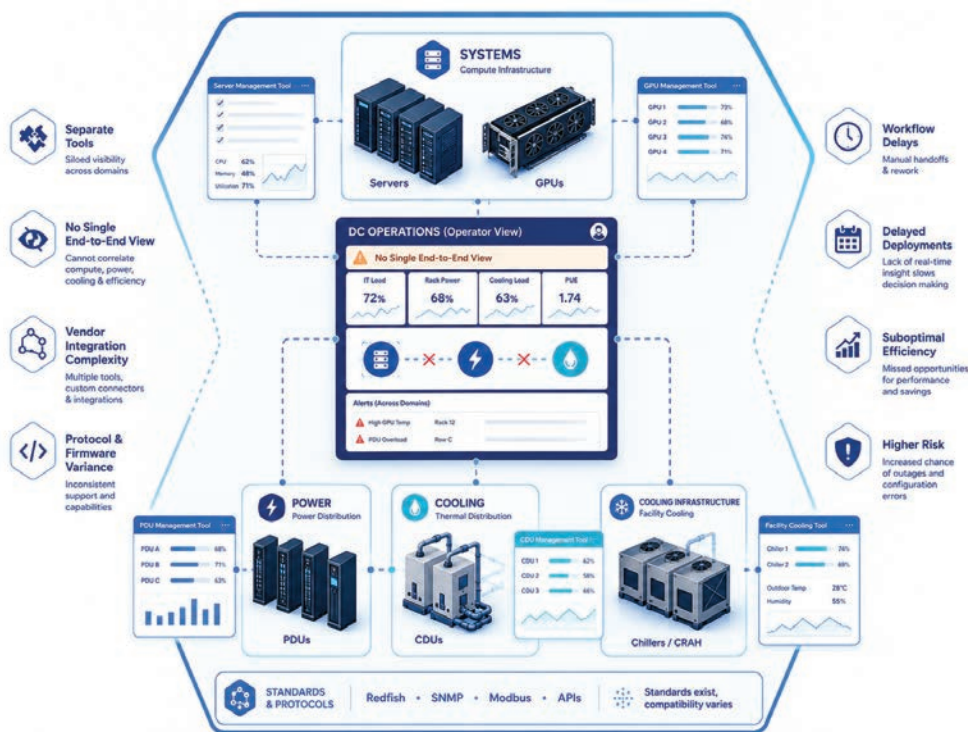


現代 AI 資料中心的 融合基礎設施

現代資料中心正超越傳統的資料中心整合管理。高密度的 AI/HPC 以及企業工作負載如今仰賴智慧電力與冷卻基礎設施。過去的传统企業運算僅依賴運算、網路和儲存，迫使資料中心營運者使用分散的應用工具來管理其作業。

如今，資料中心營運者通常透過各自獨立的工具來

管理伺服器、GPU、PDU、CDU、冷卻基礎設施以及電源供應器。這種傳統典範常常在其營運模式中造成效率低落，嚴重影響到其時間敏感的工作流程。資料中心人員可以看到個別元件，但他們可能無法獲得單一視圖，了解運算需求、機架電力消耗、液冷系統以及設施能源 PUE 如何相互影響。



挑戰在於，大多數的管理軟體原本並非設計用來處理這種程度的融合。許多平台仍將伺服器、電力與冷卻視為各自獨立的領域。同時，電力與冷卻設備來自眾多不同供應商，導致昂貴的整合挑戰，並延遲資料中心的部署。

即使存在 Redfish、SNMP、Modbus 和開放 API 等標準，實際相容性仍取決於供應商的實作、已測試的裝置支援、韌體版本的差異，以及基礎設施 OEM 之間的合作；這嚴重影響了軟體平台的整合。

三大領域

現代 AI 資料中心管理正超越僅限系統的視角。隨著系統、電力與冷卻。每個領域都有其自身角色，但
 運算密度增加，特別是 GPU 加速平台，營運模式如它們不再孤立運作。
 今仰賴三個互聯的領域架構：



三大領域

系統: 系統領域涵蓋機架內的運算與網路基礎設施，包括伺服器、CPU、GPU、儲存、交換器、韌體、連線性、拓樸以及硬體健康狀態。此領域非常重要，因為現代機架已不再是簡單的伺服器集合；它們是高密度運算區域，擁有比十年前傳統環境多得多的元件、相依性與互連。

電力: 電力領域涵蓋能源如何在機架及更廣泛的設施中被傳輸、測量、控制與保護。高密度系統會消耗更多電力，而機架層級的電力感知對於容量規劃、可用性、效率與安全運作正變得至關重要。

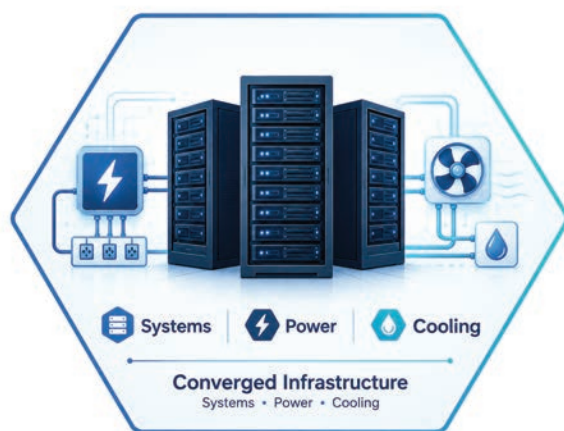
冷卻: 冷卻領域涵蓋維持現代系統安全運作所需的熱基礎設施，包括液冷系統、CDU、冷卻液流量、壓力、溫度、洩漏偵測以及設施冷卻。隨著機架變得更熱且密度更高，冷卻不再只是建築服務的問題，它已直接與系統效能和可用性緊密相連。

融合的基礎設施視圖

Supermicro 將現代資料中心基礎設施視為一個融合的生態系統，而非各自獨立的硬體孤島。隨著運算平台變得更密集且功耗更高，營運重點必須超越僅關注伺服器，擴展到支援性的電力與冷卻領域，確保這些系統保持可用、安全且高效能。

隨著這些領域變得更加互聯，軟體層也必須跟著演進。僅監控伺服器已不再足夠。現代資料中心軟體需要具備發現、監控、警報，並最終管理額外基礎設施領域的能力，例如機架電力、液冷系統、CDU、PDU 以及設施冷卻系統。

這為更整合的營運模式奠定了基礎，讓系統、電力與冷卻能夠被整體理解，而非作為獨立的環境來管理。



電力與冷卻監控不再只是關於 uptime。它們已成為安全、容量、效率與最佳化的功能性需求。在機架、PDU、CDU 以及冷卻系統層級的感知變得至關重要。平台應持續監控電力消耗、相位/負載平衡、插座層級消耗、閾值，並即時分析過載風險評估，其中包括溫度、流量、壓力、泵浦狀態、閥門

狀態、洩漏偵測以及冷卻容量等測量。透過將電力與冷卻遙測資料收集整合到單一視窗，資料中心操作人員能夠即時快速緩解營運中斷。

融合的優勢	營運理由
感測器驅動的智慧環境監控	利用來自機架感測器、PDU、CDU 以及冷卻系統感測器的資料，來偵測熱量、氣流、液冷或電力風險，在問題發生前即提前識別風險。
GPU 負載自動化	根據可用電力與冷卻資源，動態轉移、節流或排程 GPU 工作負載，讓 GPU 基礎設施能夠因應設施的限制與條件做出反應。
SLA 合規	智慧排程能在維持 SLA 承諾的同時，平衡效能、能源成本與基礎設施容量。
即時 PUE (電力使用效率)	利用即時資料顯示即時使用量與整體效率。
透過即時位置邊際定價 (Locational Marginal Pricing, LMP) 實現即時能源價格回應	幫助降低營運成本，並在尖峰時期避免對電力容量造成壓力。

解決方案

Supermicro 已超越傳統伺服器基礎設施，積極投資更廣泛的資料中心生態系統方法，以支援永續的高密度運算環境。這不僅包括運算系統，還涵蓋支援 AI、HPC 以及企業工作負載安全且高效運作所需的冷卻與電力領域。

SuperCloud Composer (SCC) 的開放分散式基礎設施管理軟體平台，為產業帶來可擴展的 Kubernetes 控制平面，將各個基礎設施領域整合成更流暢的營運儀表板。SCC 提供基礎，讓使用者能透過單一管理體驗來發現、監控、觀察與管理系統、機架電力以及冷卻基礎設施。

SuperCloud Composer 的 PDU Console 與 LCCM (Liquid Cooling Console Manager) 能夠收集重要的電力與冷卻遙測資料，用以產生警報、觸發營運事件、通知外部系統，並支援自動化工作流程。範例包括過載警告、相位不平衡警報、插座層級電力異常、冷卻效能下降事件、洩漏偵測警報、泵浦或閥門故障警報，以及容量風險通知。長期來看，這些遙測資料也能支援更高層級的協調調度決策，例如工作負載節流、工作負載放置，以及 SLA 感知的基礎設施保護。

遙測觸發	事件 \ 動作
PDU 過載風險	警示營運人員、觸發關鍵事件、透過 email/syslog/webhook 進行通知。
相位/負載不平衡	根據可用電力與冷卻資源，動態轉移、節流或排程 GPU 工作負載，讓 GPU 基礎設施能夠因應設施的限制與條件做出反應。
插座層級高電力消耗	識別受影響的機架/伺服器，並標記異常消耗。
冷卻溫度突破	產生重大熱警報，並升級通知至營運團隊。
低流量 / 壓力問題	在熱故障發生前發出冷卻效能下降事件警示。
泵浦或閥門故障	維護警報並啟動服務工作流程。
洩漏偵測	重大警報、立即升級通知，可能的工作負載保護動作。
水/冷卻液品質問題	針對液冷系統的預防性維護事件。
電力/冷卻裕度降低	容量警告以及未來工作負載放置決策。

SCC PDU Console

SCC「PDU Console」模組透過收集關鍵實體資產資訊、即時感測器資料、插座狀態與能源使用量，提供智慧 PDU 的集中式可視性與管理。透過直觀的階層式機架視圖，使用者可以快速定位資料中心內的 PDU、監控即時遙測資料、透過圖表與表格查

看歷史電力資料，並執行遠端插座電源開/關動作。這有助於團隊管理跨資料中心或遠端邊緣站點的 PDU，提升能源效率、支援 PUE 監控，並為高密度 HPC 與 AI 機架環境做出更快速的營運決策。

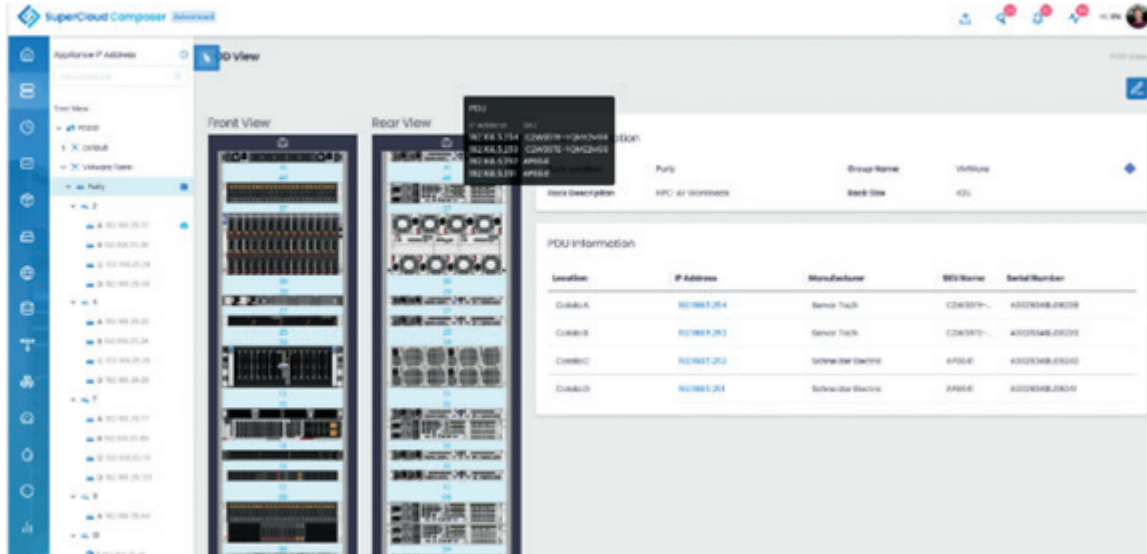


圖 1 – SCC PDU Console – 每機架顯示的 PDU

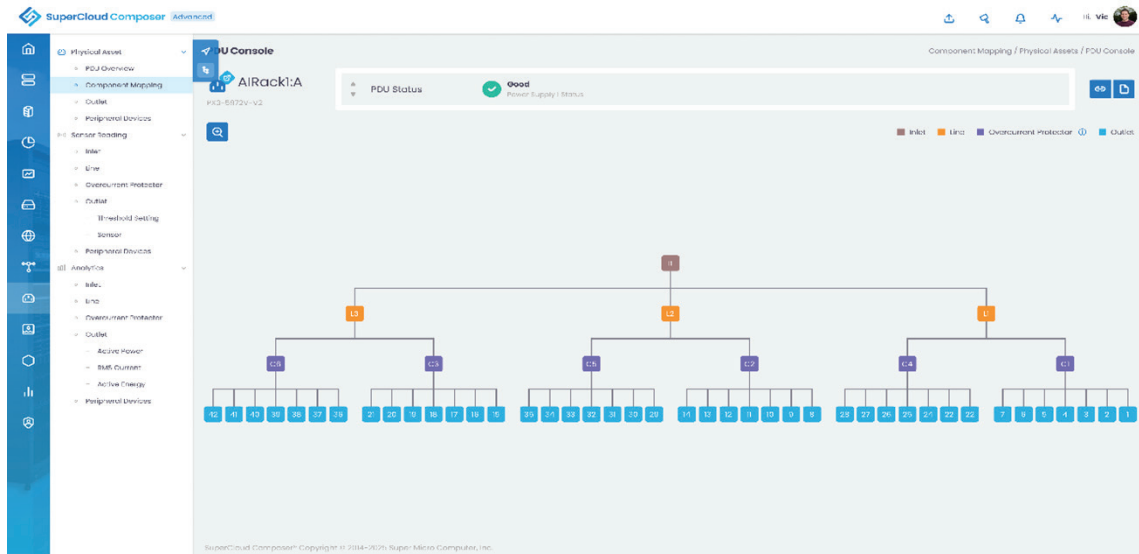


圖 2 – SCC PDU Console – 元件對應圖

- 彙總歷史 PDU 資料，並透過圖表、曲線圖與表格呈現。
- 提供隨時間推移的電力與能源使用量的細粒度可視性。
- 提供一目了然的指標、即時遙測資料，以及歷史監控功能。

SCC 冷卻基礎設施

SCC「冷卻基礎設施」模組提供跨越整個冷卻鏈的即時監控，從機架層級的液冷設備一直延伸到設施層級的冷卻塔。

對於 CDU，SCC 可以監控關鍵營運資料，例如進水與回水溫度、冷卻液流量、壓力、泵浦狀態、閥門狀態、洩漏偵測以及警報狀態。

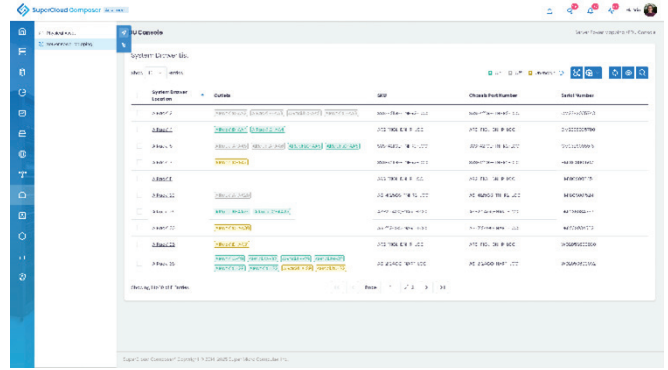


圖 3 - SCC 冷卻基礎設施，機架內 CDU



圖 4 - SCC 冷卻基礎設施，列內式 CDU



圖 5 - SCC 冷卻基礎設施，機架內 CDU



圖 6 - SCC 冷卻基礎設施，列內式 CDU

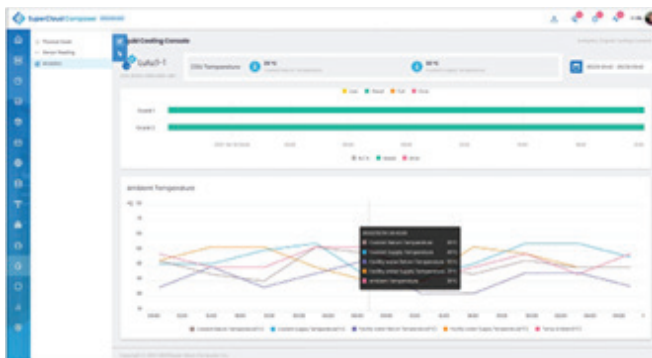


圖 7 - SCC 冷卻基礎設施，機架內 CDU

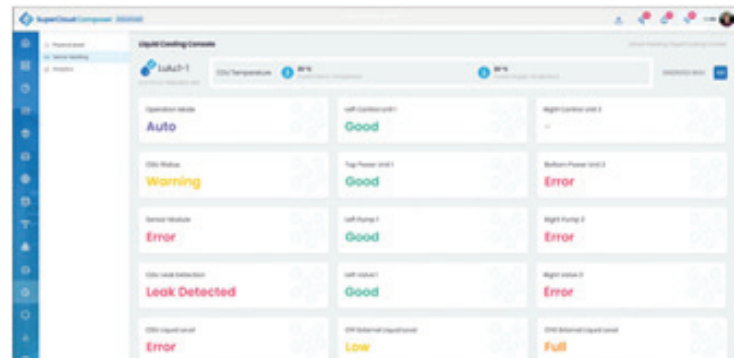


圖 8 - SCC 冷卻基礎設施，列內式 CDU

對於 RDHx 與側掛式冷卻單元，SCC 進一步提供機架與機架列層級冷卻效能的可視性，包括氣流、冷卻液溫度與流量。在設施層級，冷卻塔監控可顯示

水溫、泵浦與風扇狀態、水槽水位、水質、流量以及整體冷卻容量。這讓操作人員能夠從機架一直到整個資料中心冷卻基礎設施，獲得單一的冷卻健康狀態視圖。



圖 8 – SCC 冷卻塔

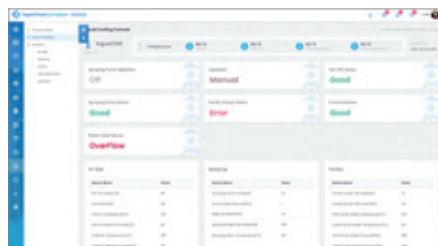


圖 9 – SCC 冷卻塔狀態



圖 10 – SCC 冷卻遙測

未來視野

融合式資料中心管理的下一個階段是協調調度，不僅將系統、電力與冷卻共同監控，還能透過政策驅動的動作進行協調，以管理以下項目：

能源來源協調調度

跨多種能源來源的協調調度，例如電網、BESS (電池儲能系統)、燃料電池、太陽能/微電網、發電機，以及未來的核能與 SMR (小型模組化反應爐) 等電力來源。

冷卻感知的運算節流 (throttling)

當冷卻容量、機架電力或設施狀況下降時，自動降

低 CPU/GPU 的電力消耗或效能，幫助防止熱故障或服務中斷。

電力與冷卻的預測性維護及風險預防

利用來自電力、冷卻、機架與系統基礎設施的遙測資料，在影響工作負載或資料中心營運之前，及早識別效能下降、容量風險或故障狀況。