

10122026 Maia 300 October News Announcement

Advances across the infrastructure layers

Saurabh Dighe – CVP, Azure Hardware Systems Planning and Architecture

Monday, October 12, 2026

SAURABH DIGHE: Maia 300 really builds on the strong foundation of software-defined dataflow architecture that we introduced with Maia 200, but we did make major advances in five areas.

First, we moved to a multi-chiplet architecture onto the N2 process node from TSMC. Second, we went to HBM4 and we added substantially greater HBM capacity and bandwidth. We added new precision data types and MX precision data types for computation and KV storage. We did advance our next-generation of scale-up network.

And then finally, we built a larger, more power-efficient rack scale system.

And so together these changes increase the resources available for long-context decode-heavy inference from a single accelerator through the entire full stack rack and data center.