



## 美光發布速度、容量傲視業界的高頻寬記憶體(HBM) 推動生成式AI創新

July 27, 2023 at 10:00 AM CST

### 採用1β 製程節點，率先業界推出第二代8層堆疊24GB HBM3

#### 實現 1.2TB/s 以上頻寬及卓越功耗效率

2023年7月26日，愛達荷州博伊西 — 美光科技 ( Nasdaq: MU ) 宣布推出業界首款第二代 8 層堆疊 ( 8-High ) 24GB HBM3，並於今日開始送樣。此產品頻寬達 1.2TB/s 以上，每腳位傳輸速率超過 9.2Gb/s，較目前市面上的 HBM3 解決方案高出 50%。此外，美光第二代 HBM3 的每瓦效能較前幾代產品提升 2.5 倍，刷新 AI 數據中心的關鍵性能、容量及功耗指標。這些改進縮短了業界訓練 GPT-4 等大型語言模型及其更高階版本所需時間，幫助 AI 推論所需的基礎硬體架構發揮最佳功效，並提供卓越的總體擁有成本 ( TCO )。

奠基於先進的 1β DRAM 製程節點，美光領先業界的 HBM 解決方案能在業界標準封裝尺寸中，將24GB的晶粒組裝為8層高度的立方體。美光 12 層堆疊 ( 12-High ) 36GB HBM3 亦將於 2024 年第一季開始送樣。透過此堆疊高度，美光可提供較競品解決方案高出 50% 的容量。美光第二代 HBM3 性能功耗比和每腳位傳輸速率的改善，對因應當前 AI 資料中心的高功耗需求而言至關重要。美光提供較現行 HBM3 解決方案多出一倍的直通矽晶穿孔 ( TSV ) 數量，又以五倍金屬密度減少熱阻抗，並採用高能效資料路徑設計，實現功耗的改善。

身為 2.5D/3D 堆疊和先進封裝技術的領導者，美光很榮幸能成為台積電 3DFabric 聯盟的合作夥伴，攜手塑造半導體和系統創新的未來。在第二代 HBM3 產品的研發中，我們與台積電的合作為順利導入和整合 AI 及高效能運算設計應用的運算系統奠定了堅實基礎。台積電已收到美光第二代 HBM3 樣品，並正與我們密切合作，以進行進一步的評估和測試，從而助力客戶的次世代高效能運算應用創新。

在生成式AI方面，美光第二代 HBM3 亦能符合多模態數兆參數AI模型所需。每組堆疊24GB的容量、大於 9.2Gbps 的每腳位傳輸速率除了能減少大型語言模型 30% 以上的訓練時間，進而降低總體擁有成本外，也能大幅提升每日查詢次數，讓訓練完成的模型在使用時更有效率。不僅如此，美光第二代 HBM3 傲視業界的每瓦效能也替現代 AI 資料中心扎實省下營運費用：若以裝設1000萬個GPU計算，每組堆疊節能5瓦，預計將能在五年內省下多達 5.5 億美元的營運成本。

美光副總裁暨運算產品事業群總經理 Praveen Vaidyanathan 表示：「美光第二代 HBM3 的研發重點，在於為客戶及業界提供卓越的 AI 及高效能運算解決方案。我們所考慮的一項重要標準，就是第二代 HBM3 產品能否輕鬆與客戶平臺整合。對此，記憶體內建的自我測試 ( MBIST ) 不僅可徹底程式化，還能以規格中的最高每腳位傳輸速率進行測試，有助於改善與客戶的測試能力，提升協作效率，並縮短上市時間。」

NVIDIA 超大規模與高效能運算副總裁 Ian Buck 表示：「生成式 AI 的核心是加速運算，而這種運算則受益於具有高頻寬及能源效率的 HBM。我們與美光在廣泛產品領域的合作已行之有年，也渴望能在第二代 HBM3 上持續合作，以推動 AI 創新。」

美光運用其全球工程組織開發這項突破性的產品，包含於美國進行設計和製程研發、於日本進行記憶體製造，及於台灣進行先進封裝。繼採用 [1α 製程節點 24Gb 單體式 DRAM 晶粒的 96GB DDR5 模組](#) 問世，以支援高容量需求的伺服器解決方案後，美光今日又宣布推出基於 1β 製程節點 24Gb 晶粒的 24Gb HBM3 產品，為美光業界技術領先地位立下新的里程碑。美光也計劃於 2024 上半年推出採用 1β 製程節點 32Gb 單體式 DRAM 晶粒的 128GB DDR5 模組。這些產品展示了美光在 AI 伺服器領域的領先技術創新。

歡迎造訪 [micron.com/HBM3](https://micron.com/HBM3)，了解更多產品資訊。

## 資源

- [HBM3 產品頁面](#)
- [Six Five Insider Podcast 節目採訪美光產品管理資深總監 Girish Cherusery](#)

## 關於美光科技

我們是創新記憶體和儲存空間解決方案的業界領導者，並且正在改變世界使用資訊的方式，豐富 *所有人* 的生活樣貌。美光持續關注於客戶、技術領先、卓越的製造與營運，透過美光 ( Micron® ) 和 Crucial® 品牌提供高性能 DRAM、NAND 和 NOR 記憶體以及儲存的豐富產品組合。每一天，我們人員提出的創新推動了數據經濟、人工智慧和 5G 應用程式的進步，激發各種機會——從資料中心到智慧邊緣以及客戶端和行動裝置使用者體驗。欲進一步瞭解 Micron Technology, Inc. ( Nasdaq : MU )，請瀏覽 [micron.com](https://micron.com)。

© 2023 Micron Technology, Inc. 版權所有。文中資訊、產品和 / 或規格若有變動，恕不另行通知。美光 ( Micron )、美光標誌及其他美光商標皆為 Micron Technology, Inc. 財產。其他商標之財產權則歸其各自擁有者所有。

## 美光媒體聯絡人

Kelly Sasso

美光科技公司

+1 (208) 340-2410

[ksasso@micron.com](mailto:ksasso@micron.com)