



慧荣科技参展GTC2026, 多层次存储方案赋能NVIDIA AI生态系统

March 17, 2026

慧荣科技 (NasdaqGS: SIMO) 作为设计和推广固态存储设备NAND闪存主控芯片的全球领先的企业之一, 今日宣布将在NVIDIA GTC 2026的3015号展台展示丰富的差异化企业级SSD主控芯片和PCIe NVMe BGA启动SSD产品组合, 专为满足NVIDIA AI生态系统不断演进的需求而打造。



随着AI模型规模的不断扩大, 推理架构正在从HBM和系统DRAM扩展到高性能NAND存储层, 正如NVIDIA的ICMS计划所体现的那样。在这种新架构中, 基于NAND的存储成为需要确定性延迟和差异化服务质量的性能关键层。

慧荣科技提供垂直整合的存储解决方案, 涵盖先进的SSD主控芯片设计、完整的固件开发, 以及符合E1.S、E1.L、E3.S、E3.L和U.2等主流企业级外形规格的紧凑型参考设计套件(RDK), 以满足AI服务器部署需求。慧荣科技还提供在AI系统中部署的具备企业级应用所需的高耐用性与长期运行稳定性的企业级PCIe NVMe BGA启动SSD。

First PCIe Gen6 Enterprise SSD Controller for Near-GPU High-performance Storage



Performance

- Sequential: 28 GB/s
- Random: 7M IOPS (4KB)

Features

- PCIe Gen6
- SBCV / NPIV
- NVMe
- IDE
- NVMe 2.0+
- QOP NVMe SSD Spec 2.3
- WRW/SWRB (Extra) WPC
- Performance™
- ESE Data Protection
- SMBIO Monitoring
- Crypto Functions: CNRA 2.0, PQC
- Data Security: TCG Opal
- Root of Trust: Caliptra 2.1
- 1st PCIe Gen6 SSD with Caliptra 2.1
- Secure Boot & Attestation

High Performance
Low Power

- TCM: 4mm
- Power: 5-13.5W (max/typical)
- Support: S.M.A. Interface
- Max: 4x 100Gbps Capacity

AIC x SMI

AIC F2032 E3.S/L NVMe JBOD Series
Powered by Silicon Motion
MontTitan™ Gen5 SSD Solutions



Model: F2032-G6
Storage: 32 x PCIe Gen5 Dual port E3.S/L

SM8366

High Performance: 16-Chan
Seq. R/W: 14+/11 GB/s
Ran. I/O: 10K/450K IOPS



EXASCEND
Inspiration to Innovation
www.exascend.com



慧荣科技总经理兼首席执行官苟嘉章(Wallace Kou)表示：“NAND 存储正成为AI基础设施中的战略层。随着推理上下文的不断扩展，存储不仅要提供高容量，还要提供确定性延迟、持续带宽和可预测的服务质量。主控芯片技术是充分发挥NAND闪存的性能潜力的关键。慧荣科技先进的主控芯片架构旨在满足NVIDIA ICMS等创新带来的需求，使NAND存储能够在AI内存层级中作为主动的、性能驱动的层级运行。”

VAST Data战略运营副总裁Avery Pham表示：“AI 基础设施需要智能数据管理软件与高性能存储硬件之间的紧密集成。VAST的统一数据平台将结构化和非结构化数据整合到一个为AI工作负载优化的单一可扩展架构中。灵活、高性能的SSD主控芯片技术，例如由慧荣科技开发的技术，对于实现跨层级激活大规模AI数据集所需的性能一致性和可扩展性至关重要。”

在GTC 2026上，慧荣科技将展示在生态系统多个层级中为AI基础设施优化的存储解决方案，包括：

用于KV缓存扩展的NVIDIA ICMS–MonTitan® SM8466/SM8366

NVIDIA推理上下文内存存储(ICMS)计划突显了将KV缓存从传统内存扩展到对延迟敏感的NAND存储层这一不断演变的需求。此类架构需要确定性的服务质量、严格的尾部延迟控制、强大的工作负载隔离、多租户支持以及持续的读取性能，以保持高效的GPU利用率。

慧荣科技的MonTitan®企业级SSD主控芯片的架构设计正是为了契合这些新兴需求。利用慧荣科技获得专利的PerformaShape™技术，主控芯片能够动态优化工作负载行为，增强并发工作负载之间的隔离，并管理延迟分布，从而在AI推理工作负载下性能表现稳定，具备良好的可控性。**近GPU高性能存储–MonTitan® SM8466/SM8366**

对于近GPU加速层，存储必须在混合和动态工作负载下保持高带宽和IOPS。MonTitan®企业级SSD主控芯片提供低延迟、强并行性和可预测的性能，以支持GPU加速环境中的推理处理、数据暂存和模型更新。



High Performance Low Power

- 15WDC @ 30m
- 100% 8 CE/CM (800000/24)
- Support SCA interface
- Up to 3.2TB Capacity

Performance

- Sequential: 28 GB/s
- Random: 7M IOPS

Features

- PCIe Gen6
- SIOV / MPF
- AEC
- IDE
- NVMe 2.0+
- ORP NVMe SSD Spec 2.0
- 4KQ/4KRB (CoCo) LBPC
- PermaShape™
- E2E Data Protection
- SMART Monitoring
- Crypto Functions: CNSA 2.0
- Data Security: TCG Opal
- Root of Trust: Caliptra 2.1
- 1st PCIe Gen6 SSD with Caliptra
- Secure Boot & Attestation

AIC x VAST DATA

Efficient, Affordable Platform
Powered by Silicon Motion
MonTitan™ Gen5 SSD Solutions

Model
Ceres V2 Dbox

Storage
22 x 122TB PCIe Gen5 QLC E1.L
8 x 3.2TB SLC Module U.2

SM8366

High Performance: 16-Channel with 8CEs
Seq. R/W: 14+/11 GB/s
Ran. R/W: 3,500K/450K IOPS
Ultra High Capacities: 128/256TB QLC SSDs

AIC | **SiliconMotion** | **MonTitan™**



计算优化存储—SM8466、SM8366和SM8308 在AI计算节点中, 存储需求强调吞吐量一致性、耐用性和系统级效率。慧荣科技的MonTitan®企业级SSD主控芯片增强了I/O调度, 减少了不必要的数据移动, 并在持续的混合工作负载下保持平衡的功耗与性能运行。

近线SSD和温数据存储—支持TLC和QLC NAND的SM8366和SM8388

对于容量驱动的AI基础设施, 近线SSD存储优先考虑高密度和成本效益。慧荣科技的企业级SSD主控芯片支持TLC和QLC NAND, 能够为温数据集和扩展的AI数据存储库实现大容量部署。凭借先进的ECC和具备完善的错误管理机制, 这些主控芯片确保了在高密度QLC环境中的数据完整性和可靠运行, 为大规模AI部署提供可扩展的存储。

适用于AI服务器的启动盘解决方案—企业级PCIe NVMe BGA启动SSD和SM8008主控芯片 在AI数据中心基础设施中, 启动存储构成了系统可用性和安全运行的可靠基础。慧荣科技提供企业级PCIe NVMe BGA启动SSD解决方案, 专为紧凑集成、强耐用性和长期稳定性而设计, 已部署于AI系统中。

慧荣科技还提供SM8008启动盘主控芯片, 使客户能够开发量身定制的灵活、高可靠性启动存储解决方案, 以满足AI服务器平台的需求。

欢迎参观GTC慧荣科技3015号展台, 了解慧荣科技的存储解决方案如何能够在整个NVIDIA AI生态系统中实现可扩展、高效且可靠的AI存储架构。

关于慧荣科技

慧荣科技 (NasdaqGS:SIMO) 是为固态存储设备提供NAND闪存主控芯片的全球领先企业之一。该公司为服务器、PC和其他客户端设备提供的SSD主控芯片在全球SSD主控芯片市场拥有广泛影响力, 并且是用于智能手机、物联网产品和汽车应用的eMMC和UFS嵌入式存储主控芯片的领先商户供应商。

慧荣科技还为超大规模数据中心、工业系统和车用SSD提供定制的高性能解决方案。其主控芯片旨在推动世界上高性能、高可靠性的存储平台, 集高性能、低功耗和成熟可靠性于一身。

我们的客户包括世界上大多数NAND闪存供应商、数据中心和企业级存储解决方案提供商、存储设备模块制造商和领先的OEM, 他们都依赖于慧荣科技成熟的主控芯片技术来实现创新性和高质量的存储解决方案。