Western Digital.

SOLUTION BRIEF

Deploy Large Memory Scientific Codes on Fewer Servers

Ultrastar DC ME200 Memory Extension Drive, NVMe U.2 and AIC HH-HL form factors

TCO Highlights

- Provide TB-class RAM per server with NVMe™
- Improve memory density
- Reduce scale-out complexity with in-memory datasets
- Optimal for very large maxtric operations, block and stream-based codes, and algorithms that require a single system image

Product Features

- Up to 24TiB software-defined memory per 2U server (U.2)[1]
- Works transparently with existing Linux® distrubutions and application stacks

Ultrastar® DC ME200 Memory Extension Drive for scientific codes enables better memory management for greater scientific computing

Datasets are growing at an exponential pace. This is great for scientific workloads since more detailed observations are made, confidence intervals can shrink and calculations can arrive at more meaningful answers. Groundbreaking discoveries in medical, astrophysical, metrological, fundamental physics, computational fluid dynamics, and more are now being made thanks to this new data.

In scientific computing, many codes can be broken down to work on small subsets of data with their results coalesced into a single, final result. Such divide-and-conquer methods have enabled the explosion of clusters of discrete servers as used in modern HPC clusters. In these clusters, problems are segmented into chunks that fit into the working memory of the individual servers in the cluster, usually somewhere between 64 and 256 GiB. As the problem grows, either the number of servers required to fit it into DRAM is increased, or it simply takes longer to get an answer as more jobs of the available DRAM-size are queued.

The Ultrastar DC ME200 Memory Extension drive from Western Digital can change that equation by expanding server RAM up to eight times the capacity of its physical DRAM through the use of one or more custom NVMe devices. This can allow certain classes of memory-constrained codes to run on fewer servers, reducing needed cluster time and allowing an HPC cluster to be more effective.

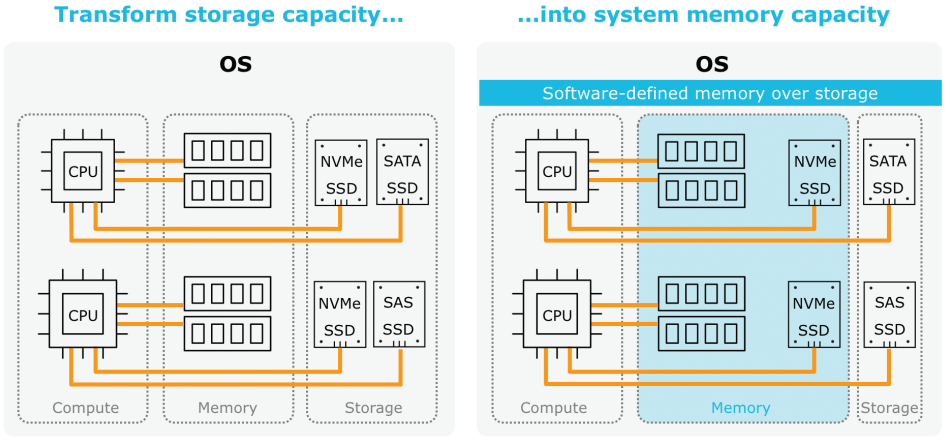


Figure – The left diagram depicts the traditional compute-memory-storage architecture, The right digram depicts how database instance configurations can take advantage of Ultrastar memory drives to augment server DRAM to create a virtualized memory pool to enable greater memory expansion.

[1] Memory capacity is indicated by GiB and TiB and based on binary values such that one gibibyte (GiB) is equal to 2^{30} bytes and one tebibyte (TiB) is equal to 1024 GiB (2^{40}) bytes. Suggested expansion of 8x DRAM based on internal performance testing across a variety of industry standard benchmarks.

Providing Terabytes of RAM in a 1-U Server

Up to 24 TiB of effective RAM space can be provided in a single 2-socket, 2U server using 3 TiB of DRAM, multiple Ultrastar DC ME200 drives and provided software. More modest memory expansion configurations are supported, for example 1 TiB of effective RAM using only 128 GiB of DRAM and 1 TiB of Ultrastar DC ME200 drives.

Standard Linux operating systems are supported, with the memory extension technology loading before and operating below the level of the OS. The operating system simply sees the combined DRAM and and Ultrastar DC ME200 capacity. Applications simply expand their memory allocations and process larger chunks of datasets per instance.

Code Classes that Can Benefit

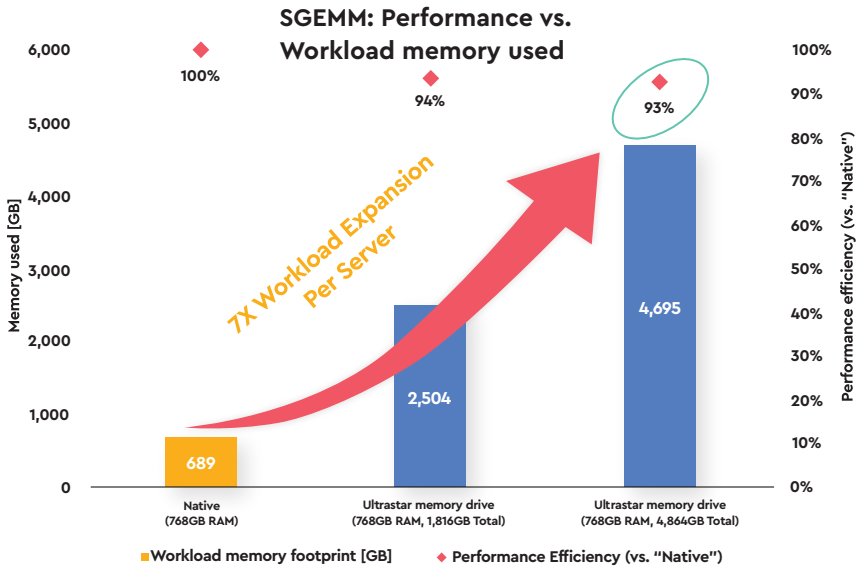
Many codes can benefit from the Ultrastar DC ME200 with little to no performance impact, such as:

- Very large matrix codes with high locality operations
- Highly parallelized, block-based memory accesses
- Streaming/striding memory-bound operations
- Codes which are complicated to split across multiple system images

Codes which are heavily CPU-limited, single-threaded, or which have a completely random access pattern may not benefit fully from the Ultrastar memory drive. However, when the cluster is running multiple single-threaded jobs on a single node, the extra RAM available with the Ultrastar memory drive may still improve node utilization by allowing multiple instances of single-threaded jobs.

Worked Example: Very Large Matrix Multiplication (SGEMM)

It is very common to have large matrix operations as a large part of scientific codes. A standard kernel used by many of these codes is the Single-precision GEneral Matrix Multiply, or SGEMM. The following test results were obtained using a segmented SGEMM of an in-memory matrix of increasing size, from 768 GiB (DRAM) up to almost 5 TiB (DRAM plus Ultrastar memory drive). The matrix multiply was performed in segments in a similar way to that which it would be broken up to execute over a small cluster of nodes, and the effective FLOPS were calculated. The results are shown in the graph: a matrix seven times larger than DRAM was operated on by the single node with less than a 10% performance degradation versus operating on one fully fitting within DRAM.



For More Information

Contact your Western Digital representative or go to https://www.westerndigital.com/products/data-center-drives/ultrastar-dc-me200-memory-extension-drive to get more detailed product information on the Ultrastar DC ME200 Memory Extension Drive.

Western Digital.

5601 Great Oaks Parkway
San Jose, CA 95119, USA
US (Toll-Free): 800.801.4618
International: 408.717.6000

www.westerndigital.com