

電熱共生效應

超高密算力系統電源架構、寬能隙半導體與能源管理

編輯部整理

過去通用雲計算時代沉澱下來的供電、配電與熱管理體系，正在面對生成式 AI 百千瓦乃至兆瓦級單機櫃算力的猛烈衝擊。當 GPU 集群功耗成倍攀升，制約 AI 性能釋放的瓶頸不再只是晶片算力、互聯頻寬，而是極易被忽略的電能傳輸鏈路：焦耳定律帶來的線路損耗、

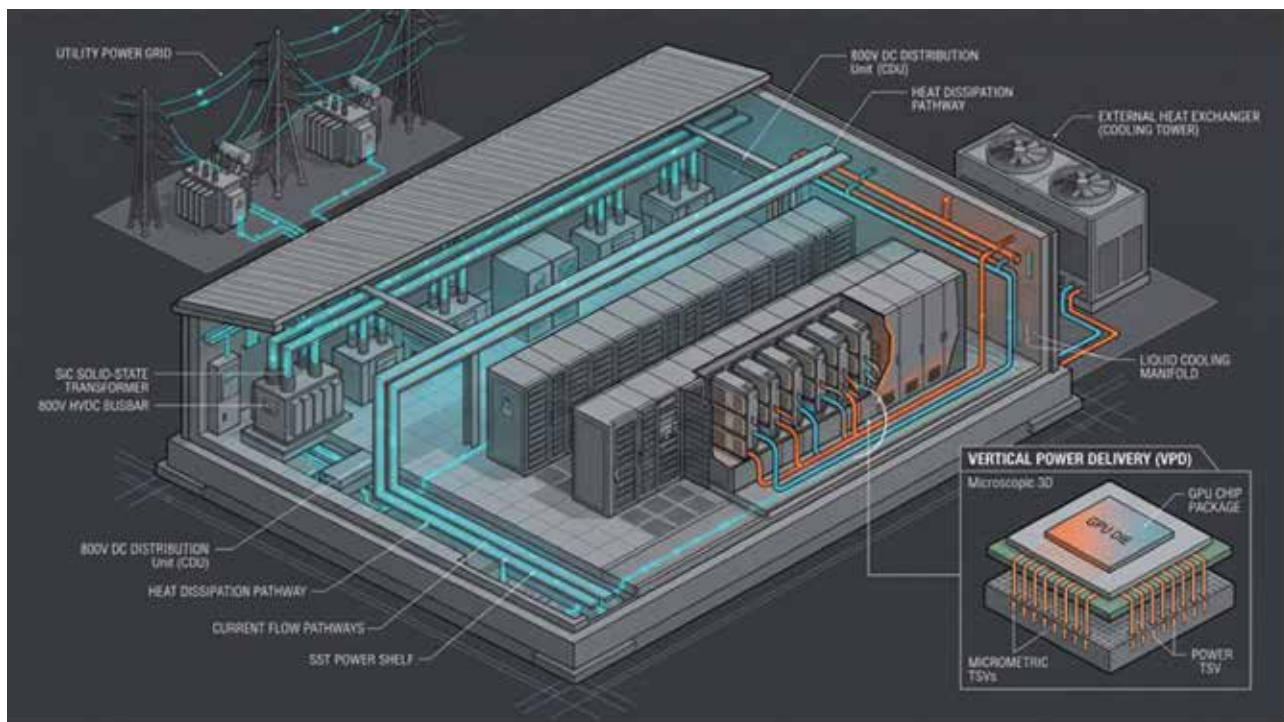
大電流引發的熱聚集、供電與散熱互相強化的電熱共生效應，正在成為萬卡智算集群必須翻越的物理高牆。

供電架構的反覆運算不是簡單電壓數值的提升，而是從電網、機架母線、功率半導體器件，延伸至晶片封裝供電、軟體動態調度的系統性變革。

焦耳定律下的線損危機與配電瓶頸

在雲計算與大中型資料中心發展的黃金時期，資料中心的電力分配長期維持著一種相對穩定的經典架構。傳統供電鏈路通常採用「交流集中式分級變換」方案：市電高壓交流電進入資料中心後，經由中壓

圖說：超高密算力系統的電源架構、寬能隙半導體與能源管理 (AI 製作)



變壓器降壓，通過大型集中式不間斷電源 (UPS) 進行整流與逆變，再經由機架電源分配器 (Power Distribution Unit; PDU) 分配至各個伺服器機架。在機架內部，傳統的伺服器電源模組 (PSU) 最終將交流電轉換為 12V 的直流電，直接輸送到伺服器主機板。這種 12V 直流供電架構在過去十多年間成為了無可爭議的行業標準。根據權威機構 Uptime Institute 發布的《全球資料中心調查報告》，在傳統通用雲計算時代，全球資料中心伺服器機架的平均功率密度長期處於 5kW 至 10kW 的區間內，極少有單機櫃功率超過 15kW。在低功率密度時代，12V 架構因其元器件技術極度成熟、供應鏈成本低廉且安全電壓易於控制，完美契合了當時的算力需求。

然而，近兩年，生成式人工智慧 (Generative AI) 全面爆發與萬卡智算集群的加速落地，以一種近乎摧枯拉朽的姿態打破了傳統電力平衡。專門針對 AI 優化的智算資料中心，其單機櫃功率密度已普遍跨入 50kW 至 100kW 的門檻，部分高密度機櫃甚至突破了 120kW，這相比於傳統資料中心翻了數十倍之多。這一劇烈變化的典型代表便是目前正在市場中大規模量產和部署的 NVIDIA Blackwell 架構集群。根據輝達官方最新的系統用戶指南披露，其旗艦級 GB200

NVL72 液體冷卻機架在滿載運行下的標稱功耗已達到了驚人的 120kW 以上。更有甚者，伴隨 NVIDIA Vera Rubin Ultra 等次世代加速計算架構的預研，未來的極限單機櫃功耗預計將直逼數百千瓦甚至 600kW。

Uptime Institute 研究總監 Daniel Bizo 在其發布的預測報告中明確指出：「由 AI 驅動的負載增長將給已經捉襟見肘的電網帶來極限壓力，超高密度基礎設施的建設成本和電力複雜性，正在將全球數位基礎設施的能效要求逼向死角。」在這種單櫃功率動輒百千瓦甚至數百千瓦的時代，沿用多年的 12V 供電系統遭遇了無法逾越的物理阻礙。傳統 12V 架構走向終結的根本原因，在於最基本的物理學定律——焦耳定律。在電路傳輸中，電能在線路上的發熱損耗 (線損) 與電流的平方成正比 ($P_{loss} = I^2 * R$)。同時，由歐姆定律可得，傳輸線纜兩端的電壓降 (IR Drop) 表達式為： $\Delta V = I * R$ 。

當我們嘗試用傳統的 12V 電壓去驅動一個功耗為 120kW 的智算機架時，主幹母線 (Busbar) 上需要承載的電流將達到極其恐怖的 10,000 安培 (A)。在工程學上，要安全傳輸 10,000A 的電流且不發生災難性的過熱，所需要的配電銅排截面積將如同成年人的大腿粗細。這不僅會佔用極其寶貴的機架內部空間，大幅阻礙液冷

管道的布線與散熱氣流，更會帶來無法承受的銅材硬體成本。此外，如此巨大的電流會導致板級和機架內部的電壓降變得難以控制，晶片端的供電穩定性將大打折扣。若是單機架功率進一步提升至 Rubin Ultra 世代的數百千瓦，12V 架構在物理上已完全不可能實現。

除了機架內部的物理約束，外部的宏觀政策與能源環境同樣在加速這場電源技術的深刻變革。隨著資料中心對電力的需求出現爆發式突增，電網的供電能力已經成為制約 AI 發展的最大瓶頸。在這一背景下，傳統供電架構中頻繁的交流 - 直流 - 交流 - 直流多級變換所產生的能源浪費 (通常每級轉換會帶來 2% 至 5% 的能源損耗) 已經無法通過合規審查。為了將能效指標 PUE (電源使用效率) 逼近 1.1 的極致水準，行業迫切需要砍掉傳統的集中式 UPS 轉換層級，將供電鏈條扁平化。這種由內而外的雙重壓力，共同催生了從 48V/54V 壓機架到 800V 高壓直流直供資料中心的技術風暴。

在超高算力密度的下半場，電力供應不再只是資料中心的後勤保障，而是直接決定了 AI 晶片能否發揮全量效能的核心瓶頸。電能傳輸中的熱效應與電路損耗，直接反作用於系統的散熱設計與空間佈局，形成了極其複雜的「電熱共生效應」。晶片功耗的快速攀升

導致供電電流急劇增加，電流增加帶來的焦耳熱，又進一步推高了伺服器內部的環境溫度；而高溫環境則會導致半導體導通電阻增大，使得導通損耗與熱量積聚呈幾何級數惡化。因此，打破傳統供電架構的層層枷鎖，從根本上重構供電物理鏈路，已成為當前智算中心硬體工程師必須面對的首要課題。

對於萬卡智算集群而言，供電系統的失誤將不再僅僅意味著額外的電費開支，更意味著系統可靠性的全面崩潰。傳統伺服器主機板在 12V 供電模式下，板級配電網絡 (PDN) 的阻抗控制相對寬鬆，但在動輒數千安培的極限電流衝擊下，PCB 走線上的微小電阻都會轉化為不可忽視的發熱源與電壓噪聲。若要確保核心 AI 晶片 (如 GPU、ASIC) 在瞬態高負載下不發生掉電或降頻，配電架構必須具備極高的響應速度與極低的傳輸阻抗。這正是推動整個產業鏈全面轉向更高電壓等級傳輸、並在靠近負載端進行高效降壓的根本動力。

在這種電力架構巨變的背景下，傳統基於矽 (Si) 基 MOSFET 的電源開關管已經逼近了其材料物理極限。矽基元件的高頻開關損耗大、導通阻抗難以降至微歐姆級別，在面對高電壓、大電流的轉換需求時顯得力不從心。這促使電源工程師將目光轉向以氮化鎵

(GaN) 和碳化矽 (SiC) 為代表的第三代寬能隙半導體材料。通過引入高頻率、高效率的寬能隙功率器件，結合扁平化的高壓配電鏈路，資料中心才能在有限的空間內實現功率密度的飛躍，徹底解決焦耳定律所帶來的線損危機與配電瓶頸。

資料中心配電架構的階梯式躍遷

正是因為 12V 供電在極端算力下的物理極限倒逼，開放運算計畫 (OCP, Open Compute Project) 在最新的 Open Rack V3 (ORv3) 標準中做出了強制性的技術修正，將機架母線電壓全面提升至 48V/54V 架構。隨著 OCP 正式將 ORv3 標準確立為全球智算中心建設的基石，48V/54V 供電架構在 2025 至 2026 年間實現了從選擇性嘗試向絕對統治的根本性轉變。NVIDIA 在其大批量出貨的 Blackwell GB200 平臺中，全面採用了基於 ORv3 標準的 48V/54V 直流母線配電方案。AMD 的 Instinct MI300/MI350 系列以及 Intel 的 Gaudi 系列加速器生態鏈，也均將 48V/54V 納為原生供電標準。

在 ORv3 架構下，輸入機櫃的高壓交流電在機架層面的集中電源機櫃 (Power Shelf) 中一次性轉換為 48V/54V 直流電，隨後通過高導電率的銅質母線排 (Busbar) 直接輸送給各

個伺服器節點。由於電壓提升了 4 倍，機架內部的傳輸電流驟降，將電流壓縮至原本的四分之一，大幅緩解了輸電線損與空間焦慮。使得機櫃能夠輕鬆承載 100kW 以上的超高功率密度，同時為液冷管道和高密度的 NVLink 網路線纜留出了至關重要的物理空間。採用 ORv3 48V/54V 方案後，伺服器機架內部的配電損耗相比傳統 12V 系統直接降低了近 90%，這為資料中心在全負載運行下省去了極大的散熱電費與減重成效。

儘管 48V/54V 供電架構在當前 AI 伺服器中已建立起穩固的統治地位，但面對算力密度近乎瘋狂的線性增長，技術專家們已經意識到，48V 依舊只是資料中心能源演進史上的一個過渡性階梯。當單個 AI 智算機架的功率正迅速從 120kW 向 240kW 甚至數百千瓦的極限跨步時，即使採用 48V 供電，單機櫃的電流依然會再次突破數千安培大關。這意味著在機架內部，厚重的配電銅排、高昂的連接器硬體以及因電流過大而產生的局部熱效應，將再度重演當年 12V 架構面臨的絕境。

為了徹底打破這一惡性循環，以 Google、Microsoft、NVIDIA) 透過 OCP 推動標準化，以及電源巨頭台達電、光寶為核心的產業聯盟，在 2025 至 2026 年間開始強力推動 800V 高壓直流 (800 VDC) 直

供電技術。該技術的變革核心在於大膽地做乘法與減法：將進入資料中心的市電經由主變壓器後，或者直接經由固態變壓器 (SST)，整流為 800V 的高壓直流電，並將其一路毫無衰減地引入智算機架內部。

從物理學角度來看，將母線電壓從 48V 提升至 800V，意味著電壓提升了 16.6 倍。在傳輸相同功率的情況下，主幹輸電線路上的電流將被壓縮至原本的十六分之一以下。這一跨越式升級，讓超大功率機架的配電系統得以脫胎換骨。

輝達執行長黃仁勳 (Jensen Huang) 在架構大會上對此發表了原話觀點：「當我們在單一機架中塞入數百顆 GPU、將功率推向幾百千瓦甚至兆瓦級時，

我們不再只是設計晶片，而是在建造一個電熱耦合的微型物理工廠。如果不將供電電壓提升至 800V 直流，機櫃內部將被厚重的銅排徹底塞滿，散熱液冷管路將毫無立錐之地。」這段話深刻點出了 800V 架構在空間與熱管理上的戰略意義。

同時，施耐德電機安全電力業務部技術長 Jim Simonelli 在 Data Center World 上也提出了明確的論斷：「當前 GPU 算力密度的激增使得傳統 400V/480V 交流配電越來越不切實際；800VDC 不是簡單的電壓提升，而是為未來超高密

表：資料中心三代供電架構 (12V vs 48V vs 800V HVDC) 電能轉換全鏈路損耗與銅材用量對比表

關鍵電力評估	傳統 12V 架構 (通用計算時代)	主流 48V/54V 裝架 (現行 AI 部署)	前沿 800V 直流架構 (次世代超智算集群)
典型機架承載功率	5 kW ~ 15 kW	50 kW ~ 120 kW	150 kW ~ 300 kW+
主幹輸電電流 (以 240kW 計)	20,000 A (理論值，工程上無法實現)	5,000 A (銅排體積巨大，發熱嚴重)	300 A (傳輸電流縮減至 48V 的 1/16)
端到端傳輸路徑效率	約 88.5%	約 93.2%	突破 98.5%
銅材用量與空間佈局	銅排用量極大，佔用大量散熱空間	銅排用量適中，連接器規格要求嚴苛	銅材用量驟降 45% 以上，物理空間大幅釋放

圖說：Delta 3MW GoCool 液對液冷卻系統 800 VDC 列間電源系統 資料來源：deltaww.com。



單櫃提供物理生存空間的底層重構。」

在 800V 高壓直流架構下，電能轉換的全鏈路損耗與銅材用量獲得了驚人的改善。由於電流大幅降低，相同功率傳輸所需的銅纜用量可減少 45% 以上。這種節省不僅減少了高昂的銅材費用，更為伺服器主機板釋放了極其寶貴的物理空間。在傳輸效率方面，800V HVDC 架構通過減少轉換環節（如淘汰傳統的多級交流變壓與 UPS 逆變過程）和極致降低線路電阻損耗，其端到端純傳輸路徑效率已成功突破 98.5% 的電氣驗證。

第三代寬能隙半導體的分工與旗艦產品

寬能隙半導體的物理演進與算力瓶頸突破

在資料中心電源架構從傳統的 12V 躍遷至 48V/54V，並進一步向 800V 高壓直流 (HVDC) 演進的過程中，傳統的矽基 (Si) 功率元件遭遇了難以逾越的物理瓶頸。矽基 MOSFET 與 IGBT 在面對極端高頻切換與極高電壓時，其開關損耗與熱聚集效應呈指數級攀升。矽材料的臨界擊穿電場僅約 0.3 MV/cm，禁帶寬度僅為 1.1 eV，這意味著在承受千伏級高壓或兆赫茲級切換時，元件內部會產生劇烈的漂移層阻抗與極大的反向恢復電荷

(Qrr)。若強行在超高密算力機架中使用矽元件，將導致電源模組體積過於龐大，佔據寶貴的運算與散熱空間。

為突破此一物理天花板，以氮化鎵 (GaN) 和碳化矽 (SiC) 為代表的第三代寬能隙半導體，在超高密算力系統中確立了明確的物理分工，並成為能源轉換的核心支柱。寬能隙材料具備更高數量級的臨界擊穿電場、更高的導熱率以及極低的導通電阻 ($R_{ds(on)}$)，使得電源設計能夠大幅提升開關頻率與功率密度，同時顯著降低熱耗散需求。

GaN 在末端負載點 (POL) 與高頻降壓中的統治力

在 48V/54V 轉換至 1V 以下超低電壓的末端負載點 (Point-of-Load, POL) 供電階段，GaN 憑藉高頻特性確立了絕對的統治地位。GaN 的禁帶寬度達 3.4 eV，電子飽和漂移速度高達 2.5×10^7 cm/s，遠超傳統矽材料。更為關鍵的是，GaN 異質結所形成的二維電子氣 (2DEG) 擁有極高的電子遷移率 ($> 2000 \text{ cm}^2/\text{V}\cdot\text{s}$)，這賦予了 GaN 元件在兆赫茲 (MHz) 級別的高頻下穩定運行的能力，且幾乎無反向恢復損耗。

開關頻率的極大提升，允許電感、電容等被動元件的體積成倍縮減，完美契合了 AI 加速卡對極致高功率密度的需

求。在傳統 100 kHz ~ 300 kHz 的矽基轉換器中，磁性元件通常佔據了電源模組 60% 以上的體積；而進入 MHz 級別的高頻領域後，電感值需求顯著下降，電感體積可縮小為原來的幾分之一，甚至可以實現板級整合與磁集成技術。

意法半導體 (ST) 針對 48V 轉換的極端挑戰，推出了基於 GaN 的高頻 LLC 諧振轉換器方案。該方案不僅支持 1MHz 以上的高頻運行，其峰值轉換效率更達到 98%，功率密度突破 2600 W/in^3 。這意味著在極其狹小的空間內，即可穩定輸出高達 12kW 的功率，徹底解決了晶片周邊空間不足的痛點。與此同時，英飛凌 (Infineon) 推出了新一代 CoolGaN 雙向極雙向開關 (Bidirectional Switch, BDS) 架構。英飛凌電源與感測系統事業部資深副總裁 Adam White 強調：「隨著機架功率邁向 120kW+ 的極限，新一代雙向 GaN 開關技術在單晶片上實現雙向控制，替代了過去複雜的分立器件，將機架配電模

圖說：意法半導體 (ST) 採用了堆疊式半橋結構的 650V 氮化鎵 (GaN) 電晶體



資料來源：st.com

組內部組件數量精簡了 40%，從根本上降低了失效率。」這項拓撲精簡不僅壓縮了空間，更將故障率降至新低。

SiC 在 800V 高壓源端與安全切換中的不可替代性

相對地，在 800V 高壓直流源端與集中整流模組中，SiC 成為了無可替代的主導者。SiC 的禁帶寬度為 3.2 eV，臨界擊穿電場強度達 3.0 MV/cm，且導熱率高達 3.5 ~ 4.9 W/m.K。這使得 SiC 不僅能承受 1200V 以上的高壓，還能將高負載下的熱量迅速傳導，極大降低了高壓大功率開關模組的能量損耗。納微半導體 (Navitas Semiconductor) 在 800V 生態

中推出了第五代 GeneSiC，採用溝槽輔助平面 (TAP) 製程。該技術融合了溝槽型與平面型的優勢，將短路耐受時間極限推升至 8 微秒，徹底解決了高壓系統易拉弧的行業難題。同時，納微推出了 10kW 800V 全磚 DC-DC 電源平臺，實現了 800V 至機架內部的極致轉換。納微半導體執行長兼技術長在技術發布會上明確指出：「在 800V 直流 AI 工廠中，GaN 與 SiC 呈現明確的物理分工——SiC 專注於源端千伏高壓整流與極速安全阻斷，GaN 專注於板級高頻降壓；兩者結合才能讓全鏈路轉換效率跨過 98.5% 的硬性門檻。」

除了能效轉換，SiC 更在 800V 高壓系統的安全防護中扮演關鍵角色。高壓直流系統一旦發生短路，產生的電弧極難熄滅。西門子 (Siemens) 與英飛凌聯合推出的 SENTRON 固

態斷路器 (SSCB)，深度整合了 1200V CoolSiC 功率模組。該元件能在微秒 (micro-s) 級別內精確切斷極限短路電流，反應速度比傳統機械式斷路器快上千倍，為高價值的 AI 硬體構築了堅實的微秒級保護防線。

第三代半導體核心參數對比

下表為針對高密算力系統的第三代半導體 (GaN/SiC) 關鍵核心參數對比。

垂直供電 (VPD) 與算電熱動態

晶片級垂直供電 (VPD) 的結構突破與物理優勢

當配電網路將高壓電能送達主機板後，最終一哩路在於如何將電流無損地送入 GPU 或 CPU 核心。在傳統的橫向供電 (Horizontal Power Delivery) 模式中，電壓調節模組 (VRM) 環繞在處理器周圍，電流需穿過 PCB 的多層導線。然而，對於功耗突破千瓦、電流高達數百安培的 AI 晶片，橫向傳輸會產生嚴重的走線電阻與壓降 (IR

圖說：Navitas GeneSiC Module



資料來源：Navitas Semiconductor

參數	氮化鎵 (GaN)	碳化矽 (SiC)	算力系統工程應用定位
典型工作電壓區間	40V ~ 100V (少部分至 650V)	650V ~ 1200V+	GaN 主攻低壓端；SiC 主攻高壓主幹。
開關頻率極限	1 MHz ~ 數 MHz	100 kHz ~ 300 kHz	GaN 實現無源元件微型化；SiC 兼顧高壓與效率。
短路與過載耐受度	較弱，依賴微秒級控制迴路保護	極強 (如 TAP 製程達 8 微秒)	SiC 在高壓故障時提供關鍵的物理緩衝時間。
核心旗艦方案代表	ST 高頻 LLC 模組、Infineon CoolGaN BDS	Navitas 10kW GeneSiC、Siemens SSCB 斷路器	構成「SiC 源端整流 + GaN 末端負載點」的黃金組合。
熱導率 (W/m*K)	約 1.3 (需依賴封裝散熱)	3.5 ~ 4.9 (本體導熱極佳)	SiC 天然適合高散熱壓力的高功率密集型模組。

Drop)，導致超過 10% 的能量以焦耳熱形式浪費，並佔用大量訊號佈線空間。

為打破此物理限制，晶片級垂直供電 (VPD) 技術應運而生，成為次世代算力供電的終極解方。VPD 技術的本質是將高頻 GaN 電源模組與微型電容直接疊層堆疊於晶圓背部，或嵌入封裝基板正下方。根據 IEEE 功率電子學會數據顯示，這種立體佈局使電流能夠以垂直微米級的距離直達矽晶片內部，徹底消除了橫向供電走線電阻。此技術不僅將輸電損耗降至微乎其微，更為晶片四周釋放了 100% 的物理空間，極大促進了高頻寬記憶體 (HBM) 的擴展與超微距互連。例如 Empower / ADI 推出的 Crescendo HD 平臺，其電壓調節器率先突破了每平方毫米 5 安培的電流密度屏障，展現了 VPD 在超高密供電中的巨大潛力。

軟體定義電源 (SDP) 與算電熱動態閉環管理

然而，單純的硬體結構重塑仍不足以應對 AI 算力極端動態的負載波動。大模型在進行預訓練或推理時，運算負載會在毫秒之間發生數倍的劇烈起伏，這種瞬態負載浪湧不僅會造成嚴重的電壓跌落，更會引發晶片內部的瞬態熱衝擊。為了平抑這種波動，算電熱動態閉環管理系統成為高階機櫃的標配。

軟體定義電源 (SDP) 技術結合數字多相控制器，構建了這個動態閉環的大腦。以德州儀器 TPS536C9 或亞德諾半導體 LTC3888-1 為代表的數字多相控制器，具備超高解析度的脈寬調變 (PWM) 與微秒級的動態響應能力。這些控制器透過 PMBus 或其他高速總線，與基板管理控制器 (BMC) 及算力晶片深度聯動。當演算法預測到即將發生的高算力爆發時，SDP 系統能夠在毫秒級別內動態微調電壓振幅 (AVS)，提前增加供電相位以承接電流浪湧；而在負載降低時，迅速關閉冗餘相位以降低開關損耗。

這種算電熱的動態協同，使得電源系統不再是單向的電力輸送通道，而是與冷卻系統、運算排程高度融合的智慧網格。透過即時監測晶片結溫、電壓降與電流波動，系統能在熱點形成前主動幹預，將整機功耗與發熱維持在最優化區間。在 VPD 技術消除了物理電阻障礙的基礎上，SDP 技術進一步在時間維度上實現了能量的精準分配，構築了超高密算力系統在極端工况下穩如泰山的能源與熱管理防線。

小結：

算力密度的持續抬升，正在把資料中心的競爭從晶片層面下探到能源實體層。從 12V 的落幕、48V/54V 的過渡，再

到 800V 高壓直流的產業推進，整套配電體系的反覆運算，本質是人類通過抬高傳輸電壓，對抗焦耳定律帶來的線損桎梏。而 GaN 與 SiC 寬能隙半導體分別承擔源端高壓變換與板級高頻降壓，解決功率變換器件的效率瓶頸；晶片級垂直供電 VPD 縮短電流傳輸路徑，軟體定義電源 SDP 完成算 - 電 - 熱動態閉環，軟硬協同共同緩解電熱共生帶來的惡性循環。

當看到，800V 高壓直流、垂直供電等前沿技術目前仍處在產業快速演進階段，還面臨器件可靠性、高壓安全保護、系統標準化、成本控制等多重現實挑戰，並不會立刻全面替代 48V/54V 現有部署。但技術方向已經清晰：未來的智算中心，供電不再是附屬後勤系統，而是與算力、散熱深度耦合的核心組成部分。

隨著下一代 AI 加速硬體持續突破單機櫃功耗上限，配電網路、寬能隙功率器件、封裝供電與動態能源管理技術持續演化。只有打通從電網到 AI 晶片內核的全鏈路能效，馴服電熱共生效應帶來的損耗與風險，萬卡、十萬卡規模的超大規模智算集群，才能夠實現高效、穩定、低成本的規模化落地。

參考資料：

■ Uptime Institute 發布的《全球資料中心調查報告》