

慈濟學校財團法人慈濟大學採購規格書

採購標的名稱：AI服務效能管理系統 林信漳 7/6

採購單號：114010354

項次	品名及規格	數量	單位
一	- 具管理及執行 Docker Containers 提供 AI 運算服務。可利用 Kubernetes 進行 Docker Containers 集群管理與部署機制，至少能提供 GPU 資源管理(含分配、收回及下線維護)、AI 運算服務排序及儲存系統掛載等功能。 - 提供專案申請審核啟用與取消功能，於專案審核啟用時，可提供 GPU 條件設定，再設定範圍內的申請單可以不需要管理者審核。 - 後台管理員可增修刪專案內容與成員，提供勾選專案後批次刪除專案功能。可對專案設定可使用之特定 GPU 型號，設定完成後，該專案使用者只能看到限用之 GPU 型號。 - 需提供多管理者設定功能，並可以有使用者與管理員在平台上的操作記錄。管理員可於專案中進行資源用途調整與 GPU/CPU/RAM 額度設定。 - 提供從 NVIDIA GPU Cloud/Docker Hub 拉取鏡像功能。 - 提供管理者選擇 SSH 登入容器方式(密碼/金鑰)。 - 提供管理者圖形界面進行 GPU 資源分配管理，可兩種模式，獨佔式：支援 1 人使用多張 GPU 卡。分享式：支援多人使用 1 張 GPU 卡。 - 依 GPU 型號支援細粒度記憶體切割，最小顆粒度可達 1GB 並支援多位使用者共享同一張 GPU (涵蓋 NVIDIA 及 AMD)。 - GPU 獨立切割，可依 GPU 型號等資源設定單張 GPU 切割分享記憶體並定義該系統主機之 CPU 與記憶體數量供使用者選擇。 - 可對 GPU 資源設定每小時費用計算機制，並可為各專案設定使用額度。當專案額度不足時，於建立容器時提出警訊。 - 報表查詢，提供集中式管理圖形介面可監控各 GPU 伺服器硬體使用率(CPU、RAM、GPU、GPU RAM 等)。 - 支援 MinIO 物件儲存檔案系統。 - 可以同時管理 NVIDIA 和 AMD 的 GPU 資源。 - 依需求可採高可用性 (HA) 設計以避免單點失效。 - 快速容器部署，平台需提供使用者進行鏡像服務部署功能，同時部署之服務可以不必使用 GPU。 - 支援 Nvidia 完整深度學習開發環境，包含 Ubuntu 22.04 LTS(含)以上、NVIDIA Driver、Nvidia CUDA 11.7~12.8、cuDNN、NVIDIA Docker、jupyter。 - 支持 Elastic Training:在訓練任務中，支持根據容器集群計算資源的變化，動態擴容或縮減工作容器，提高資源利用效率。 - 平台商業軟體為一年買斷制授權 - 提供支援 NVIDIA GPU H200 *4、RTX6000 *1 之原廠授權軟體。 - 提供軟體安裝與教育訓練。 以上所有規格得同品級或提供優規。 (※允許投標廠商於投標文件內預先提出同等品之廠牌、價格及功能、效益、標準或特性等相關資料，以供審查。)	1	式 <u>林信漳</u> <u>7/6</u>

本採購案是否有後續擴充 ☐ 是，擴充期間_____、金額或數量_____。
☒ 否。

交貨期限：☒ 簽約日起 90 日內。☐ 民國_____年_____月_____日內。
☐ 合約生效日(決標日)起_____日內。☐ 收到信用狀日起_____日內。

保固期限：自驗收合格之日起保固 1 年。

請購單位：電算中心 林信漳 7/6 請購單位主管簽章：林信漳

請購人簽章：林信漳 聯絡電話：22349

本案規格內容係依政府採購法第 26 條訂定。

※政府採購法第 26 條：

機關辦理公告金額以上之採購，應依功能或效益訂定招標文件。其有國際標準或國家標準者，應從其規定。

機關所擬定、採用或適用之技術規格，其所標示之擬採購產品或服務之特性，諸如品質、性能、安全、尺寸、符號、術語、包裝、標誌及標示或生產程序、方法及評估之程序，在目的及效果上均不得限制競爭。

招標文件不得要求或提及特定之商標或商名、專利、設計或型式、特定來源地、生產者或供應者。但無法以精確之方式說明招標要求，而已在招標文件內註明諸如「或同等品」字樣者，不在此限。