Massive Memory for AMD EPYC-based Servers at a Fraction of the Cost of DRAM

The ever-expanding digital economy has created significant demands for both real-time and batch processing of large data sets. IT organizations across the world are leveraging in-memory computing to drive superior application performance and obtain meaningful insights using advanced business analytics. However, in-memory computing can become bottlenecked by set limitations on the amount of memory available to the server, as well as prohibitive DRAM pricing.

This paper describes how system memory for AMD EPYC™ processor-based servers can be transparently expanded to address the growing demand for in-memory computing and higher memory-per-core ratio in a scalable and cost-effective way.

Scaling System Memory and Memory/CPU ratio

In-Memory databases are becoming ubiquitous as they deliver the required performance for real-time analytics and business insight, analyzing vast streams of data from Internet users, transaction events, and IoT devices. High concurrency environments, such as container-based applications, are an additional key example where memory usage can quickly outpace processing capabilities, at which point the options are limited and expensive, generally requiring adding more servers (scale-out) to house the extra memory.

The growth in compute power of a single system and of workload memory usage have only been partially met by developments in the DRAM market, and server systems remain limited in the number of DRAM slots per processor. As of the time of writing, the $/GB sweet spot is the 32GB DIMMs; 64GB DIMMs come with a premium, and 128GB DIMMs are even more expensive and are in short supply with long lead times.

With the introduction of EPYC server processors, AMD is providing excellent compute density with high core counts, and large memory capacity with support for up to 2TB per socket or 4TB per dual-socket server. However, as data-sets grow in size, even this amount of memory may not be sufficient for demanding use-cases. For some workloads, more servers can be added to a cluster for a given application, solely for the sake of providing more memory, but this results in lower server utilization and a corresponding increase in TCO.

As such, having the ability to break through the memory capacity limitations is critical for efficient and cost-effective data-center infrastructure.
Memory Expansion

AMD EPYC™ processors support memory expansion to address the issue of DRAM limitations and cost. AMD has partnered with Western Digital Corporation to bring to market a transparent memory expansion solution: Ultrastar® DC ME200 Memory Extension Drive for AMD EPYC. The Ultrastar® DC ME200 can be used to scale existing system memory, promote server consolidation, and reduce the complexity of splitting multi-terabyte data sets across multiple servers.

Ultrastar® DC ME200 provides applications with large amounts of system memory at a fraction of the cost of DRAM as at its base it uses high-performing enterprise-grade NAND Non-volatile Memory (NVM), which is low cost and high density. Advanced software algorithms work at the firmware level to enable byte-addressability for DRAM-like operation, and maintain near DRAM levels of performance across a variety of applications, especially targeting highly parallel workloads with high numbers of transactions. Once installed, the solution is transparent, requiring no changes to the existing OS and application stacks. See Figure 1.

This results in a more efficient data-center, by enabling more work to be accomplished within each server, or more value to be extracted from the server by hosting more users or customers.
Use-Case Scenarios & Cost Savings

There are three key scenarios in which it is beneficial to use Storage Class Memory:

- Where very large system memory is required – more than the DRAM capacity of a commercial-of-the-shelf (COTS) server system by specification
- Where the cost savings of replacing DRAM with NVM outweighs the performance difference between the two technologies
- Where both of the above can be combined, i.e. where the software-stack is distributed and can use scale-out, but a lower number of scale-out servers with much large system memory in each significantly reduces cost

For the purposes of this paper, we shall refer to the first case as “Memory Expansion”, the second as “Memory Replacement”, and the last as “Footprint Reduction”.

Examples of Memory Expansion

Consider in-memory database engines, which require a shared-memory (non-distributed) system. With Ultrastar DC ME200, an in-memory database of larger than 10TB can easily be run on a COTS dual-socket server system. In fact, in-memory database deployments of any size larger than a server's DRAM limit may benefit from using the Ultrastar DC ME200. Additionally, there are many workloads that cannot be run without sufficient system memory: in scientific computing, in computer-aided engineering (CAE), and in business domains where users analyze or run simulations on very large data sets, such as in Financial Institutions.

Examples of Memory Replacement

Multitenancy scenarios are very common in enterprise IT, and even more popular with cloud service providers. The more workloads (e.g. multiple containers, multitenant databases, etc.) one can place into a single physical server, the better the utilization or yield of the infrastructure. In most multitenant cases, system memory limits the number of workloads per server, so there is reason to deploy systems with maximum commodity-memory capacity. With Ultrastar DC ME200, less DRAM is used per node, as it is replaced by NVM. In some cases, even more system memory is enabled by Ultrastar DC ME200 to further leverage the economic benefits.

Examples of Footprint Reduction

Many in-memory software stacks assume scaling out using more servers in order to host larger data in-memory in a cluster. As data size grows, servers need to be added to the cluster. This is common with in-memory key-value and with in-memory processing frameworks. In many cases, CPUs are severely underutilized in such clusters.
Using Ultrastar DC ME200, the amount of system memory per node can be significantly increased, and the cluster size can be correspondingly reduced. For example, a cluster holding 30TB of data in memory, using 30 nodes of 1TB each, can be reduced to only 8 nodes with 4TB of system memory each, and as a side-effect CPU utilization is increased.

**Performance**

Figure 2 shows results from tests performed with a Segmented-SGEMM (Single precision floating General Matrix Multiply) – a form of computation which is used in Artificial Intelligence applications. We compare the computational efficiency of a system with 512 GB of DRAM across the three configurations:

- using only DRAM for memory, using ~20 GB of the memory,
- using only DRAM for memory, using more than 80% of the available memory (~440 GB),
- using DRAM plus 4 TiB of Ultrastar DC ME200 memory, where the data set is expanded to take up more than 80% of the total memory (~3.4 TiB of memory)

![Segmented-SGEMM GFLOPS](image)

*Figure 2: Performance (in Giga-FLOPS) of DRAM-only vs. Ultrastar Memory*

The results demonstrate that the computational efficiency of the Ultrastar DC ME200 configuration is 89.2% that of the DRAM-only benchmark. It is important to note that workload run on the Ultrastar DC ME200 configuration is eight times larger than the DRAM-only configurations. It simply is not possible to run a workload this large on the DRAM-only systems.
Economic Model

Figure 3 compares two Ultrastar® memory configurations to a baseline configuration of the same dual socket AMD EPYC™ based system with 2.0 TiB of DDR4.¹

- **Configuration A** shows that reducing DRAM from 2.0 TiB to 256 GB and adding two 1 TiB Ultrastar DC ME200 devices can reduce total system cost by 40%.

- **Configuration B** shows that reducing DRAM from 2.0 TiB to 512 GB and adding two 2.0 TiB Ultrastar DC ME200 devices, maintains same system cost, while providing applications with 100% more memory.

Both Configuration-A and Configuration-B showcase how EPYC processor-based servers can deliver larger system memory than DRAM-only servers using Ultrastar DC ME200. This enables cost-effective memory expansion and server consolidation by allowing for the increase in size of data sets that can be handled within a single server, thereby reducing the total number of servers needed, server and core-based application license expenses, and many associated OPEX and CAPEX costs.
Summary

AMD EPYC™ based servers, combined with Ultrastar® DC ME200 enable increasing the amount of memory per server (and the GB/CPU ratio), in order to provide a more cost-effective data-center for in-memory applications with minimal impact on performance.

For More Information

For more information about AMD’s EPYC line of processors visit: [http://www.amd.com/epyc](http://www.amd.com/epyc)


Footnotes

1. DRAM Pricing as of JAN-2019, Source DRAMXchange; Ultrastar DC ME200 street/volume price, source Western Digital