

雲端 AI 架構 對開機碟的需求

企業級開機碟解決方案

在 AI 雲端環境中，開機碟是關鍵任務的核心元件之一，其決定了伺服器開機上線、重新加入叢集，以及提供工作負載的速度及可靠性。

開機碟是經常被忽略的重要基礎儲存層級，在系統初始化、受信任開機以及營運效率方面皆扮演關鍵角色，對 AI 基礎設施的效能是不可或缺的關鍵。

隨著雲端服務供應商持續擴展由 GPU、CPU 及各類 AI 加速器所組成的異質運算平台，Boot Storage 必須具備低延遲、高 IOPS、高耐用度、可靠性及安全性，才能避免成為整體 AI 基礎架構的效能瓶頸。

慧榮科技的企業級 PCIe NVMe 開機 SSD 解決方案，是 AI 開機碟的最佳選擇。結合我們的控制晶片與韌體技術，其能確保系統開機的一致性、斷電防護能力，以及整體機群的可管理性。

開機碟在雲端服務供應商 AI 架構中的關鍵角色

開機碟在雲端服務供應商 AI 架構中的關鍵角色

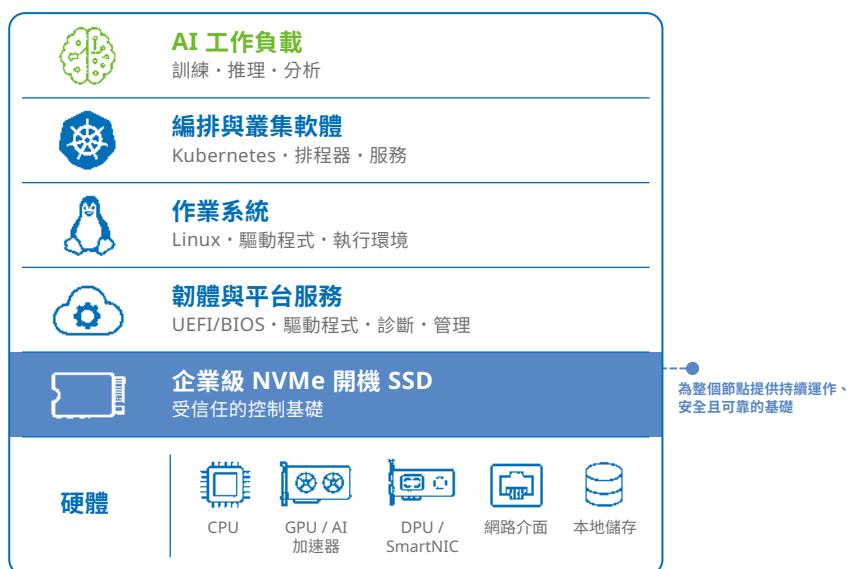


圖 1：開機碟提供受信任的控制基礎，負責初始化 AI 雲端節點，並將這些節點連接至編排與儲存服務，同時確保工作負載資料與系統磁碟區各自獨立。

在雲端 AI 環境中，開機碟是儲存架構中的關鍵基礎架構之一，讓運算執行個體能正常啟動並運作。

開機碟包含作業系統、啟動程式以及核心平台代理，用於初始化硬體並啟動 AI 工作負載所需的軟體堆疊。它們儲存關鍵組態資料，例如網路設定、身分憑證、儲存掛載點以及編排參數，讓系統在開機後能立即連接雲端服務，並存取所需的運算與資料資源。

開機碟所支援的其他功能還包括啟動編排、網路與儲存軟體堆疊，藉此建立 AI 訓練與推論所需的通訊與資源管理層。

在雲端 AI 環境中，開機碟的可靠性與效能至關重要。雲端平台必須持續快速建立與釋放運算執行個體，以因應動態變化的 AI 工作負載。因此，開機碟若能提供更快的啟動速度與更高的可靠性，便可有效降低啟動延遲，協助 AI 服務快速上線，並確保每次部署皆具備一致且可預期的運作表現。

開機碟可建立一致且可重現的系統狀態。透過標準化的開機映像與系統配置，雲端 AI 服務供應商能自動將不同規格的硬體初始化為已驗證且一致的運作環境，無須人工介入。這種標準化與抽象化機制，使平台能跨越不同實體主機快速擴展大規模運算叢集，同時維持一致的部署、管理與營運品質。

將開機碟與資料儲存分離，有助於實現架構最佳化。開機碟會針對作業系統與平台服務進行精準配置與調校，而資料碟或物件儲存則針對容量、耐用度以及特定工作負載的 I/O 存取模式進行設計。如果應用程式資料占滿或損壞開機碟，作業系統可能無法正常運作，並造成停機。將應用程式資料存放於獨立的儲存空間，可避免資源競爭，並保護平台服務的完整性。

隨著雲端 AI 架構持續發展，開機碟也必須具備更大的容量、更高的資料吞吐量、更低的存取延遲、更高的耐久性，以及更佳的功耗效率。

AI 需求對雲端服務供應商的開機碟帶來更大壓力

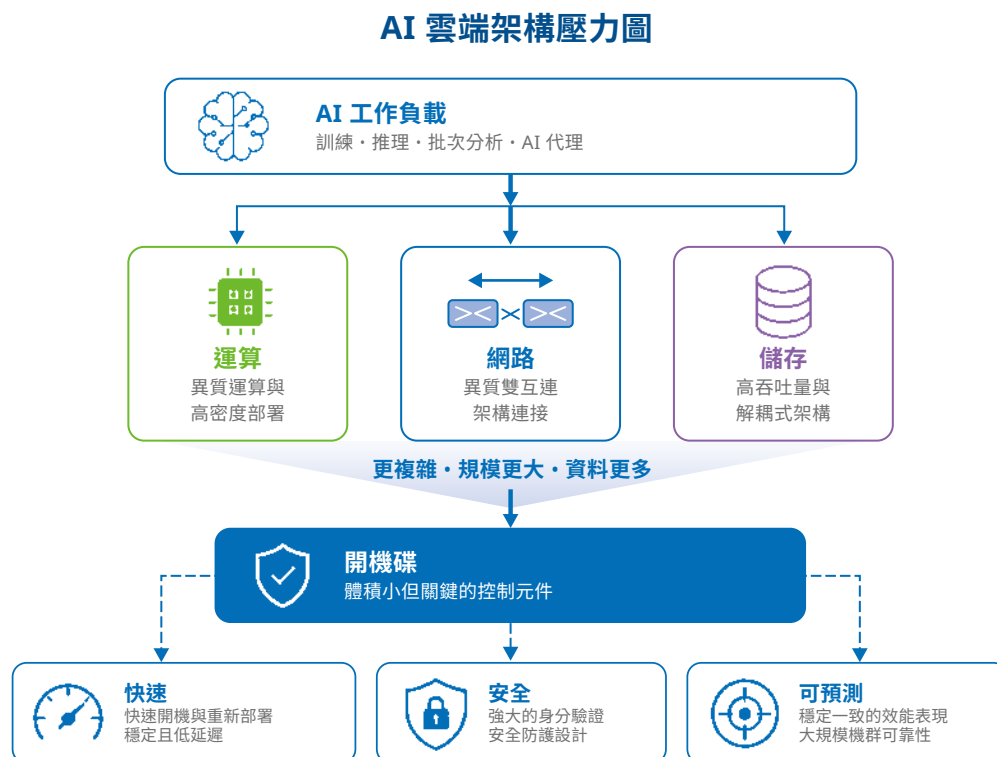


圖 2：AI 工作負載驅動了雲端基礎架構朝向異質運算、雙互連架構網路與高吞吐量儲存發展，也使具備快速、安全且可預測的開機碟，成為大規模機群營運不可或缺的關鍵元件。

為因應 AI 工作負載日益提升的需求，雲端服務供應商正重新塑造資料中心架構，全面優化運算、網路與儲存三大核心基礎設施。高密度 AI 加速器如今採用更高的記憶體頻寬、小型晶片（Chiplet）或多晶片模組（MCM）設計，以及更緊密的硬體與軟體協同設計，使其能在液冷、高功耗環境下維持高效率運作，並將擴展能力從單一伺服器延伸至整個機櫃層級。

AI 工作負載也帶來更劇烈且更難預測的資源需求波動，運算需求經常集中於配備大量 GPU 的特定區域或運算叢集。這使得容量預測與負載平衡更加困難，尤其是在 AI 模型訓練、推論服務流量激增，以及批次分析工作快速變化的情況下，更增加了資源調度的複雜度。為因應這些挑戰，雲端服務供應商正採用異質運算節點，結合 CPU、GPU、DPU 與 FPGA 等不同運算元件，依據不同 AI 訓練與推論工作負載的需求，提供最佳化的運算資源配置。

這些架構變革也促使資料中心採用雙網路架構（Dual-Fabric Networking）：一方面透過高頻寬的東西向（East-West）網路，支援 GPU 與 GPU 之間的大量資料傳輸；另一方面則透過南北向（North-South）網路，處理對外連線及一般系統流量。同時，儲存系統也必須具備高資料吞吐量，以持續為 AI 加速器提供資料，避免運算資源閒置；資源調度器（Scheduler）則需能根據不同工作負載的需求，將每項作業分配至最合適的運算資源。此外，分散於不同地理位置的資料中心，正逐漸被視為單一的 AI Factory。透過 Scale-across 網路互連技術，可將低延遲連線能力延伸至從數百公尺到數百公里的距離，讓跨資料中心的 AI 運算資源得以協同運作，如同一座大型 AI 基礎設施。

隨著 Scale-out 與 Scale-across 架構逐漸融合，越來越多 AI 基礎設施開始採用相同的硬體與軟體平台，讓網路資源得以在資料中心內（Intra-Data Center）與跨資料中心（Inter-Data Center）之間彈性共享與調度。這種統一架構可支援跨多個資料中心的大規模 GPU 叢集，以及分散式 AI 推論等應用，包括部署於邊緣至雲端環境的代理式 AI 工作負載。

儲存架構同樣正因應 AI 工作負載而持續演進。雲端服務供應商廣泛採用 NVMe SSD，將儲存與運算分離，並將資料直接傳輸至 GPU 記憶體等技術，最大化資料吞吐量並降低存取延遲。此外，Checkpoint 機制已逐漸成為 AI 基礎架構的重要效能瓶頸。隨著 AI 模型規模持續擴大，每次 Checkpoint 所需儲存的資料量已達數 TB 等級；同時，大量平行工作負載也大幅增加了寫入壓力，而日益嚴格的服務等級要求（SLOs）則要求更快速地完成快照建立與系統還原，使儲存系統面臨更大的效能挑戰。

最後，資料生命週期管理正逐漸演進為以雲端原生架構為核心的持續性資料流程（Pipeline），涵蓋資料蒐集、清理、治理、重複利用，以及模型快速迭代與回饋等各個環節。這種架構要求資料能即時共享，並具備更完善的可追溯性與更嚴密的安全防護，確保 AI 系統始終使用最新且可信賴的資料。在此趨勢下，開機碟的設計也逐漸朝向精簡化、高速化，以及無狀態（Stateless）架構發展。

開機碟面臨雲端 AI 技術需求的新挑戰

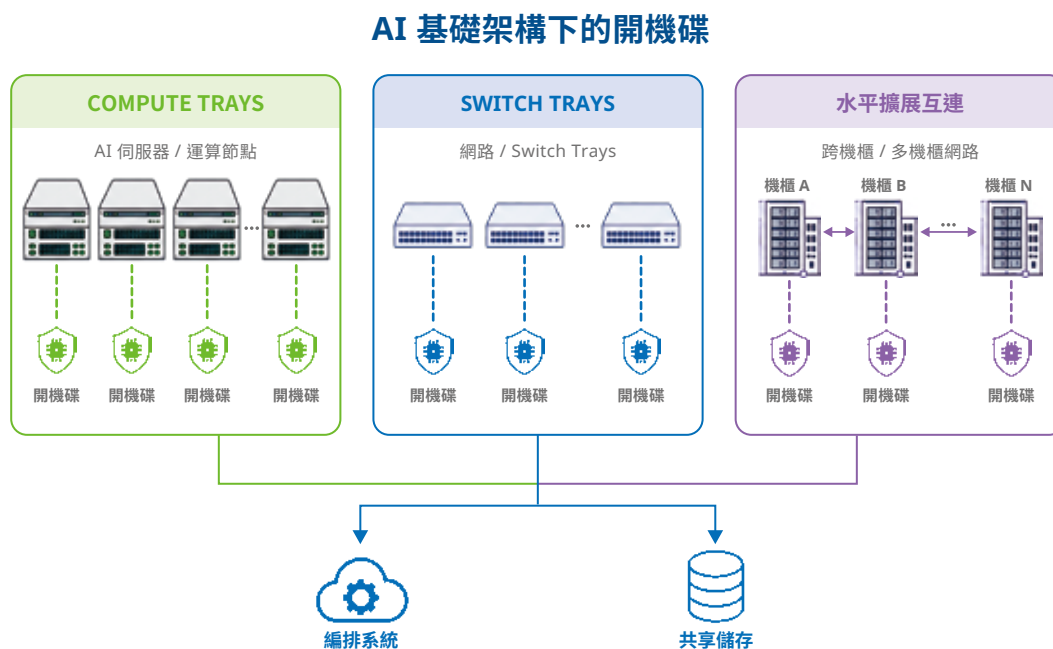


圖 3：無論是在 Compute Trays、Switch Trays，還是水平擴展互連架構中，開機碟都扮演精簡而關鍵的控制中樞，負責初始化各項裝置、建立受信任的系統組態，並將基礎架構連接至編排平台與共享儲存服務。

在以 AI 導向的雲端服務供應商環境中，開機碟的角色雖然有限，但卻扮演著關鍵的角色：其用於承載作業系統與必要的系統軟體，用來初始化每個節點，並使其能夠執行 AI 工作負載。開機碟和用於處理資料集、模型檢查點及訓練流量的高效能儲存分層是彼此獨立的。

實務上，開機碟通常只包含一個精簡的作業系統映像，提供系統啟動所需最基本的環境。節點可從本機小型映像啟動，或將作業系統載入記憶體執行，隨後連接至網路儲存設備，例如 NVMe-over-Fabrics 或其他共享儲存系統，以承載實際工作負載。

這類環境通常涵蓋三種場景：Compute Trays、Switch Trays，以及水平擴展 (scale-out) 互連。這些設備具備相同的基本需求，但實務上的運作重點各有不同，端視該裝置是負責啟動 GPU 伺服器、互連架構交換器，或大型分散式伺服器機群而有所不同。

對於 Compute Tray 而言，通常會配置一顆專用的 NVMe Boot SSD（多採用 M.2 或 E1.S 規格），用於儲存作業系統映像，使運算節點能夠在不依賴外部儲存系統的情況下獨立完成啟動。完成開機後，Compute Tray 便會加入叢集管理平台，並將系統初始化至已驗證且一致的軟體狀態，再開始參與分散式 AI 運算。

對於 Switch Tray 而言，本地端開機碟負責儲存 Switch 作業系統、組態設定檔及管理軟體。這項設計至關重要，因為網路 Fabric 必須先穩定完成初始化，GPU 伺服器才能建立低延遲的通訊環境。開機碟不僅可協助 Switch 快速恢復運作，也能支援網路架構的平順擴充。

在水平擴展互連架構中，開機碟扮演控制核心的角色，而非單純的儲存引擎。其負責載入作業系統、容器執行環境、驅動程式與加入網路架構所需的網路組態，而大部分資料與模型參數則保留在可獨立擴展的共享儲存系統中。

開機碟同時也是實現大規模節點管理與安全防護的重要基礎。標準化系統映像、安全開機、快速重新部署、本地端認證資訊、韌體管理，以及系統組態基準等關鍵功能，都仰賴穩定且可預測的本地端儲存。因此，當節點發生故障或需要重新配置用途時，可快速完成系統重置與重新部署，而無須影響整體 AI 基礎架構的正常運作。

總而言之，對於雲端服務供應商而言，開機碟的設計重點應放在 IOPS、延遲、可靠性與可管理性，而非單純追求儲存容量；同時，更應將開機碟視為 AI 叢集信任根（Root of Trust）的一部分，作為整體安全架構的重要基礎。上述應用情境充分說明，在高度動態的 AI 雲端環境中，SSD 是開機碟的最佳選擇。

為何 SSD 是 AI 雲端環境中的開機碟最佳選擇

為何 AI 雲端開機碟要採用 SSD？

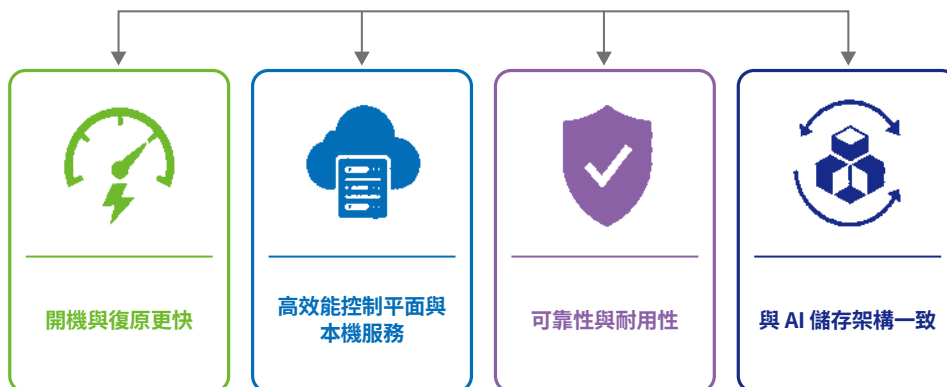


圖 4：相較於機械式 HDD，企業級 PCIe NVMe SSD 可提供更低延遲、更高 IOPS、更佳耐用性與穩定行為，以支援 AI 雲端基礎架構在大規模環境下的開機、復原與管理需求。

AI 基礎架構需要超低延遲、高 IOPS 與持續穩定的效能，而這些都是傳統機械式硬碟無法滿足的需求。開機碟必須能支援數千台伺服器同時開機、頻繁的作業系統重新映像（OS Re-imaging），以及快速的故障復原，同時避免成為整體系統的效能瓶頸。

SSD 具備多項關鍵特性，可支援典型開機碟功能：

- **開機與復原更快：**開機碟必須在發生故障後迅速使伺服器恢復上線，以降低叢集停機時間並維持 GPU 使用率，也需能在大規模環境下，提供可靠的復原能力。因為即使是微小延遲或重複重建，也可能導致作業瓶頸。
- **高效能控制平面與本機服務：**開機碟必須提供超低延遲、高 IOPS 的儲存效能，才能在 AI 雲端環境中快速且可靠地啟動編排、身分驗證、監控及其他關鍵系統服務。採用本地端 SSD 作為開機碟，可大幅降低存取延遲，有效避免系統啟動過程中的效能瓶頸。
- **可靠性與耐用性：**開機碟一旦發生故障，可能對整體基礎架構造成立即性的服務中斷。由於這些系統需要頻繁寫入，且通常具備長期運作週期，因此儲存裝置必須能承受持續性的工作負載，同時確保資料完整性與系統可用性。
- **符合 AI 儲存資源配置需求：**符合 AI 儲存資源配置需求：若 I/O 存取模式與儲存資源配置不一致，容易造成效能下降與不可預測的延遲。SSD 能提供可預測且一致的效能、穩定的功耗表現，以及優異的資料完整性，確保作業系統載入、韌體執行及持續性遙測等關鍵作業能穩定可靠地運作。

SSD 提供 GPU 基礎架構所需的高速效能、一致性與穩定性，讓系統能夠快速啟動、迅速復原，並在大規模部署環境下維持不中斷的穩定運作。

慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案專為 AI 基礎架構打造

慧榮科技開機 SSD 技術堆疊



圖 5：慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案，結合控制晶片以及 IntelligentSeries™ 功能，提供快速、可預測、安全且具備電源異常保護能力的開機儲存，滿足大規模 AI 雲端基礎架構的部署需求。

在 AI 雲端環境中，開機儲存已不再是次要元件，而是關鍵控制節點，直接影響系統運行、資源使用率與整體叢集效率。

慧榮科技結合專有控制晶片與韌體技術，兼顧高效率、一致性與安全性，全面滿足 AI 伺服器對開機碟的需求，同時降低大規模部署環境中的啟動風險與整體營運成本。其核心優勢不僅在於卓越的效能，更在於能確保數千台 AI 伺服器快速且一致地完成啟動、維持系統安全，並在長期運行中有效降低功耗。

- **快速且可預測的開機效能：**慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案結合專為開機儲存設計的儲存架構、具備可預測性的韌體運作機制、穩定的電源管理，以及完善的資料完整性保護，確保作業系統能可靠地完成載入與初始化。
- **AI 節點持續運作的可靠性：**慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案，以穩定的開機行為、斷電保護及高度的資料完整性為核心，即使在頻繁的日誌寫入及持續運作的環境下，仍能確保作業系統載入與韌體執行的穩定性。
- **斷電與惡劣環境耐受力：**電源異常與嚴苛環境韌性：慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案結合先進的電源異常保護技術，透過硬體與韌體協同運作，在突發斷電時安全地將傳輸中的資料寫入非揮發性快閃記憶體，並確保映射表的完整性，有效維持系統資料一致性與可靠運作。
- **安全防護與可信任開機：**慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案結合硬體加密技術與整合於控制晶片中的安全機制，包括韌體驗證保護、安全開機，以及符合 TCG Opal 等安全標準，確保開機碟不受未經授權的存取或惡意竄改，建立可信任的系統啟動環境。

慧榮科技專有 IntelligentSeries™ 技術透過多種機制進一步提升開機 SSD 解決方案效能，包括

- IntelligentLog™ 可全面記錄 SSD 健康狀態，讓系統開發者了解其當前狀況並採取適當措施，並記錄關鍵事件日誌，以提升除錯速度與效率。
- IntelligentScan™ 透過自動掃描與修復機制，確保安全區域資料讀寫正確性，同時延長資料保存時間並提升 SSD 壽命。
- IntelligentThermal™ 可監控開機碟溫度，並在超過一定臨界點時啟動散熱機制。有效地在主機控制溫度管理與硬碟控制溫度管理間做出選擇，以控制溫度與防止在高溫下造成 SSD 損壞。
- IntelligentShield™ 可確保所有正在傳輸的資料安全地儲存在指定的安全區域，以防範因電力中斷造成的資料遺失。當偵測到電源供應不穩定時，該開機碟將立即啟動此機制。
- IntelligentFlush™ 能以定速進行資料轉移，讓資料能快速地移轉到安全的儲存區塊。

- IntelligentGuard™ 技術透過強化開機碟韌體的身分驗證與保護機制，確保開機的安全性。
- IntelligentImage™ 為嵌入式系統提供預先載入機制，可在封裝前將資料預先載入指定的安全區塊，能保護資料免受像是回錫時的高溫等製造風險的損害，並大幅提升產線效能。
- IntelligentZone™ 支援多命名空間設計，讓 SSD 可根據不同系統設計者的需求自訂。SSD 被區分為不同的儲存區塊，資料則依據設計放置於適當的區塊內。這能確保資料完整性，以及有效利用 SSD 儲存空間，以獲得最大的經濟利益。

開機碟：AI 雲端環境中的關鍵控制節點

開機碟雖然容量不大，卻是影響 AI 雲端基礎架構效能、可靠性與安全性的關鍵元件。在大規模 AI 雲端環境中，任何停機都將直接影響資源利用率、營運成本與服務品質，因此開機碟的重要性遠超乎其容量大小。它已不再只是系統的輔助儲存裝置，而是支撐整個 AI 基礎架構穩定運作的關鍵控制節點。

隨著雲端服務供應商持續擴大分散式 GPU 叢集的部署規模，快速完成系統啟動、在故障後可預測且一致地恢復運作，以及維持可信任的系統狀態，已不再只是系統背景功能，而是 AI 基礎架構不可或缺的核心營運能力。PCIe NVMe SSD 憑藉低延遲、高 IOPS、高耐久性及優異的能源效率，能充分滿足這些需求，協助 AI 平台維持高可用性與營運韌性。

慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案進一步整合我們專有的 IntelligentSeries™ 技術，結合可預測且一致的運作機制、電源異常保護、安全開機功能，以及智慧化管理能力，全面提升開機碟的可靠性、安全性與可管理性，滿足 AI 雲端基礎架構大規模部署的需求。

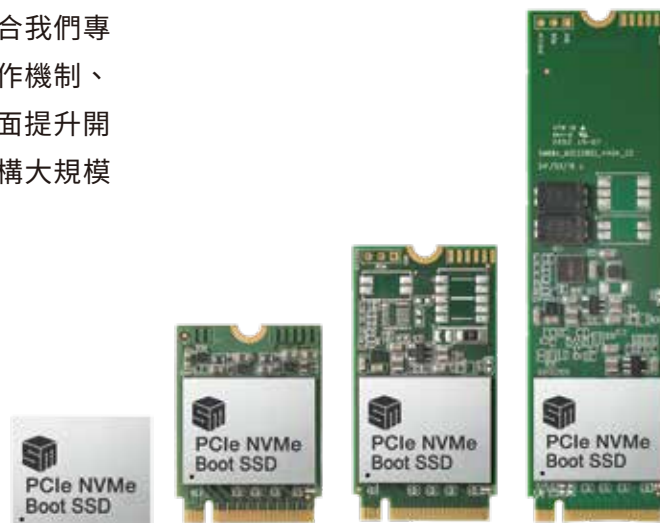


圖 6：慧榮科技企業級 PCIe NVMe 開機 SSD 解決方案支援各種尺寸規格，包括 BGA 與 M.2 設計。

更多資訊請造訪 www.siliconmotion.com