficient deployment of next-generation AI Factories.

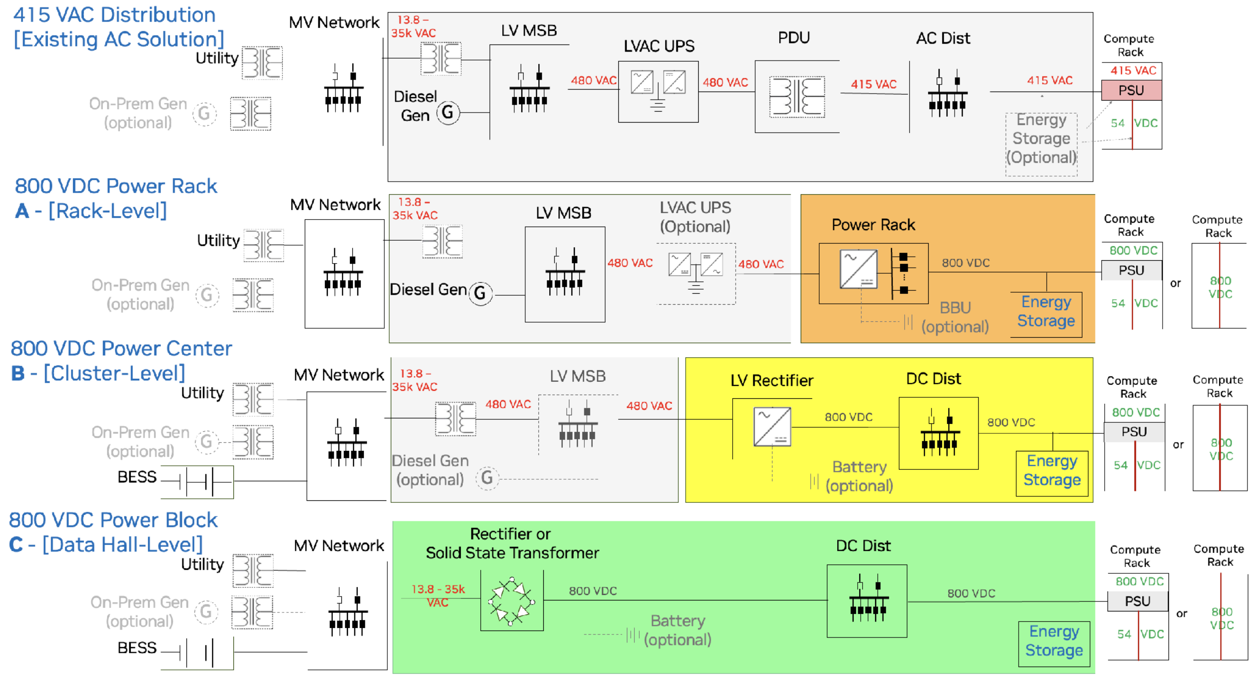

Figure 1: Flexible Power Architecture Options with AC & 800 VDC for AI Factory Deployment

Progress & Ecosystem Momentum

Since shifting from feasibility studies to deployment planning, ecosystem alignment has matured into concrete development milestones across compute, facility, and equipment stakeholders.

Key design parameters such as power block sizing, system equipment scope, compute rack interfaces, and redundancy models have converged into standardized requirements. This reduces ambiguity allowing equipment manufacturers to focus resources on interoperable production-ready designs.

To accelerate ecosystem readiness, the industry is adopting an open OEM development model. Key requirements such as, power quality, power smoothing, interconnects, and end-to-end system stability criteria are being standardized. By openly sharing these requirements, equipment vendors can develop interoperable solutions with greater confidence while ensuring compatibility with current and future compute platforms. This focused approach reduces unnecessary design variations allowing for quicker development cycles and shared learnings. The Open Compute Project (OCP) is a key industry forum for sharing lessons learned, discussing data hall-level 800 VDC architectures, and driving broader ecosystem participation through open standards and collaborative development. Figure 2 demonstrates the joint collaboration among Google, Microsoft, and NVIDIA announced during 2026 GTC Taipei.

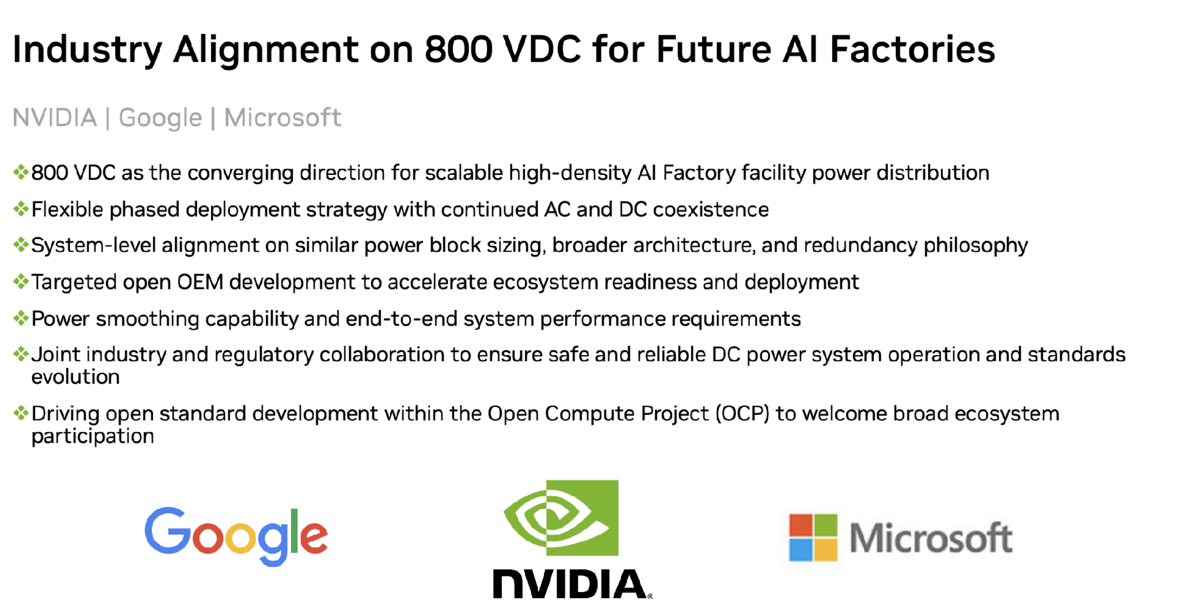

Figure 2: Building the 800 VDC Ecosystem: CSP, OEM, and Regulatory Collaboration

Rack Power Roadmap

NVLink allows for multiple GPUs to connect over copper acting as one large synchronous GPU. The limit of this copper reach requires that more GPUs are put into a small area to maximize performance. This increases local power demands as more GPUs are pushed into the compute rack. Even though local power density increases, this gives an overall performance improvement for every watt used for compute giving improved overall performance and power utilization.

To support this scaling trajectory, NVIDIA is evolving rack-level power architectures in parallel with compute development. Illustrated in Figure 3, near-term generations will continue leveraging the proven 54 VDC distribution plane while adopting fully liquid-cooled power delivery components (Gen 3), including power shelves, busbars, and connectors, to support increasing current requirements. However, as rack power approaches and eventually exceeds practical low-voltage distribution limits, the future native 800 VDC rack architecture (Gen 4) is expected to provide a more scalable path forward by significantly reducing current, simplifying power delivery, and enabling continued compute density growth beyond the capabilities of traditional 54 VDC designs.

The rack power roadmap reflects this continued growth in rack-level power density. The Gen 1 power racks support up to 145 kW utilizing 60A power whips in an 8 to make 6 configuration. The Gen 2 newly released power architecture with Vera Rubin NVL72 increases the rack power capability to 330 kW per rack with both AC input and DC input options. Gen 3 introduces 100% liquid-cooled power shelves which can support up to 570kW in a single rack position. Gen 4 is the planned transition to 800 VDC in rack distribution taking power directly from the facility and down converting it as close as possible to the point of load.

In this context, 800 VDC has emerged as the confirmed and practical path for scalable, high-density power delivery. It represents the highest voltage level that can be broadly adopted at the rack while meeting the densities required to be converted locally within the compute rack. Operating at 800 VDC significantly reduces current and shrinks the power distribution footprint compared to lower-voltage systems (e.g., 54 VDC). It also enables alignment between rack- and data hall-level architectures, minimizing power conversion stages and simplifying overall system design.

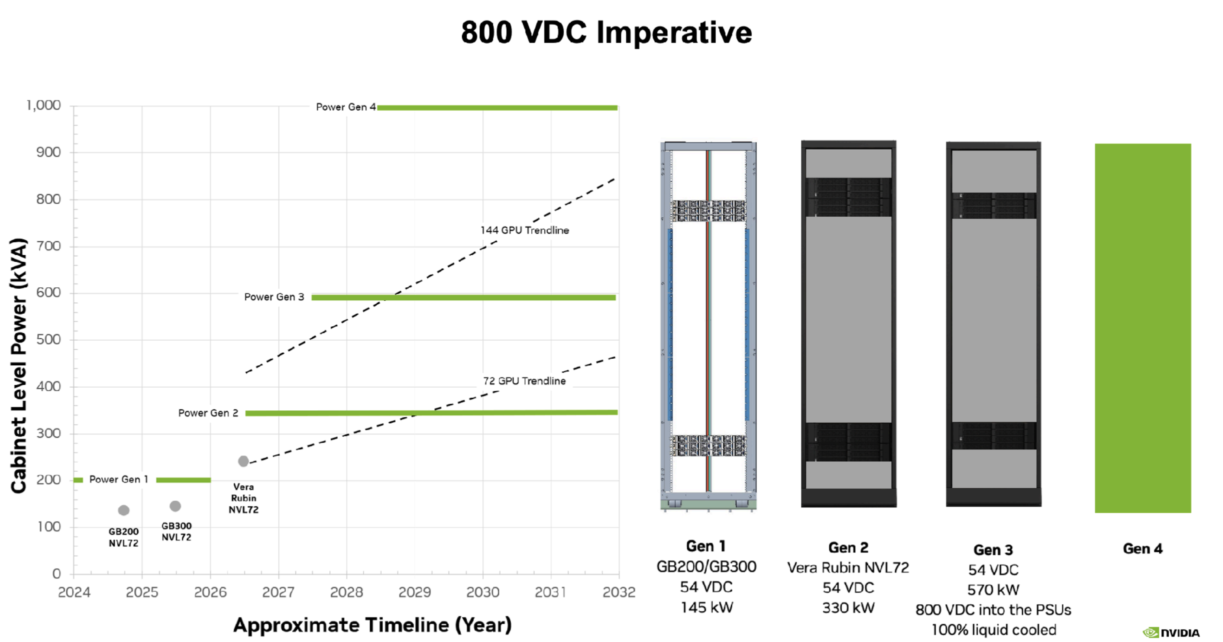

Figure 3: NVIDIA Rack Power Roadmap and the Path to Native 800 VDC Architecture

Converged Architecture & Deployment Options

Following the industry alignment described in the previous section, 800 VDC system design has converged into a focused set of practical and deployable architectures supporting a wide range of AI Factory deployment conditions. These architectures are not sequential requirements, but flexible deployment options enabling operators to adopt 800 VDC based on existing infrastructure, deployment timeline, business priorities, and targeted compute density. These converged architectures provide a flexible and non-disruptive path for 800 VDC adoption as shown in Figure 4:

- Existing Option - shows the default solution to host NVIDIA compute racks in AC distribution AI Factories such as DSX
- Option A enables 800 VDC input racks in facilities designed for AC and enables BBUs
- Option B supports effective and selective cluster-level 800 VDC deployment
- Option C delivers data hall-level 800 VDC power distribution system

The DSX Reference Architecture enables faster deployment, improved interoperability, and standardized infrastructure planning across AI Factory deployments. Today, DSX primarily standardizes AC-based GPU compute pod architectures utilizing 480 VAC distribution from utility interconnection through the compute rack.

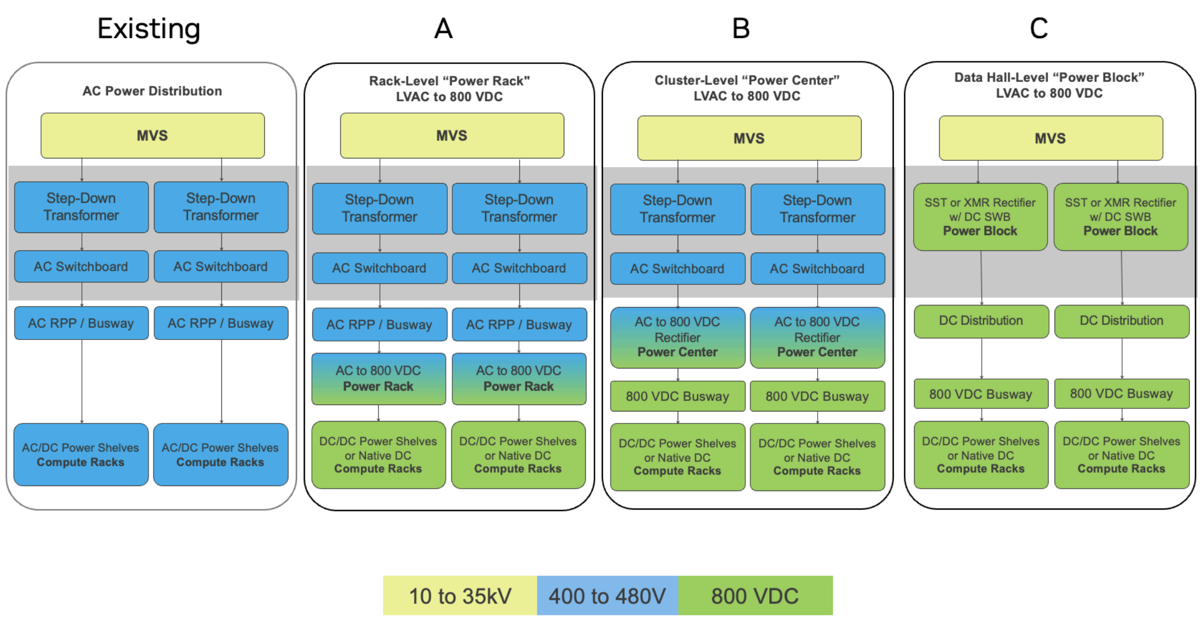

Figure 4: 800 VDC Power Distribution Deployment Options – GPU Compute Load Focused

As AI Factory power density continues to scale, DSX is evolving to incorporate 800 VDC deployment options to provide additional flexibility for future high-density compute infrastructure. As illustrated in Figure 5, the initial 800 VDC deployment focus is targeted within the GPU Compute category, where rack power growth is the most aggressive and the benefits of higher-voltage power distribution are most impactful. This targeted approach allows 800 VDC architectures to be introduced at the rack, row, or pod level while leveraging existing upstream facility infrastructure, minimizing deployment risk and accelerating adoption.

This approach enables gradual adoption of 800 VDC while maintaining AC and DC coexistence across future AI Factory deployments. It also provides a common design framework for system architecture, interoperability requirements, power quality expectations, and future scalability as the industry transitions toward increasingly higher-density compute platforms.

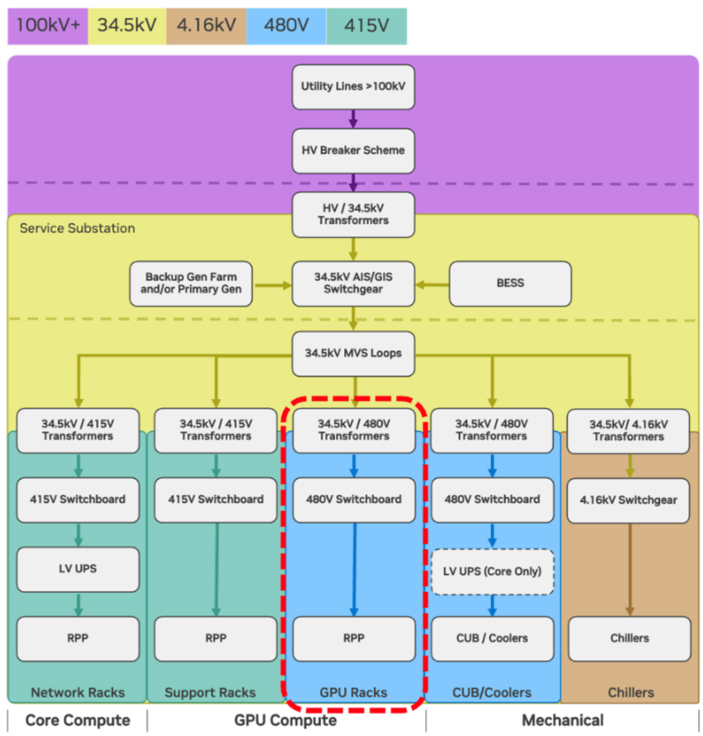

Figure 5: DSX Electrical Service Plan and 800 VDC Implementation Scope (circle in red)

Existing Option – AC-Based Deployment Architecture (Current Baseline)

Existing option represents the current standard deployment architecture for NVIDIA compute infrastructure using traditional AC power distribution systems. This deployment model remains the default and broadly adopted solution for today's AI Factory deployments and continues to be fully supported within NVIDIA's DSX reference design framework.

Current high-density compute racks are powered through:

- 480 VAC power distribution
- 100A AC power whips
- 4-to-make-3 power feed architecture for rack-level redundancy and availability

Under this architecture, as shown in Figure 6, AC power is directly delivered to compute rack power shelves, where conversion to low-voltage DC occurs within the rack. The overall pod-level deployment layout, electrical topology, cooling integration, and infrastructure design methodology are standardized and covered under NVIDIA DSX reference architectures. DSX provides a system-level deployment framework for scalable AI Factory infrastructure, including:

- Compute pod architecture
- Rack deployment methodology
- Power and cooling integration
- Mechanical and electrical layout guidance
- Deployment scalability and operational consistency

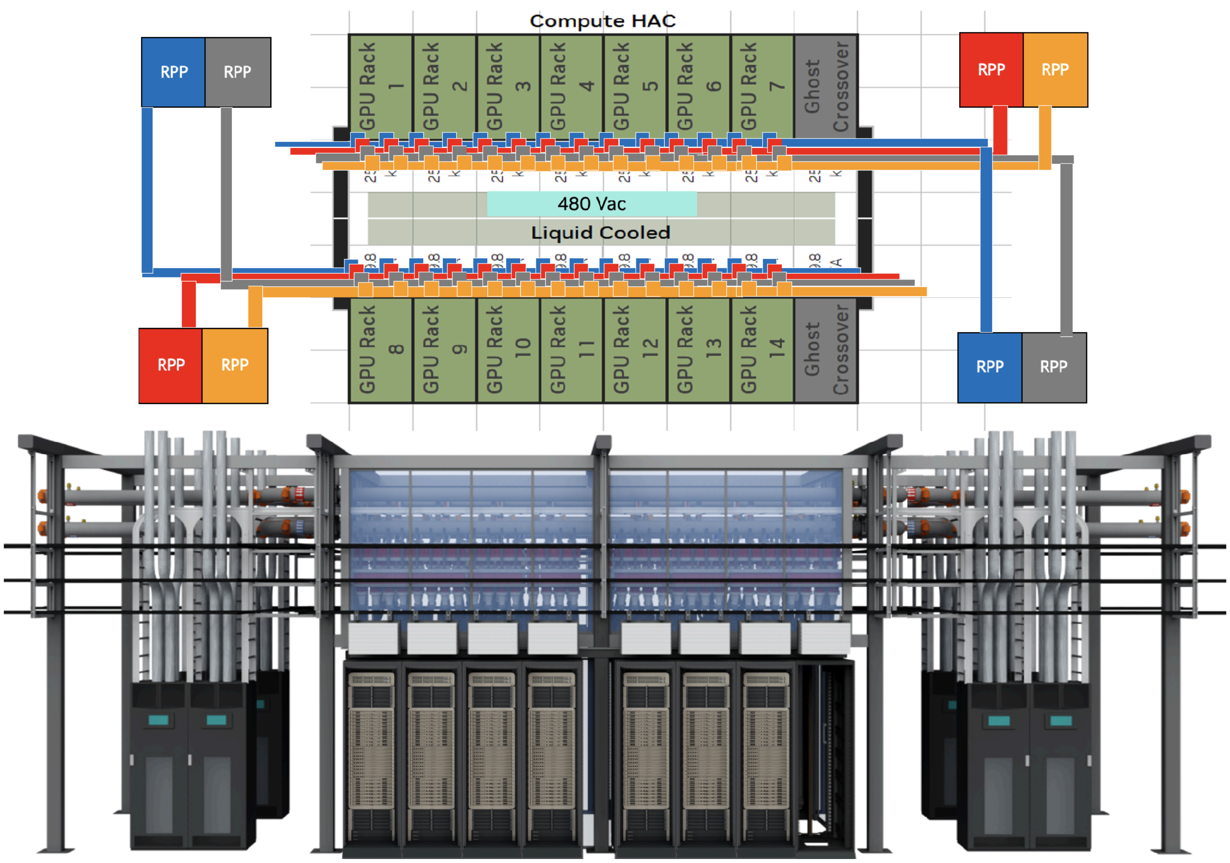

Figure 6: Existing VR NVL72 Compute Scalable Unit Equipment Layout

Option A – Rack-Level Integration (Power Rack)

Option A introduces 800 VDC in the whitespace through localized conversion using dedicated Power Racks positioned adjacent to compute racks. Initially, the architecture supports approximately 660 kW per Power Rack and enables early adoption of 800 VDC while maintaining compatibility with existing AC facility infrastructure.

Each Power Rack accepts AC power from the existing facility distribution system and converts the power to 800 VDC for adjacent compute racks. The initial design adopts a standard 19-inch rack form factor and uses point to point wiring to not require an AC or DC rack power distribution unit or DC busbar. Figure 7 below illustrates the cluster-level layout for Option A.

Power racks are a necessary step in the deployment of 800 VDC. They are required to support rack configurations where the total power capacity cannot fit into a single compute rack in the short term. Initially, this includes Battery Backup Units (BBUs) and systems requiring N+N feed redundancy. Longer term, this capability will be needed for compute racks with power densities that can only be supported by 800 VDC inputs, while still enabling deployment within existing AC data centers, test facilities, laboratories, and other environments that may not be upgraded to 800 VDC distribution. Compared to traditional AC-powered compute racks, this approach significantly reduces compute rack power shelf space consumption while enabling standardized 800 VDC rack interfaces using 125A DC power whips and interlocked connector systems.

Key deployment considerations include:

- Additional row footprint required for Power Rack placement
- Existing row-level AC capacity may require validation
- 12x to 24x 100A AC power whips per Power Rack need to be managed
- Power Rack density may become a future scaling constraint

This option provides the fastest path for enabling near-term 800 VDC ecosystem readiness while supporting continued AC and DC coexistence within future AI Factory deployments. The Power Rack is entering production in Q3 2026.

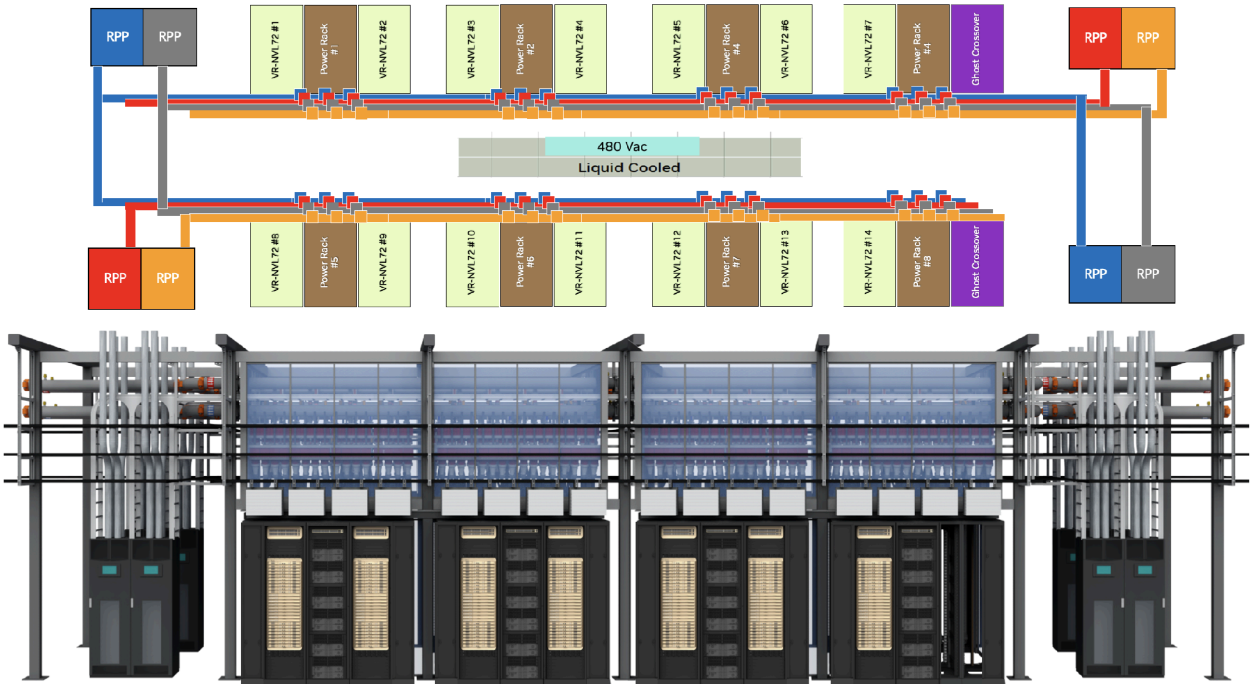

Figure 7: Option A - VR NVL72 Compute Scalable Unit Equipment Layout

VR Power Rack Overview

The initial Vera Rubin power rack design is shown in Figure 8 below. AC power comes into the rack through 12x 100A power whips, which feed through the top of the rack. These whips connect to the facility with IEC60309 4100P6W or 4100P7W connectors. The inputs to the power supply provide a true 3 phase input at 415-480VAC with no Neutral connection required. The whips feed directly into AC to 800 VDC power shelves.

The 800 VDC power shelves are air-cooled with 6x18.3kW PSUs. These PSUs take AC and convert to 800 VDC. Each PSU has 1200 J of energy storage for AC input current shaping to smooth the AC input with dynamic loading.

The figure below shows the rack option with Battery Backup Units (BBUs) and N+1 redundancy, but there is also an option without BBUs and only a 3 RU power shelf height. The BBUs support 80 kW for 60 seconds with 4x20kW BBUs. The BBUs and PSUs then tie to a 125A 800 VDC cross cable used for connecting from the power rack into the compute rack.

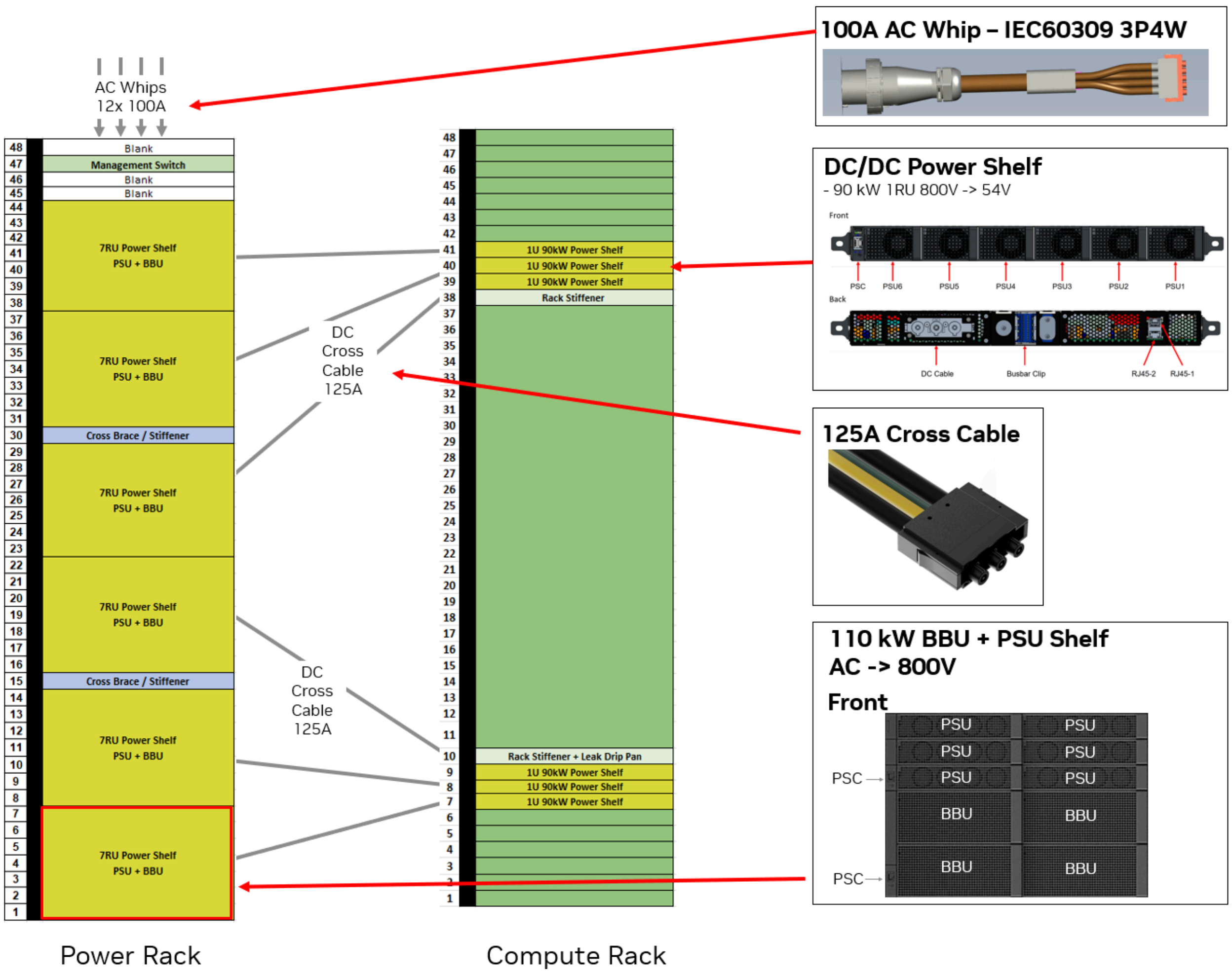

Figure 8: Gen 1 Power Rack

This approach of taking AC directly into the power shelves and outputting DC was chosen to simplify the overall rack design. This helps to eliminate a rack AC power distribution unit, DC power distribution unit and allows leveraging the existing AC to 54 VDC 110kW power shelf solution developed for Vera Rubin.

The 125A 800V cables are then routed to the individual compute racks. The rack above shows a single compute rack, but it is also possible to power racks in a 2:1 compute to power rack ratio as shown in Figure 9. The cables are then routed through cutouts within the power and compute rack. To ensure that the cables are safe for plugging, there is a mechanical interlock and short pin that will power off the cable if there is any break in continuity. This ensures that disconnecting DC whips is not done under load and if anything comes loose, the cable will de-energize.

Internal to the compute rack is an 800 VDC to 54 VDC power shelf with 6x15 kW PSUs. This DC/DC power shelf allows for any compute rack designed around the MGX form factor to then be enabled for 800 VDC data center distribution.

The initial Power Rack architecture is designed to accelerate early deployment readiness by leveraging existing IT equipment certification methodologies under UL 62368-1 while limiting data hall-level 800 VDC exposure. Current certification activities focus on rack-level DC integration. This power rack can be certified now for both UL and IEC region deployments.

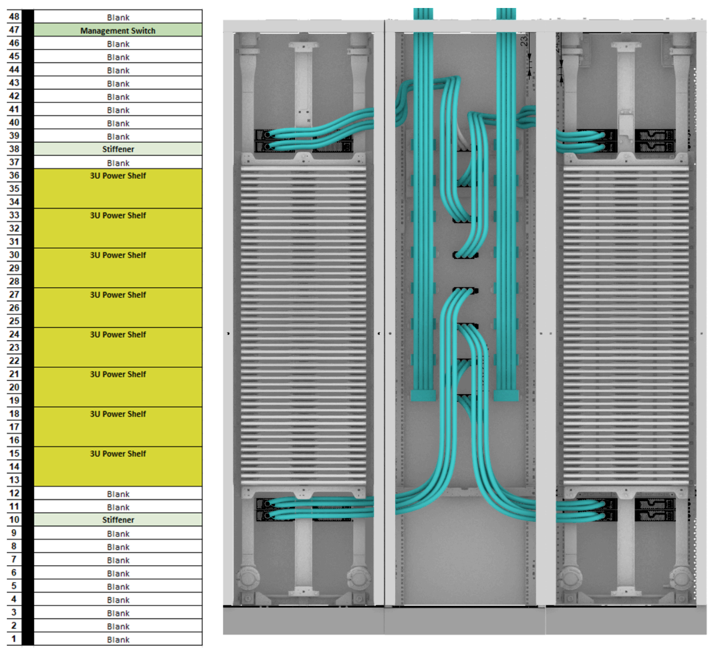

Figure 9: [Left] Non-BBU version of the power rack. [Right] Back of compute and power rack showing a 2:1 compute to power rack ratio and routing

Next Generation Power Rack

As power densities continue to increase, the power rack will continue to evolve. The next generation of power racks will transition to 100% liquid-cooled designs to minimize facility cooling needs driven by power rack density increases. In addition, the BBU power density and power shelf density will continue to increase to 55kW per RU.

Option B – Cluster-Level Deployment (Power Center)

This approach introduces 800 VDC at cluster-level using centralized (end-of-row) rectification and row-level DC distribution. The architecture utilizes centralized DC Power Centers with AC power input and 800 VDC output, through DC busways/Tap Cans or DC Remote Power Panels (RPPs), to support scalable high-density compute deployment while maintaining compatibility with existing upstream/Grey Space AC infrastructure.

Figure 10 below illustrates the cluster-level layout for Option B. The Power Center architecture enables selective 800 VDC deployment within targeted compute clusters while preserving the existing facility electrical room and gray-space design. Compared to traditional AC deployment, the architecture significantly simplifies compute rack power interfacing through reduced power whip quantity, simplified cable routing, and scalable row-level distribution.

Each DC Power Center targets approximately:

- Adjustable up to 2 MW output capacity
- AC power input protection with proper sizing

Multiple rectification architectures are compatible, including:

- Aggregated modular power shelf architectures, Silicon-Carbide (SiC) based, utilizing Power Rack architecture
- Monolithic IGBT-based rectifiers derived from existing UPS systems

The output 800 VDC distribution section is designed with:

- DC feeder breaker(s)
- Preferred 1250A aggregated feeder rating for future scalability
- Overhead 800 VDC busway systems
- 125A tap cans for Day-1 deployment, adopting MCCB or SSCB breakers

The downstream distribution architecture includes:

- Integrated power metering
- Protection relay systems
- Shunt-trip functionality
- Interlocked DC Power Whip and Connector System for Safe Rack-Level DC Interfaces

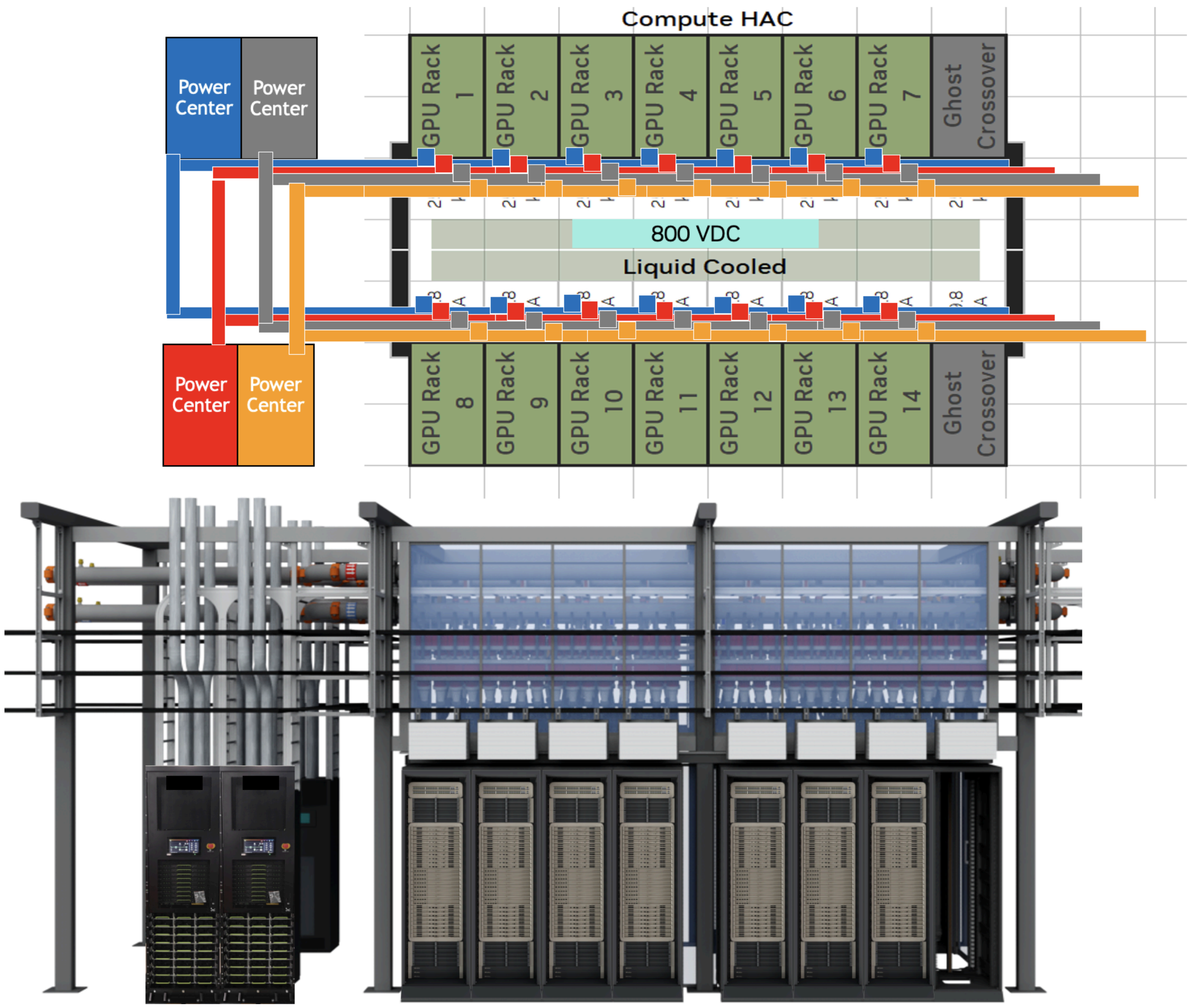

Figure 10: Option B - VR NVL72 Compute Scalable Unit Equipment Layout

Branch Protection Architecture

For the initial 800 VDC tap can design, SSCB or MCCB-based protection are both considered. MCCB may be the most practical Day-1 approach due to broad product availability, supply chain maturity, and established deployment familiarity. At the 125A branch level, MCCBs can support initial tap can protection needs, although clearing time may be up to approximately 20 ms depending on device selection and fault conditions requiring either feed redundancy or additional holdup time to support fault clearing. Compute rack interface connections leverage design principles from EV charging systems to improve operational safety and maintenance workflow. Each DC power whip and connector incorporates:

- Power whip interlock design for safe rack-level DC interfaces
- Mechanical interlocking between the rack connector and upstream tap can contactor

- Normally-open (NO) contactor supervising connector position and lock status
- De-energized connection and disconnection operation

This architecture enables:

- Selective deployment of 800 VDC compute rows
- Mixed AC and DC compute environments within the same data hall
- Retrofit deployment without modification to upstream AC electrical rooms
- Flexible power density scaling at cluster level
- Improved white-space utilization by eliminating side Power Racks

Compared to rack-level conversion approaches, Option B centralizes rectification and distribution functions while maintaining controlled 800 VDC deployment boundaries and standardized rack interface methodology.

Current engineering focus areas include:

- DC breaker selection and certification
- Tap can protection coordination
- Grounding and protection philosophy
- Arc fault and ground fault detection integration
- Busway and connector certification
- Operational maintenance methodology

Option B benefits from limited 800 VDC exposure zones, mature rectifier technologies, scalable Power Center configurations enabled by Power Rack modular power shelves, and a flexible installation methodology suitable for a wide range of deployment scenarios. As a result, this architecture represents high-confidence pathways for scalable near-term 800 VDC AI Factory deployment readiness.

Current industry development activities for Option B include:

- 800 VDC rectifier development and certification
- DC busway and tap can validation
- Connector and interlock safety testing
- Protection coordination methodology
- End-to-end system simulation and grounding evaluation

Multiple OEMs are actively developing products aligned with this deployment architecture to support upcoming pilot deployment readiness. The expected Option B deployment will be as soon as Q3 2027.

Option C – Data Hall-Level High-Density Deployment (DC Power Block)

Option C introduces a data hall-level 800 VDC distribution architecture designed for future ultra-high-density AI Factory deployments. Compared to Option B, this approach further centralizes power conversion and distribution into large-scale DC Power Blocks to improve deployment scalability, modularization, operational simplicity, and long-term infrastructure optimization.

The architecture, shown in Figure 11, targets centralized power blocks at approximately 4.8 MW and beyond, enabling direct data hall-level 800 VDC distribution through high-capacity DC switchboards and busway systems. The design is intended to support future native 800 VDC compute racks and higher-density AI infrastructure without requiring rack-level sidecar deployment.

The initial implementation covers both Transformer Rectifier Unit (TRU) or Solid-State Transformer (SST) based architectures integrating:

- MV AC input section
- Step-down transformer
- LV rectification stage (Adopting variety of AC/DC conversion designs)
- Centralized 800 VDC switchboard (Optional with alternative direct multi-port outputs)
- High-current DC distribution output

The complete system is targeted to be integrated into modularized outdoor containerized deployment packages to support:

- Prefabrication
- Factory testing
- Simplified site installation
- Reduced commissioning complexity
- Reduced indoor electrical room footprint

The architecture also improves future integration flexibility for:

- DC UPS deployment with battery integration
- Battery Energy Storage Systems (BESS)
- Renewable energy interconnection
- DC microgrid capability

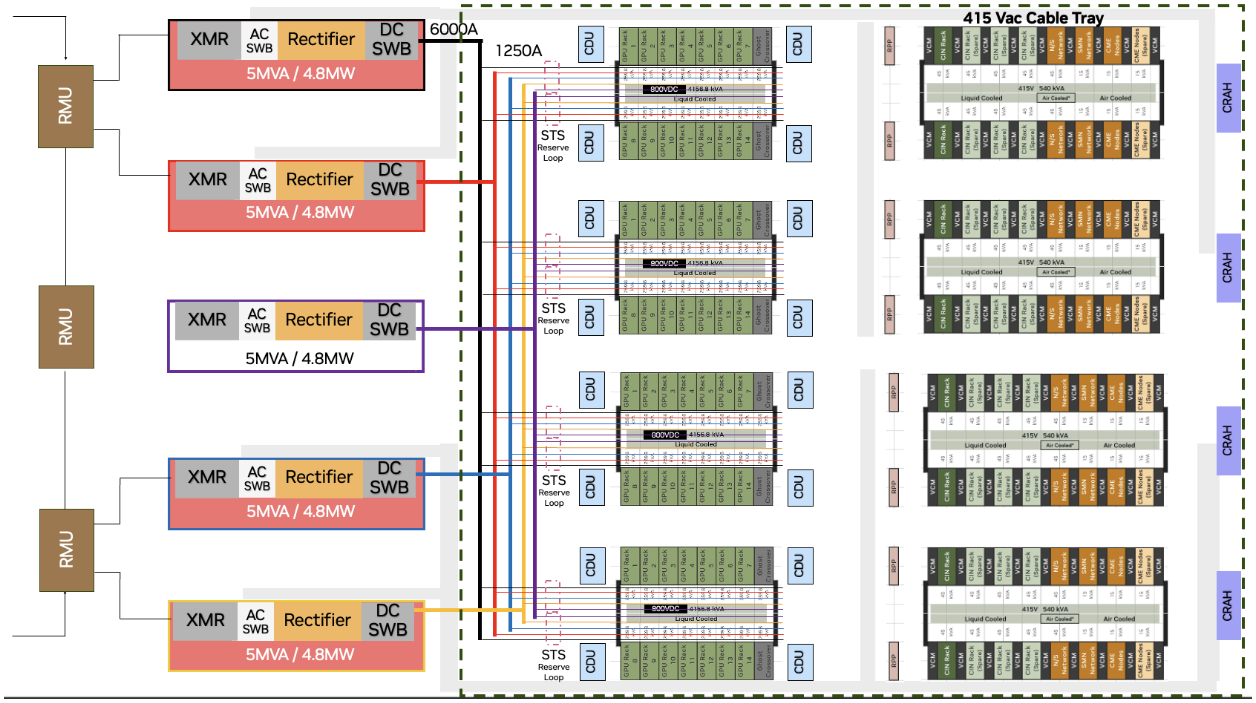

Figure 11: Option C - VR NVL72 Deployment Unit Equipment Layout

Standardized Power Block and Cluster Architecture

A standardized 4.8 MW capacity per cluster is targeted as the ideal deployment building block for Option C. This aligns with the Option B deployment philosophy utilizing approximately 1.6 MW Power Centers configured in a 4-to-make-3 redundancy architecture, resulting in the same effective cluster power capacity at 4.8MW.

To achieve similar deployment consistency at larger scale, Option C adopts a block-redundant architecture with catcher capability. As illustrated in Figure 12, a 4+1 catcher configuration can be implemented using a DC Static Transfer Switch (STS). Unlike AC systems, which require phase synchronization prior to paralleling or transferring between sources, DC sources can be paralleled and switched more directly through voltage regulation and current-sharing controls. This simplifies STS implementation, reduces control complexity, and improves overall system reliability. To coordinate power sharing and fast source transfer between primary and catcher power blocks, the STS architecture may utilize:

- Semiconductor-based OR-ing architectures
- Adjustable DC/DC voltage-controlled catcher paths

With each power block rated at approximately 4.8 MW, each cluster maintains the same standardized power capacity and redundancy philosophy across deployment architectures. This alignment simplifies:

- Infrastructure planning
- System scalability
- Operational consistency
- OEM equipment development

Future-proofing is further enabled through adoption of:

- 4.8 MW power blocks
- 6000A-class DC busduct distribution
- 1250A row-level busway systems

Beyond individual cluster deployment, Option C is designed around a standardized 20 MW Deployment Unit (DU) consisting of four Scalable Units (SUs). This approach establishes a repeatable building-block architecture that can be replicated across a campus to support hundreds of megawatts and ultimately gigawatt-scale AI Factories. By maintaining consistent power block sizing, distribution architecture, and redundancy philosophy, the DU becomes a scalable foundation for future expansion while minimizing engineering variation between deployments.

The modular nature of the architecture also enables prefabricated and factory-integrated solutions, allowing major power infrastructure components to be assembled, tested, and commissioned before site delivery. This approach has the potential to reduce construction complexity, improve deployment speed, and accelerate large-scale AI Factory buildouts.

This architecture can efficiently support current NVL72 deployments while maintaining readiness for future native 800 VDC compute racks approaching 1 MW. As illustrated in Figure 12, future higher-density cluster deployments, such as an 8 MW cluster configuration with 1+1 rack-level redundancy, can be supported without requiring major changes to the upstream 800 VDC facility distribution architecture under Option C.

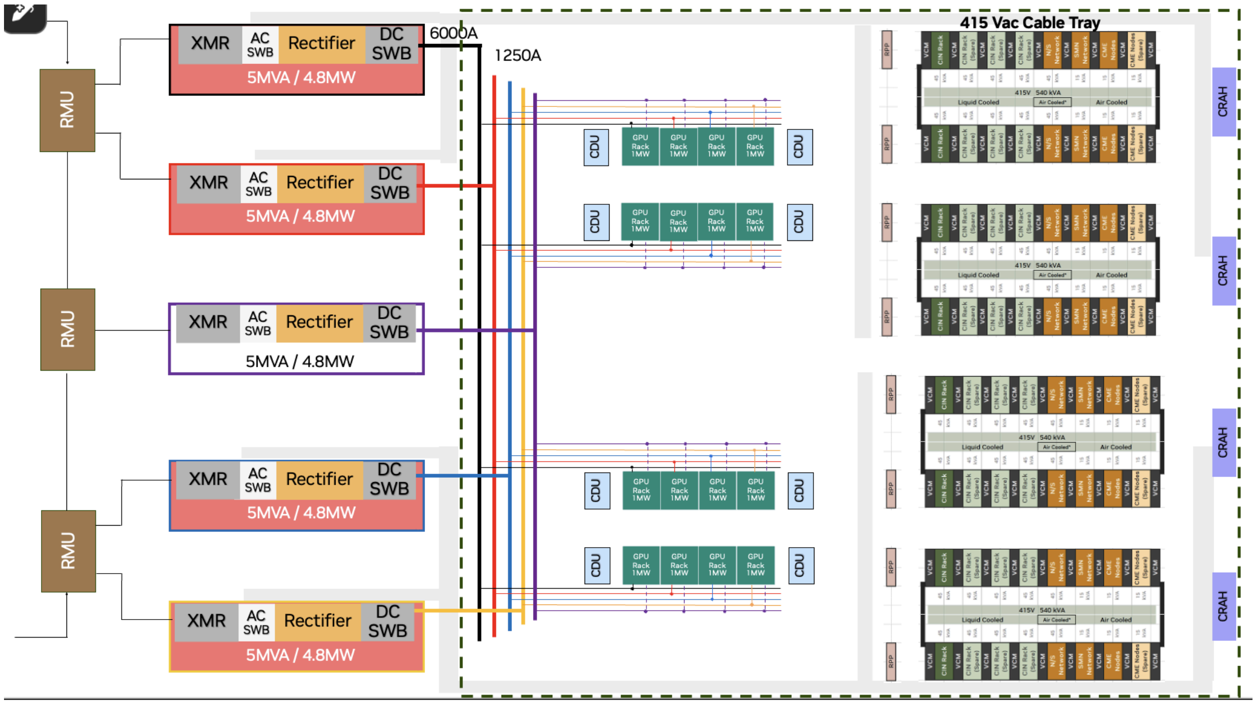

Figure 12: Option C – 800 VDC Native Compute Rack Conceptual Deployment Unit Equipment Layout

Distribution Architecture

The output distribution architecture utilizes:

- High-current DC busbars
- 6000A-class DC switchboard distribution (Optional)
- Parallel 1250A busway distribution
- Standardized tap can interfaces consistent with Option B deployment methodology

This standardized downstream distribution philosophy enables alignment between Option B and Option C deployment models while simplifying future ecosystem development and compute rack integration.

Transformer Rectifier Unit (TRU)

TRU architectures leverage mature line-frequency transformer technologies combined with centralized rectification systems. Similar approaches are already utilized in large-scale industrial and utility power conversion systems. This approach benefits from:

- Mature transformer technologies
- Strong fault isolation capability

TRU benefits (continued):

- Higher deployment familiarity
- Improved compatibility with existing 34.5kV utility distribution infrastructure

Current evaluations indicate TRU architectures remain highly practical for approximately 5 MW-class power block deployment.

Several implementation approaches are being evaluated within the TRU architecture family. One approach leverages proven IGBT-based power conversion technologies commonly used in large UPS systems, replacing the traditional AC output stage with DC/DC conversion to directly generate and regulate 800 VDC. A second approach aggregates higher-density versions of existing rack-level power shelves, utilizing front-end transformers and multiple rectifier modules operating in parallel to achieve multi-megawatt power block ratings. This approach benefits from leveraging mature power conversion building blocks already deployed within NVIDIA compute infrastructure.

Another concept is the “Panama Architecture” utilized in large-scale utility and industrial power systems. The approach uses a line-frequency phase-shifting transformer at the medium-voltage entry point to provide system isolation, voltage transformation, and multiple phase-shifted AC outputs before centralized rectification and DC distribution. Compared to actively controlled converter architectures, the Panama approach relies more heavily on passive magnetic components, offering a simpler and highly familiar design philosophy while maintaining the capability to support large-scale 800 VDC power distribution.

Solid-State Transformer (SST)

SST architectures utilize high-frequency power electronics and high-frequency isolation transformers to achieve compact and highly integrated power conversion systems. Unlike conventional TRU architectures that rely on line-frequency transformers, SST systems leverage high-frequency switching and resonant converter topologies to electronically perform voltage conversion, isolation, and regulation.

Compared to traditional TRU systems, SST architectures may provide:

- Higher power density
- Improved dynamic response
- Advanced controllability
- Reduced equipment footprint
- Increased functional integration

SST technology is being actively evaluated as a potential future implementation path for Options C power blocks. However, most commercially available SST platforms today remain focused on approximately 15kV-class input systems. Since many hyperscale AI Factory deployments utilize 34.5kV medium-voltage distribution, additional development is required to address insulation coordination, protection, thermal management, and utility interconnection requirements at higher voltage levels.

Option C – Next-Gen Architecture (MV Direct Conversion)

Option C will continue to evolve and innovate beyond current deployment-ready 800 VDC solutions. While Options A, B, and C focus on near-term deployment readiness.

The vision centers on eliminating intermediate low-voltage conversion stages by directly converting utility-scale medium-voltage distribution, such as 34.5kV AC, into data hall-level 800 VDC distribution. Mature SST technologies are considered one of the most promising enablers for this architecture due to their potential for high power density, compact footprint, and integrated control capability.

Future power blocks are expected to scale toward approximately 10 MW and beyond, enabling larger standardized deployment units and simplifying infrastructure expansion for multi-hundred-megawatt and gigawatt-scale AI Factories. Advanced research efforts are ongoing across the industry to improve SST scalability and enable higher-voltage applications, making SST architectures increasingly attractive for next-generation AI Factory infrastructure. The next generation SST is expected to be launched toward 2029.

If successfully matured, this architecture may provide:

- Direct 34.5kV AC to 800 VDC conversion
- Higher power density and system efficiency
- Reduced infrastructure footprint
- Simplified power conversion architecture
- Faster deployment through modularized power blocks

In addition to supporting future high-density compute deployments, direct MV-to-800VDC architectures could further enable tighter integration between compute infrastructure, battery energy storage systems, renewable energy resources, and DC microgrids. As AI Factory scale continues to grow, these capabilities may become increasingly valuable for improving system flexibility, resiliency, and energy utilization.

Option C is therefore positioned as a future roadmap direction for continued industry research, ecosystem development, and standards evolution beyond the deployment horizons of current 800 VDC architectures.

Engineering Design Updates

AC vs. DC Electrical Fault Behavior

Table 1 summarizes several key differences between traditional AC power systems and future 800 VDC distribution architectures from a protection and operational safety perspective. While DC systems introduce unique challenges associated with sustained arcing, grounding, and fault interruption, they also enable faster fault current control through converter-limited architectures and emerging solid-state protection technologies. The comparison also highlights the importance of system-level mitigation strategies for future 800 VDC deployment, including:

Key mitigation strategies (continued):

- Interlocked Connector Mechanisms
- Insulation Monitoring Devices (IMD)
- Residual Current Monitoring (RCM)
- Solid-State Circuit Breakers (SSCB)
- Fast fault isolation methodologies

These protection approaches are expected to become foundational design elements for future high-density 800 VDC AI Factory infrastructure.

Topic	AC System Behavior	DC System Behavior	Key Consideration
Fault Current	Natural current zero-crossings every cycle	Continuous current with no zero-crossing	DC faults require faster interruption and current-limiting strategies and arc-fault detection methods
Fault Energy	Higher fault contribution from utility and rotating machines	Often lower in converter-limited systems	Depends on source type and control architecture
Shock Hazard	Can cause muscle lock-on effect	Often causes a single muscle contraction	Both require strict personnel protection
Protection Devices	Mechanical breakers widely used	SSCBs and hybrid breakers enable ultra-fast clearing	Faster clearing reduces fault energy
Ground Fault Detection	Overcurrent and ground-fault relays	IMD and/or RCM commonly required	Depends on grounding architecture
Series Arcing	Arc is periodically weakened; easier to self-extinguish	Arc can sustain without zero-crossings; harder to extinguish	Connector design and arc detection become critical
Parallel Arcing / Arc Flash	High fault current can produce significant arc flash energy	Converter-limited systems may reduce incident energy	Current limiting and rapid fault isolation are key

Table 1: AC vs. DC Fault Behaviors

Grounding & Protection Zoning Strategy

Options B and C introduce 800 VDC distribution at the facility level, making grounding and protection system design a critical engineering consideration. Personnel safety and reliable operation require coordinated electrical fault control, detection, and isolation across the power distribution architecture. Figure 13 shows four grounding approaches under evaluation: high-resistance midpoint grounding, high-resistance return grounding, floating-ground, and solid grounding.

800 VDC distribution utilizes a two-wire architecture with positive and return conductors. In this configuration, minimizing ground-fault risk to personal safety while maintaining reliable fault detection becomes a primary design objective.

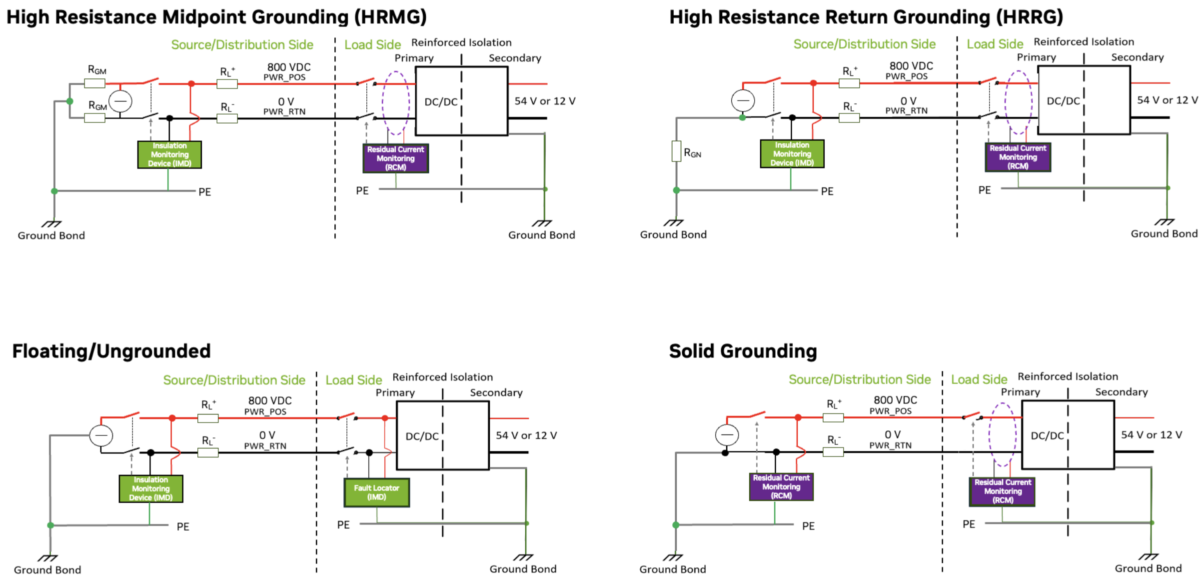

Figure 13: 800 VDC System Grounding Options

High-resistance grounding methods, including High-Resistance Midpoint Grounding (HRMG) and High-Resistance Return Grounding (HRRG), provide an engineering approach for controlling ground-fault current through intentionally sized grounding resistors. Proper resistance sizing allows the system to maintain ground-fault current below personnel safety limits while providing sufficient fault magnitude for reliable detection using residual current monitoring (RCM) and protection devices. Both HRMG and HRRG support a balance between fault-current control, fault detection capability, operational continuity, and personnel protection.

While both approaches offer similar ground-fault current limiting benefits, HRMG provides symmetrical grounding of both positive and return conductors through a common midpoint resistor network. This results in balanced conductor-to-ground voltages, improved EMC performance, and more consistent ground-fault detection on both conductors. HRRG provides effective detection for positive conductor faults but may experience delayed detection of return conductor faults due to the grounding reference location. As a result, HRMG provides a more uniform fault monitoring, earlier fault identification, and improved system reliability while maintaining low ground-fault current levels and continued operation following a first ground fault.

Floating-ground systems can theoretically minimize first ground-fault current to near zero while allowing continued operation following a single ground fault. However, insulation leakage and system capacitance may still create fault current paths. In addition, extremely low fault current levels can make fault detection and fault location more challenging, requiring insulation monitoring devices (IMD) and fault locator systems to identify and isolate faults.

Solid grounding represents the simplest grounding architecture enabling fast and sensitive fault detection due to the higher ground-fault current. It also simplifies the protection architecture by requiring isolation and interruption primarily on the positive conductor during fault conditions. This can reduce system complexity and may provide advantages for future native 800 VDC compute rack designs. However, the resulting fault current can be substantially higher than floating or High-Resistance Grounding systems, increasing fault energy exposure and requiring robust protection coordination. To support safe operation, solidly grounded systems require fast protection devices, and minimal ground bond impedance to ensure a safe and controlled fault current return path.

The preferred grounding approach will ultimately be determined through system-level studies evaluating fault-current behavior, protection coordination, personnel safety, operational requirements, equipment availability, and overall deployment practicality.

Protection Zones & Detection Philosophy

The 800 VDC protection architecture is divided into multiple protection zones with different operational and personnel exposure requirements.

Facility Distribution Zone (Source Output to Tap Cans)

The upstream distribution network, from transformer rectifier unit (TRU) or solid-state transformer (SST) output through DC busway to tap cans' input, is treated as part of the facility electrical infrastructure. Access to energized equipment within this zone is expected to be restricted to qualified electrical personnel.

For this protection zone, insulation monitoring devices (IMD) or insulation resistance monitoring (IRM) systems are recommended to continuously monitor system insulation integrity and detect possible ground degradation or fault conditions. Upon insulation degradation detection, alarms may be generated without immediate shutdown to maintain service continuity. If high overcurrent conditions are detected, protective relays can initiate fast source shutdown through coordinated control with upstream rectifier systems. This protection philosophy is applicable to both floating-ground and HRG systems.

Rack Interface Protection Zone (Tap Can to Compute Rack)

The rack interface protection zone represents the portion of the 800 VDC distribution system with the highest likelihood of human interaction during installation, maintenance, and equipment replacement. As a result, the protection philosophy for this zone prioritizes fast fault isolation, sensitive fault detection, and reliable operation while minimizing personnel exposure to energized conductors. Independent fault detection methods, such as Residual Current Monitoring (RCM), are preferred to provide rapid detection of insulation degradation, ground faults, loose connections, and abnormal leakage currents. Combined with local protection devices and monitoring systems, this approach enables early fault identification, selective fault isolation, and improved operational reliability.

Figure 14 illustrates the 800 VDC rack power interface and protection architecture. The 125A DC power whip and connector system is designed to remain de-energized during installation and removal. Each power whip is hardwired to the tap can and protected by an upstream normally-open contactor. The rack-side connector incorporates a mechanical locking mechanism that verifies proper connector engagement. Once the connector is fully inserted and mechanically locked to the compute rack, an enable signal is sent to close the upstream contactor and energize the power whip. During disconnection, the release mechanism first opens the upstream contactor to remove power before unlocking the connector. This interlock sequence minimizes arcing risk, improves personnel safety, and supports reliable operation of rack-level 800 VDC power distribution.

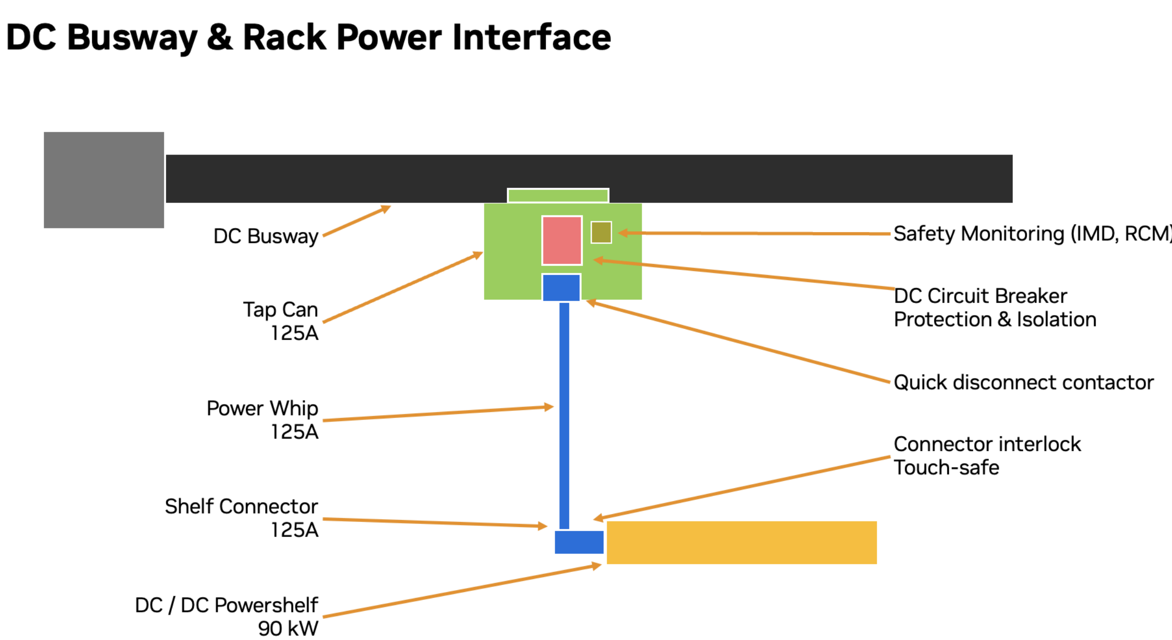

Figure 14: 800 VDC Rack Power Interface and Protection Architecture

Ground Fault & Arc Flash Detection

Additional protection mechanisms are required within the rack interface zone.

For HRG and solid-grounding systems, residual current monitoring (RCM) devices installed at the tap can compare positive and return conductor current to detect downstream leakage current and ground faults. Controlled ground fault current levels improve detection sensitivity and protection reliability.

For floating-ground systems, fault location systems integrated with upstream IMD platforms may be used to identify insulation degradation and potential fault locations. While floating systems can maintain very low ground fault current levels, fault detection sensitivity and location response time may be reduced compared to HRG implementations.

Arc fault mitigation is also required throughout the system. Connector temperature monitoring can identify overheating caused by loose connections, while high-frequency sensing technologies may be used to detect series arcing conditions downstream toward compute racks. Similar sensing approaches may also be implemented at TRU or SST outputs to monitor system-level DC distribution faults.

Source Fault Current Control

Transformer rectifier units (TRUs) and solid-state transformers (SSTs) should support fast shutdown capability through shunt-trip operation and internal disconnect isolation to rapidly clear fault current contribution.

In addition, deployment of DC/DC converter stages or equivalent current-limiting technologies at DC sources is required to reduce available fault current during abnormal conditions. Similar current-limiting approaches should be used for batteries, capacitors, and other energy storage systems connected to the DC network.

Effective source fault current limitation is expected to play a critical role in future DC arc flash mitigation, protection coordination, and safe operation of large-scale 800 VDC power distribution systems.

Solid-State Circuit Breakers (SSCB)

Solid-state circuit breakers (SSCBs) are emerging as a key enabling technology for future high-density 800 VDC power systems. Unlike conventional fuses or thermal-magnetic circuit breakers, SSCBs detect faults electronically and interrupt current before the fault reaches its maximum value. By limiting both the peak fault current and fault duration, SSCBs significantly reduce the let-through energy (I^2t), minimizing equipment damage, reducing arc-flash risk, and improving personnel safety. As illustrated in Figure 15, conventional protection devices may allow fault current to continue rising for several milliseconds before interruption, whereas SSCBs can isolate faults within sub-millisecond timeframes, preventing the high fault currents that drive incident energy. This fast, sensitive, and selective protection capability makes SSCBs well suited for future high-density 800 VDC power distribution systems. To accelerate ecosystem development and reduce OEM implementation burden, NVIDIA is initially focusing SSCB development on two target ratings:

- 125A air-cooled SSCBs for DC/DC (800V-to-54V) Power Shelves deployed within compute racks
- 1250A SSCBs, likely utilizing liquid-cooled designs, for future native 800 VDC compute rack power interfaces

In parallel, NVIDIA is evaluating the practical upper current limit of air-cooled SSCB technologies. The results will help establish future rack power interface requirements and guide next-generation native 800 VDC compute rack architectures.

While SSCB technology is technically mature at lower current ratings, broader deployment remains dependent on supply chain readiness, regulatory certification, product availability, and cost competitiveness. To support near-term deployment schedules, MCCBs are expected to serve as the primary branch protection device for initial production systems and lower-density compute rack applications.

As AI compute density continues to increase as well as the size of power domains, SSCBs are expected to become the preferred long-term protection solution due to their ultra-fast fault clearing capability, reduced fault energy, selective protection coordination, and improved compatibility with future high-density DC power distribution architectures.

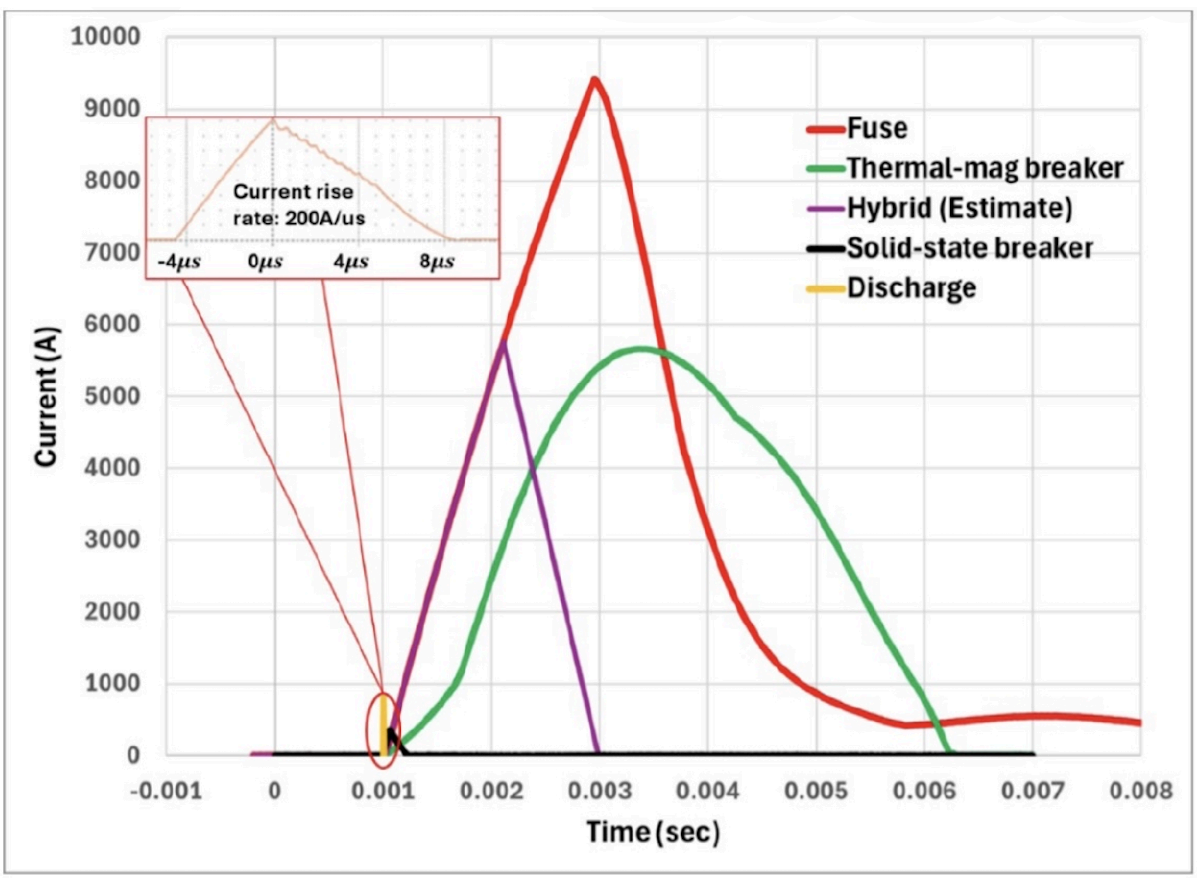

Figure 15: DC Short circuit fault clearing performance among fuse, mechanical circuit breaker, and solid-state circuit breaker (Source: Siemens - Protections for Data Centers Powered by Direct Current)

Execution Framework — Industry Alignment and Ecosystem Acceleration

As previously mentioned in Section 3, a key milestone in this transition has been the convergence of system-level architecture direction among major AI infrastructure stakeholders. Rather than pursuing fragmented deployment models, the industry is increasingly aligning on common deployment philosophies, including:

- Standardized 800 VDC facility distribution targets
- Similar power block sizing approaches
- Common rack power interface requirements
- Shared redundancy philosophies
- Flexible AC and DC coexistence strategies

This alignment significantly reduces architectural ambiguity across the ecosystem and enables more focused infrastructure development for next-generation AI Factories. At the same time, OEM development efforts have become increasingly targeted and execution-oriented. By narrowing the deployment scope to a smaller set of converged architectures and standardized interfaces, equipment suppliers can prioritize engineering resources toward production-ready systems instead of supporting multiple incompatible designs. This enables accelerated development of:

- Rectifiers and power blocks
- DC switchboards and protection devices
- Busways and tap cans
- Rack interface connectors and power whips
- Monitoring, protection, and control systems

The converged deployment framework also simplifies the regulatory and certification pathway. A more focused equipment scope enables standards organizations and certification agencies to concentrate on the highest-priority gaps required to support safe deployment and equipment listing. Industry engagement with UL, NFPA, IEEE, IEC, and other standards bodies is now centered on:

- DC grounding and protection methods
- Arc flash evaluation methodologies
- Equipment certification pathways
- Busway, connector, and breaker qualification
- Installation and operational safety requirements

This coordinated ecosystem approach is intended to accelerate deployment readiness while ensuring safe and reliable operation of large-scale 800 VDC power systems.

Certification Strategy and Regulatory Development

The transition toward large-scale 800 VDC AI Factory deployment requires coordinated development of system requirements, equipment certification, and deployment practices. Significant progress has been made through direct collaboration among NVIDIA, cloud service providers (CSPs), OEM partners, and UL Solutions to establish a practical and deployment-focused certification strategy.

Rather than pursuing broad regulatory changes across all DC applications, the current industry approach focuses on a narrowed equipment scope aligned with the deployment architectures defined in this paper. This targeted strategy reduces complexity, accelerates certification activities, and enables regulatory efforts to focus on the equipment and system requirements most critical to near-term AI Factory deployment.

System-Level Requirements First

A key lesson learned during industry alignment activities is that equipment certification alone is insufficient. System-level performance and safety requirements must first be established to ensure interoperability between power infrastructure and NVIDIA compute platforms.

Current workstreams focus on defining:

- Power quality requirements
- Voltage regulation and transient response limits
- Power smoothing expectations
- Grounding methodologies
- Protection coordination requirements
- Fault current management strategies
- Compute rack interface requirements
- End-to-end system stability criteria

These requirements establish a common engineering framework that enables multi-vendor interoperability while guiding OEM product development, certification, and deployment.

System Simulation & Validation Framework

The 800 VDC data center distribution is a new emerging standard with multiple vendors requiring compatibility and interoperability. To quickly navigate through issues and develop standards and requirements initially validation and limits will be set based off simulation and data center modeling.

Simulation Focus Areas

Initial models will define rack, distribution, and load requirements for:

- Dynamic response
  - Frequency Domain Stability Analysis
  - Source Output Impedance / Module Sharing
  - Distribution Inductance Restriction
  - Power quality – Voltage fluctuation / Rippling
- Define minimum equipment performance requirements
  - Rectifier behavior (impedance, response)
  - Stability criteria
- Purpose:
  - Multi-OEM compatibility
  - System reliability validation

DC Arc Flash and Fault Current Control

DC arc flash incident energy is primarily determined by the available fault current (I) and the fault clearing time (t). The proposed 800 VDC architecture addresses both factors through a layered protection philosophy to improve personnel safety and equipment protection.

The first objective is to reduce the available fault current (I). Future 800 VDC systems can leverage current-limited power sources and power conversion technologies, including rectifiers, DC/DC converters, solid state transformers (SSTs), battery energy storage systems (BESS), and other distributed energy resources. Applying current limiting capability across all fault current contributing sources significantly reduces the maximum short circuit current and corresponding arc flash incident energy. To fully realize this benefit, both the current limiting devices and the associated fault current calculation methodology should be validated, certified, and accepted by regulatory bodies so that the reduced fault current can be recognized in equipment certification, protection coordination, and arc flash studies.

The second objective is to minimize the fault clearing time (t). Parallel arc faults can be detected using high-speed protection methods such as overcurrent protection, differential protection, and protective relaying, followed by rapid interruption using solid state circuit breakers (SSCBs) or other high-speed interrupting devices. Series arc faults generally do not produce sufficient overcurrent to operate conventional protection devices and therefore require dedicated arc fault detection techniques based on electrical signature analysis, thermal monitoring, or other advanced sensing methods before the affected circuit is isolated.

This layered protection philosophy provides a scalable foundation for safe 800 VDC deployment while supporting future certification and large-scale commercial deployment.

Regulatory Alignment and AHJ Enablement

NVIDIA is collaborating with UL Solutions, OEM partners, and data center operators to accelerate the commercialization of 800 VDC AI Factory infrastructure. Rather than creating entirely new standards, the current strategy leverages existing certification frameworks wherever practical while identifying targeted updates needed to support emerging 800 VDC equipment and deployment architectures. This approach enables rapid product certification and ecosystem readiness without delaying deployment through extensive regulatory development.

Current collaboration focuses on:

- Equipment certification and applicability of existing standards for 800 VDC products
- Development of installation, commissioning, and operational guidance
- Global certification alignment, including coordination with IEC requirements for international deployment
- Collaboration with industry partners to support consistent certification and deployment practices

In parallel, Authority Having Jurisdiction (AHJ) engagement has become an important enabler for large-scale deployment. Industry efforts focus on developing installation best practices, inspection guidelines, commissioning procedures, and technical training to provide AHJs, inspectors, and commissioning teams with a consistent understanding of 800 VDC AI Factory infrastructure.

Collectively, these efforts represent the industry's transition from technology development to commercial deployment. By aligning product certification, regulatory acceptance, and field implementation, the ecosystem establishes a practical and scalable pathway for widespread adoption of 800 VDC AI Factory infrastructure.

Ongoing Development & Next-Step Focus Areas

The 800 VDC ecosystem has progressed from concept development to deployment planning. The next phase focuses on validating system performance, establishing engineering requirements, and enabling large-scale commercial deployment. Key development activities include:

System-Level Performance Requirements

Establish comprehensive system-level specifications that define power quality, dynamic response, system stability, grounding, protection coordination, and grid compliance requirements. These requirements will serve as the common engineering framework for OEM equipment development while ensuring compatibility with NVIDIA compute platforms and safe end-to-end system operation.

End-to-End System Validation

Develop an integrated validation framework covering the complete power path from medium-voltage AC input to the GPU chip. Hardware testing, system integration, and digital simulation will be used to validate interoperability, stability, protection coordination, reliability, and overall system performance across multi-vendor deployments.

Certification and Deployment Readiness

Continue working with industry partners and regulatory organizations to establish consistent practices for equipment certification, system approval, installation, commissioning, and operational readiness. The objective is to streamline product qualification while supporting repeatable deployment across different regions and jurisdictions.

Operations and Lifecycle Best Practices

Develop standardized procedures for factory acceptance testing (FAT), site acceptance testing (SAT), commissioning, startup, operation, maintenance, and troubleshooting to improve deployment quality, operational safety, and long-term system reliability.

Business Value and Reliability Assessment

Perform comprehensive cost, reliability, and lifecycle analyses to quantify the technical and economic benefits of 800 VDC deployment. The study will evaluate capital investment, operational efficiency, scalability, equipment utilization, availability, and total cost of ownership to establish a compelling business case for future AI Factory infrastructure.

Call to Action

The rapid growth of AI infrastructure is fundamentally reshaping data center power architecture. As AI Factories continue scaling toward gigawatt-class campuses and megawatt-class compute clusters, the industry must evolve beyond traditional AC distribution to support the next generation of accelerated computing.

Since the publication of the initial OCP 2025 white paper, the industry has made significant progress. What began as an architectural concept has evolved into an execution-focused initiative spanning compute platforms, power infrastructure, OEM product development, system simulation, validation, and regulatory alignment. Cloud service providers (CSPs), AI Factory operators, OEM partners, and industry organizations are increasingly converging on 800 VDC as a practical and scalable power distribution architecture for future AI infrastructure.

This alignment provides the foundation for the next phase of ecosystem development. The industry's focus is now shifting from defining the architecture to validating system performance, accelerating product development, establishing certification pathways, and preparing for large-scale commercial deployment. Continued collaboration across the ecosystem will be essential to define common system-level requirements, validate end-to-end interoperability, streamline certification and regulatory approval, and develop standardized deployment, commissioning, and operational practices.

Importantly, 800 VDC is not intended to replace existing AC infrastructure overnight. AC and DC architectures are expected to coexist for the foreseeable future, allowing operators to adopt the deployment model that best meets their workload requirements, facility constraints, and business objectives while providing a scalable migration path toward higher power densities.

The transition to next-generation AI power infrastructure cannot be achieved by any single organization. Success depends on continued collaboration among NVIDIA, CSPs, OEM partners, utilities, standards organizations, regulatory bodies, and the broader industry to develop interoperable products, validate system performance, and enable safe, reliable, and repeatable deployment at scale.

The opportunity ahead extends beyond introducing a new distribution voltage. Together, the industry is establishing a common power infrastructure foundation that will enable future AI Factories to achieve greater scalability, higher efficiency, improved deployment velocity, and long-term sustainability. The work ahead is no longer about proving the feasibility of 800 VDC. It is about executing together to make it the trusted and deployable foundation for the next generation of AI infrastructure.

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051
http://www.nvidia.com