August 4, 2026

Marvell Advances AI Memory Infrastructure Portfolio to Accelerate Agentic AI Inference

- **Industry's broadest portfolio:** New Marvell® Bravera® and Photonic Fabric™ innovations build on Marvell Structera™ CXL memory leadership across AI storage, rack-scale memory expansion and pooling, and pod-level optical shared memory.
- **Overcoming memory bottlenecks:** Increased access to high-capacity, high-bandwidth memory across AI infrastructure helps address growing KV cache demands, reduce GPU stalls, improve utilization and increase token efficiency in next-gen inference workloads.
- **Optimized for hyperscale deployments:** Bravera SC6 SSD controller features NAND-agnostic architecture from multiple suppliers, providing cloud providers with the flexibility to build and scale AI memory infrastructure.

SANTA CLARA, Calif.--(BUSINESS WIRE)-- **FMS 2026 –** Marvell Technology, Inc. (NASDAQ: MRVL), a leader in data infrastructure semiconductor solutions, today announced new innovations across its AI memory infrastructure portfolio, advancing its position spanning server-level AI storage, rack-scale CXL memory expansion and pooling, and pod-level optical shared memory. The portfolio helps hyperscalers and cloud providers scale memory more independently from compute, improving infrastructure utilization, scalability and token efficiency for agentic AI inference.

This press release features multimedia. View the full release here: https://www.businesswire.com/news/home/20260804094436/en/

As agentic AI inference scales, memory capacity, bandwidth and connectivity are becoming as critical as compute. Larger models, longer context windows and growing KV caches are driving unprecedented memory demands and exposing the limits of traditional server-attached architectures. Increasingly, AI performance and token efficiency depend on how efficiently processors can access and move data across memory resources.

Memory disaggregation addresses this challenge by enabling memory to be expanded, pooled and shared more independently of compute. By making memory more accessible and reducing data movement and latency, the Marvell portfolio helps reduce GPU stalls and improve model FLOPs utilization, allowing AI infrastructure to generate more tokens within the same data center footprints and power envelopes.

"AI infrastructure is moving beyond isolated servers to systems where compute, memory and connectivity operate seamlessly together," said Will Chu, executive vice president and general manager, Custom Cloud Solutions, at Marvell. "As AI scales, memory must scale more independently of compute so resources can be deployed where they deliver the greatest value. With the industry's broadest AI memory infrastructure portfolio, Marvell is



Marvell Bravera® SC6 PCIe 6.0 SSD controller

helping customers improve utilization, boost token efficiency and scale AI without compromising performance, power or cost."

"As AI workloads grow larger and more complex, memory capacity, bandwidth, latency and data movement are becoming primary constraints on AI performance," said Alan Weckel, co-founder and technology analyst at 650 Group. "Marvell's memory and storage portfolio gives hyperscalers and cloud providers a strong foundation for building scalable, efficient AI systems capable of supporting increasingly advanced workloads."

Server-level AI Storage

The Marvell Bravera SC6 PCIe® 6.0 SSD controller advances storage performance for AI inference, enabling more KV cache to move from high-bandwidth memory to SSD to improve infrastructure efficiency. Doubling the performance of the widely deployed Bravera SC5 PCIe 5.0 SSD controller, Bravera SC6 helps cloud providers support demanding AI, cloud and enterprise workloads while improving performance, reducing write amplification and extending NAND endurance. Optimized for hyperscale deployments, Bravera SC6 supports NAND from multiple suppliers, providing cloud providers with greater flexibility in storage sourcing and deployment. Visit the Marvell blog for more information here.

Rack-level Memory Expansion and Pooling

The Marvell Structera X memory expansion solutions help hyperscalers scale memory more efficiently for increasingly memory-intensive AI workloads. Developed in close collaboration with leading hyperscalers, Structera X enables customers to expand, optimize and better utilize memory resources, including extending the value of existing memory investments while supporting demanding AI inference applications. As larger models, longer context windows and growing KV-cache requirements make memory capacity and bandwidth critical infrastructure constraints, Structera X provides a path toward more flexible and

disaggregated memory architectures. By enabling larger memory pools and more efficient sharing of resources across servers, the platform helps improve infrastructure utilization, operational efficiency and total cost of ownership while laying the foundation for future advances in CXL-based memory pooling and sharing. Visit the Marvell blog for more information here.

Pod-level Optical Shared Memory

The Marvell Photonic Fabric memory modules, Photonic Fabric NIC and Photonic Fabric chiplets are foundational elements of a multi-rack optical shared-memory architecture for AI inference. The solution creates a new shared-memory tier across multiple XPUs and racks up to 50 meters, enabling up to 32TB of warm KV cache offload with high bandwidth and extremely low latency. By expanding memory access across AI clusters and enabling KV cache to be loaded from a shared memory tier rather than storage, Photonic Fabric helps increase inference throughput, support larger models and longer context lengths, and increase token efficiency by delivering up to 2-3x higher token throughput within existing data center footprints and power envelopes. Visit the Marvell blog for more information here.

Availability

The Marvell Bravera SC6 PCIe 6.0 SSD controller is expected to begin sampling in Q4 2026.

Visit the Marvell booth #805 at FMS 2026 to learn more about the company's latest accelerated memory and storage offerings.

About Marvell

To deliver the data infrastructure technology that connects the world, we're building solutions on the most powerful foundation: our partnerships with our customers. Trusted by the world's leading technology companies for over 30 years, we move, store, process and secure the world's data with semiconductor solutions designed for our customers' current needs and future ambitions. Through a process of deep collaboration and transparency, we're ultimately changing the way tomorrow's enterprise, cloud and carrier architectures transform—for the better.

Marvell and the M logo are trademarks of Marvell or its affiliates. Please visit www.marvell.com for a complete list of Marvell trademarks. Other names and brands may be claimed as the property of others.

This press release contains forward-looking statements within the meaning of the federal securities laws that involve risks and uncertainties. Forward-looking statements include, without limitation, any statement that may predict, forecast, indicate or imply future events, results or achievements. Actual events, results or achievements may differ materially from those contemplated in this press release. Forward-looking statements are only predictions and are subject to risks, uncertainties and assumptions that are difficult to predict, including those described in the "Risk Factors" section of our Annual Reports on Form 10-K, Quarterly Reports on Form 10-Q and other documents filed by us from time to time with the SEC. Forward-looking statements speak only as of the date they are made. Readers are cautioned not to put undue reliance on forward-looking statements, and no person assumes any obligation to update or revise any such forward-looking statements, whether as a result of new information, future events or otherwise.

View source version on businesswire.com:
https://www.businesswire.com/news/home/20260804094436/en/

Media Contact:
George Millington
pr@marvell.com

Source: Marvell Technology, Inc.