REPORT REPRINT

MemVerge builds a memory-converged platform using Optane NVDIMMs

JUNE 14 2019

By Tim Stammers

The company is beta-testing software it says dramatically boosts performance for memory-hungry applications such as analytics by pooling and managing DRAM, Intel Optane memory and flash across server clusters. Intel's Optane NVDIMMs are a key element in the system, which MemVerge is calling memory-converged infrastructure. Has the startup created a new category of infrastructure product?

451 Research®

Introduction

The prospects for 3D XPoint (3DXP) memory have been significantly boosted by Intel's recent GA launch of what it is branding as Optane DC Persistent Memory, or what might be more simply called Optane NVDIMMs. The performance of Optane memory – as Intel has begun calling 3DXP – is far better exploited in NVDIMMs than it has been to date in Optane NVMe drives. Existing applications can work with Optane NVDIMMs immediately on existing operating systems and virtualization platforms. However, to make the most of the devices by exploiting the memory's persistence or non-volatility, applications and operating environments need reworking. Enter MemVerge, a startup that has created a platform that ties together and manages DRAM, Optane NVDIMMs and flash capacity across clusters of servers to achieve what it says is the optimal use of those resources – and eliminates the need to modify applications to exploit the persistence of Optane memory.

MemVerge's software is fundamentally a hyperconverged infrastructure (HCI) platform, running on commodity hardware – apart from the obvious exception of Optane NVDIMMs and the presence of upmarket versions of Intel's latest Cascade Lake Xeon processors needed to support the NVDIMMs. But the MemVerge platform is unlike any other HCI product because it is also a memory management system. For this reason, MemVerge has coined the term memory-converged infrastructure (MCI) to describe its platform, which it says boosts both IO speeds and effective memory size tenfold. Storage services such as snapshots and replication are being developed for the platform, which is called Distributed Memory Objects (DMO).

Beta testing begins this month and GA is expected in 2020. Alpha testers have included LinkedIn, Tencent Cloud and JD.com, and the major target applications are analytics, machine learning (ML) and data warehousing. DMO will support any alternative, mainstream storage-class memories that may appear in the future as rivals to Optane memory, according to MemVerge.

451 TAKE

451 Research believes MemVerge's product is category-busting, and we think the coining of the MCI label is perfectly justifiable. We're unaware of any other platform that is claimed to create a data structure in the same way as the MemVerge platform. The company's prospects depend on the strength of demand for mechanisms to boost performance for applications such as analytics and ML, and the effectiveness of its software in doing that. For the first point, analytics and ML are growth markets. As for the second point, we are not qualified to judge the technical merits of MemVerge's software, but the presence of LinkedIn and Tencent among its alpha testers is noteworthy. We're confident that the ability of the software to allow un-modified applications to exploit the persistence of Optane NVDIMMs will also be a key selling point for some while, until ISVs have finished reworking their applications to do that.

Company background

Founded in 2017 and based in San Jose, MemVerge has raised $24.5m funding in a series A round from investors Gaorong Capital, Jerusalem Venture Partners, LDV Partners, Lightspeed Venture Partners and Northern Light Venture Capital. Former VMware and EMC executive Charles Fan and Caltech professor Shuki Bruck founded the company. Fan and Bruck co-founded Rainfinity and Bruck also co-founded XtremIO. Rainfinity, a NAS virtualization specialist, was bought for what EMC said was less than $100m in 2005. All-flash storage pioneer XtremIO was bought by EMC in 2012 for a reported $430m.

MemVerge technology

MemVerge's DMO software uses DRAM as a cache front end to the Optane NVDIMMs and tiers data across Optane NVDIMMs and NVMe flash drives. MemVerge says it plans to develop storage services such as snapshots and cluster-to-cluster replication for data protection and disaster recovery alongside multi-tenancy features and data reduction mechanisms functions.

The software is designed to boost performance for applications that are 'memory constrained'; i.e., data-intensive applications for which IT organizations have attempted to boost performance by maximizing the amount of DRAM in servers but have either hit physical limits or cost limitations. Using its software to bring shared Optane NVDIMMs into play, MemVerge says a cluster of up to 128 servers can access a shared pool of up to 768TB of memory, with up to 72GB/sec read bandwidth per server, at less than one microsecond latency. Latency is key to performance for many applications that access data randomly. As a comparison, enterprise flash drives typically impose about 50 to 100 microseconds' latency, while all-flash arrays impose hundreds of microseconds' latency. MemVerge says its software also increases bandwidth for sequential data access.

When one server accesses resources that are physically located on another server, there is a penalty in the form of the latency imposed by the network link. MemVerge says that when RDMA-enabled RoCE Ethernet is used, the additional latency is between only one and five microseconds, and that its software localizes data and needs to reach across the network only when local resources are exhausted.

Applications and performance boost

As a measure of overall performance boost, MemVerge cites the experiences of its alpha customers over the last 12 months. It says one customer running GraphSage machine learning for a social network replaced a 50-server, 24TB DRAM cluster with a single 8-way server, while also cutting the time taken to load data for training. In another case, MemVerge says its product accelerated training of a TensorFlow ML model by allowing checkpoints to be made more rapidly and frequently, providing quick recovery from failures, and by faster loading of data. Overall training speed was increased sixfold, while data was imported 350 times faster, according to MemVerge.

In a third case, an implementation of Spark analytics had reached the limits of DRAM capacity, and as a result was suffering from slow IO to disk, while flash drives were being burned out quickly by frequent intermediate data writes. Using DMO, Spark Terasort speed was increased fivefold, while RDD caching speed was boosted sevenfold, according to MemVerge. DMO also allowed intermediate data to be moved off the Spark servers, making the implementation more scalable. According to MemVerge, Kubernetes can be used to 'easily' deploy containerized versions of applications such as Spark, Presto and Tensorflow on top of DMO.

No app modifications

Optane NVDIMMs work in two modes. The first is called Memory Mode, which requires no modifications to applications. That is a major advantage, and as a result, Intel expects that initially, Memory Mode will be the most popular way to use Optane NVDIMMs. However, Memory Mode does not fully exploit the benefits of Optane memory, and as a result, Intel expects that over time, the other mode, called Application Direct (AppDirect), will become more popular. This is despite AppDirect Mode's need for applications to be modified, and it is because Memory Mode does not enable persistent writes to Optane NVDIMMs.

The fact that Optane memory is far cheaper and offers much greater capacities than DRAM is a major virtue, but not using its persistence is leaving a large part of the story on the table. Persistence is part of what qualifies Optane as a storage-class memory. In Memory Mode, data writes to an Optane NVDIMM are not safe against power loss, and depending on the application, cannot be acknowledged

until they have been copied to safety in an NVMe flash or Optane drive. Those NVMe drives are orders of magnitude slower than the NVDIMM, making this a performance-sapping requirement. Putting the persistence of Optane memory to work avoids this problem.

AppDirect Mode exploits Optane's persistence and is used by any application running on the MemVerge platform – whether or not the application has been modified. However, without MemVerge's DMO software, AppDirect requires applications to be modified, and it is not clear how long software vendors will take to do this. Intel says Windows, Linux and VMware vSphere have already been modified to support AppDirect, alongside SAP HANA, Aerospike, GigaSpaces, BaoSight and Apache's Cassandra, Spark, HBase Bucket Cache and HDFS Cache, and the Cloudera version of Apache HBase Bucket Cache. (All these applications use AppDirect with the fast byte-level transfers that the Optane NVDIMMs were designed to use, and not the slower block-level transfers that are also an option.)

451 Research is very confident that many software vendors will support AppDirect mode. As one example, Oracle has said it strongly committed to using AppDirect, which it says will be 'extremely' useful to database vendors in particular and will drive major architectural changes to their software. However, we also believe the process of modifying databases could take some time. MemVerge predicts that over the next two years, only a small number of applications will have been modified, and that while others will follow later, some applications will never be modified to support AppDirect.

Competition

451 Research is aware of only one other server-based software product that allows the persistence of Optane NVDIMMs to be exploited by unmodified applications. That product is NetApp's MAX Data, which is similar but different from MemVerge's software. MAX Data is installed on host servers and boosts performance by integrating a layer of Optane NVDIMM capacity with back-end NetApp storage; it allows unmodified applications to use Optane NVDIMMs in AppDirect Mode; and it exploits the persistence of the NVDIMMs for fast application recovery after server failures. But while MemVerge's product uses DRAM as cache and tiers data across NVDIMMs and NVMe drives within a cluster of servers, MAX Data does not involve the DRAM layer, and only tiers data between Optane NVDIMMs and NetApp storage systems, and only for individual host servers.

SWOT Analysis

STRENGTHS

MemVerge's technology ties DRAM, Optane NVDIMMs and NVMe flash drives into a single distributed memory and storage system and allows the maximum benefits of Intel's Optane NVDIMMs to be realized using unmodified applications.

WEAKNESSES

Although the startup's founders have successful track records, the company itself and its software is as yet little known and introduces an extra layer of complexity and vulnerability into operating environments.

OPPORTUNITIES

The analytics and machine learning markets are growing fast. MemVerge's DMO software was designed to meet the needs of these and other data-intensive applications for faster data access and hence faster overall performance.

THREATS

The success or failure of DMO is tied to that of Optane NVDIMMs. If those devices are not priced attractively compared with DRAM DIMMs, they will not be popular. Currently, DRAM prices are falling, putting Intel under pressure with respect to this issue.