cādence®

Solutions   Products   Services   Support & Training   Alliances   **Community**   About Cadence

Home > Community > Blogs > Industry Insights > memcon keynote why hybrid memory cube will revolutionize system memory

Share   Subscribe   Contact   Print   Search Community

MemCon Keynote: Why Hybrid Memory Cube Will "Revolutionize" System Memory

By Richard Goering on September 19, 2012
Comments(0)

Filed under: Industry Insights, TSVs, Micron, stacked die, 3D, memory, DRAM, MemCon, 3D-IC, DDR4, DDR3, 2.5D, hybrid memory cube, stacked memory, Scott Graham, HMC Consortium, HMC

DDR3 and DDR4 aren't enough - it's time for a "revolution" in system memory that will offer exponential improvements in bandwidth, latency, and power efficiency, according to Scott Graham (right), general manager of Hybrid Memory Cube technology at Micron. In a keynote speech at the MemCon 2012 conference produced by Cadence Sept. 18, Graham made a strong case for the Hybrid Memory Cube (HMC) architecture as the next big move in system memory.

Here's some quick background. The HMC architecture, originally developed by Micron, uses a small, high-speed control logic layer that sits beneath vertical stacks of DRAM die. Compared to 2D memory technologies, it promises greatly reduced latency, increased bandwidth, power reduction, and a smaller form factor. The HMC Consortium is working on a draft HMC interface standard.

"I don't mean to downplay DDR3 and DDR4, but this [HMC] technology is an exponential jump in memory performance and capabilities," Graham said. "There really is a flat line as far as performance goes in DDR3, DDR4 and other memories. We really need to take a jump up."

Graham reviewed memory challenges by application, noting that servers need high performance, networking applications need reduced latency, and mobile devices need lower power. He noted, however, that "HMC is not architected today to be a mobile memory solution - it's targeted at high-performance computing and networking solutions. But, this gives us an opportunity to use similar architectures to go into mobile applications."

A Look Inside

So what's in the stack? Graham noted that the HMC architecture combines a fast logic layer and advanced DRAMs in one optimized package. He said HMC provides an "exponential jump" in two directions - extreme high performance, and lower power per bit. "For the first time, we have a [memory] technology that will allow us to go faster and use less fuel."

One HMC enabling technology is an abstracted memory management layer. While the DRAM uses the traditional DRAM core cell architecture, it's been restructured so it uses memory vaults rather than arrays. A logic controller is placed at the base of the DRAM stack. The assembly is interconnected with through-silicon vias (TSVs) that go up and down the entire stack. The final step is advanced package assembly.

Here are some more details about the architecture:

- Micron re-partitioned the DRAM into 16 partitions and stripped away the common logic that exists in normal DRAMs
- The HMC stacks DRAMS in 4-high or 8-high configurations
- There are 16 independent vaults in an HMC - these can be thought of as channels
- A high-speed SerDes interface connects the memory cube and the processor
- The link controller interface includes 16 transmit lanes and 16 receive lanes, each running at 10Gb/second each
- HMC supports both near memory and far memory configurations. Multiple cubes can be connected together
- The packaged solution is slightly larger than a U.S. quarter dollar

In a first-silicon demo last year, Graham said, an HMC operated at 121 Gbytes (yes, bytes) per second.

DDR Versus HMC

So how does HMC "stack up" against DDR3 and DDR4? Graham first showed a reliability, availability, and serviceability (RAS) feature comparison, and concluded that HMC is far ahead, given that it not only yields well but has extensive self-repair and error correction capabilities. What really got my attention, however, was the "extreme performance comparison" that showed what it takes to support 1.28 Terabytes/second performance. Here are the requirements of HMC versus DDR3L-1600 and DDR4-3200:

Active Signals

- DDR3 requires 14,300
- DDR4 requires 7,400
- *HMC only needs 2,160 - 85% less than DDR3*

Operating Power (including CPUs)

- DDR3 requires 2.25KW
- DDR4 requires 1.23KW
- *HMC system only needs 350W - 72% less than DDR4*

Board Space

- DDR3 requires 165,000 sq mm

- DDR4 requires 82,500 sq mm
- *HMC only needs 8,712 sq mm - 90% less than DDR4*

What does it take to support 60GByte/second performance? Here the comparison showed that HMC uses fewer channels, less board area, and fewer active pins, and provides much more bandwidth per pin than DDR3 and DDR4.

Micron decided to not keep HMC technology to itself. "We realized we had something very unique and very special, and we wanted to make sure we could share it," Graham said. "We thought it had enough runway to solve problems over the next 10-15 years and be *the* memory standard going into the future." So, Micron and Samsung launched the HMC Consortium, which now has 49 "adopter" members, including Cadence. A draft interface spec has been delivered to adopters, and the consortium expects to complete the first industry spec by the end of 2012.

Sneak Peek

Graham concluded with a "sneak peek" of a derivative product that will be based on HMC technology. It's a 2D or 2.5D package that places the processor next to, not under, the DRAM stack. (In Micron's terminology 2D uses a lower cost, lower performance substrate compared to 2.5D). This technology is designed to full a performance gap between HMC, which is expected to reach speeds up to 160 Gbytes/second, and conventional memories that run much more slowly. Graham said Micron will target performance around 60-65 Gbytes/second with this new 2.5D technology.

"In summary, we think both 2.5D and HMC are revolutionary shifts, and we need help from the industry in thinking about how to use this technology," Graham concluded.

Richard Goering

Related Blog Post

MemCon Keynote: Cloud, Mobility Disrupt Semiconductor Memory Ecosystem

Comments(0)

- -

Leave a Comment

- -

Name	
E-mail	(will not be published)
Comment	

☐ I have read and agree to the Terms of use and Community Guidelines.

[POST COMMENT ▶]