White Paper

Unleashing the Power of In-Memory Computing: Intel Optane DC Persistent Memory for SAP HANA

Sponsored by: Intel and SAP

Carl W. Olofson
April 2019

IN THIS WHITE PAPER

This white paper reviews the power and promise of in-memory database systems and their applications and considers the limitations imposed up to this point on those systems, including memory size per CPU, operational complexity due to the nature of volatile memory, and limitations to database performance made necessary by the volatile nature of main memory. It then examines the benefits of new persistent memory (PMEM) technology for in-memory database management systems (DBMSs), including larger memory capacity, and greatly accelerated start-up and recovery times — all of which ensure greater business continuity than can be expected today and add up to a lower total cost of ownership (TCO). Finally, it considers a leading example of this technology — Intel Optane DC persistent memory (Intel Optane DC PMM) — and its use in combination with SAP's in-memory database, SAP HANA.

SITUATION OVERVIEW

Current Common Implementations of In-Memory Databases

Formerly, all database systems were storage optimized. This means that the data organization and layout were designed to optimize their placement on disk for the purpose of minimizing the biggest impact on performance: storage I/O. Over the first two decades of the 21st century, database systems that optimize data organization and layout for management in memory rather than storage have emerged. This technology took advantage of the advent of 64-bit addressability and falling memory prices to enable databases to drastically reduce their I/O rates and offer much better performance.

Today, most such memory-optimized database systems (commonly called "in-memory databases") come in one of two forms: either row oriented for transaction optimization (for simplicity of retrieval because transactions usually include most columns in row by row processing) or column oriented for analytical workloads (because analytics tend to involve searches for specific values or ranges in columns, execute complex joins, and retrieve select column values for the result set). It should be noted that in a full in-memory database, the performance difference between row- and column-oriented data organization for transactions is negligible, but the performance gains for analytics in a column-oriented system are quite significant because analytics tend to use a limited set of columns and require accelerated searches and complex joins for aggregation of result sets.

Such column-oriented systems tend to feature heavily compressed and carefully ordered columns that facilitate rapid lookup and obviate the need for indexing, enabling the system to do something called "vector processing," in which data is ordered in such a way that a whole relevant collection of data comes into the processor cache optimally and can be processed altogether using SIMD (single instruction, multiple data) instructions for very rapid calculations and comparisons.

Very few such systems support both analytics and transactions on the same data (such systems are commonly called "HTAPs," or hybrid transaction and analytic platforms). HTAP systems can support transactional and analytic applications using the same database at the same time, but many cannot support the embedding of analytic queries in transactional applications because they keep the transactional data in rows, the analytic data in columns, and cannot ensure that they are completely consistent at all times. Support for embedding analytic queries in transactions requires a database system with a continuously consistent view of the data, ideally kept in a single form.

How Such In-Memory Databases Work

In the simplest case, in-memory databases keep all their data in main memory, or DRAM. If they are transactional, they must execute periodic backups, or memory dumps, and also maintain a log file containing a record for each committed transaction, since DRAM is volatile, and the database must be reconstructed in case of a system restart or failure. Some in-memory databases maintain multiple replicas instead of implementing these safeguards on the assumption that not all the replicas will fail at the same time.

When the database size exceeds the available memory, many vendors support various techniques to manage data volume, performance, and cost constraints. Some in-memory databases provide for a cluster configuration in which the data is spread among multiple nodes. In the transaction case, this requires the nodes to coordinate data updates to account for overlaps or cross-node data dependencies. Such dependencies can be limited by partitioning and careful application design, but not eliminated.

Some databases deal with the size issue by applying a "heat map" approach. They keep the most active data (the "hot" data) in DRAM while keeping the data that is often referenced but seldom updated in flash (the "warm" data) and data that is seldom accessed at all (but must still be available) on spinning disk (the "cold" data). This approach is called "data tiering."

A number of in-memory databases apply some combination of the clustering and data tiering approaches.

Common Limitations of In-Memory DBMS Operation

Large-scale in-memory databases involve both configuration and operational complexity because of the need for software data tiering solutions due to memory capacity and cost constraints. Significant planning and ongoing maintenance are commonly necessary to manage the data life cycle and data placement.

For example, clusters must be carefully managed, and it is easy to impede performance or introduce an increased risk of failure through seemingly innocuous errors in configuration or operation.

Also, up-front planning and ongoing maintenance of a "heat map" strategy shouldn't be underestimated. Poor "heat map" data assignment can lead to unnecessary overhead because of excessive flash or disk I/O and impact performance.

It's important to note that many in-memory DBMSs are simplifying these solutions by building intelligence into the systems to identify data usage patterns and automate these processes.

Challenges Regarding Transaction Performance

One would think that transaction performance would be limited only by the speed of DRAM, since highly active data is kept in main memory. But this is not quite true. To ensure recoverability, a transaction must be either recorded in a log file or replicated to another system so that in case of failure, the committed data is not lost. If the data is written to a log file, the write must be complete before the commit can be recognized, and even with a log file on flash, that represents a delay. If the data is replicated to another system, the mutual update must be assured before the transaction can be recognized as committed. Again, this entails a delay.

It has been estimated that 80–90% of the average database transaction involves data reads. If the database is implemented using the data tiering strategy, then at least some of these reads will involve warm or even cold data, which means they will be slowed by the I/O time required for those retrievals. Even though this effect is mitigated by buffering, it still represents some overhead.

Start-Up and Recovery

Normally, with an in-memory system, restarting a database after either a normal or abnormal shutdown requires reloading the data from a backup or dump file into memory. If it is a transactional database recovering from an abnormal shutdown, it must also then roll forward the transactions captured in the log file up to the point of failure. The total time to reload all data into memory for optimal performance can range from minutes to hours, depending on the type and size of system and the data volume.

Using In-Memory Databases with Persistent Memory

Persistent memory is memory that sits on the memory bus on the motherboard like conventional DRAM but does not lose data when the power shuts off or the process is terminated. Since the operating system and database heap data require traditional DRAM, persistent memory does not represent a complete substitute but can be used in important ways to dramatically improve performance, ensure greater availability, and lower operational complexity of in-memory database systems. All of this results in lower operating costs and greater user satisfaction.

Systems using the data tiering approach can keep warm data in PMEM rather than flash, dramatically increasing the performance of queries and overall transaction throughput. Writing log records to PMEM rather than a log file on flash results in much shorter transaction commit times, again boosting transaction performance. This does not, however, obviate the need to use conventional block storage to ensure recoverability; it simply enables deferring the writes so that commits are not held up by I/O.

Rapid Recovery Delivers Greater Availability

With persistent memory in its early stages, systems will have a combination of DRAM and PMEM. If applications are properly optimized, they will be able to take advantage of the "persistency" characteristic of PMEM while maintaining high performance (low latency). Systems will be able to quickly get to a state of optimal, in-memory performance at start-up because data will already be in main memory (PMEM) versus having to go to traditional storage and reload all of the data into memory. This can reduce the data load time at start-up, from hours to seconds. This will dramatically help organizations reduce their recovery time objective (RTO) and improve overall business continuity.

Improved SLA Satisfaction and Business Benefits

Clearly, greater performance and availability translate into better SLAs for users, generating greater user satisfaction and enterprise productivity. Simpler operations, resulting from fewer requirements for ensuring recoverability and database start-up, also deliver lower operational cost. The combination means lower TCO and greater overall business benefits to the enterprise.

Intel Optane DC Persistent Memory

Intel has introduced a persistent memory module that delivers on the aforementioned benefits to in-memory databases plus much more. The product is called Intel Optane DC persistent memory. In addition to offering a memory capability that operates on the memory bus and does not lose its memory setting when power shuts off, it also has the following two additional features:

- **Memory mode versus App Direct mode.** In Memory mode, the Intel Optane DC PMM behaves just like ordinary DRAM; the memory controller uses the DRAM as near memory and the Intel Optane DC PMM as far memory, and it may be accessed using normal memory access instructions. This makes it compatible with all existing Intel-based applications. In App Direct mode, it enables applications, using a special protocol, to write data to it that is nonvolatile (i.e., it remains present when power is shut off). DRAM and the Intel Optane DC PMM make up the total online memory, but the data in Intel Optane DC PMM is persistent like traditional storage.
- **Densely packed data.** The Intel Optane DC PMM holds much more data than standard DRAM, greatly increasing the memory capacity of the system. Initially, Intel Optane DC PMMs are available in capacities of 128, 256, and 512GB.

Greater Memory Sizes Mean Better Performance for Larger Data Sets

Intel Optane DC persistent memory's greater memory capacity means that in-memory databases can keep much more data in memory, greatly reducing the need to use clustering and data tiering techniques as a means of managing larger databases. The result is greater operational simplicity, less exposure to the kind of risks inherent in clustered systems, and much better overall performance.

Benefits for Memory-Intensive Problem Spaces: AI/ML, Spatial, Graph, and More

Some advanced analytic data models, such as spatial and graph data, require tremendous amounts of data to operate. Artificial intelligence (AI) and, in particular, machine learning (ML) systems can require large amounts of data as well. Now, instead of forcing a mixed model of operation that necessarily involves clustering, data tiering, or both to apply such advanced analytics, it becomes possible to do much of the data processing for these advanced analytics data models in a single system. If the database is accessing Intel Optane DC PMM in the App Direct mode, the database can also ensure that, having built those data structures in memory, they will remain in case of failure and restart and will greatly speed up and simplify operations in case of planned downtime.

SAP HANA with Intel Optane DC Persistent Memory

SAP HANA is a memory-optimized relational database management system that offers true embedded analytic transaction capability. It keeps data in a columnar format, allowing the power of vector processing for complex analytic queries. At the same time, it uses a differential store to accept data changes from transactions and blend them into the columnar organization. As a result, SAP HANA delivers both high-speed transaction processing and powerful query processing and supports a blending of the two in a single database with only one instance of the data at any time. SAP HANA offers various data tiering capabilities to assign data to the right level of memory/storage to help customers balance volume, cost, and performance characteristics of their system. Data that is very frequently referenced and changed, along with various kinds of working memory, remain in DRAM; other frequently accessed and changed data is kept in PMEM. The SAP HANA–native storage extension capability was recently introduced to add another level of simplicity and agility for data tiering. Coupled with the use of Intel Optane DC PMM, SAP HANA provides robust capabilities to cost-effectively manage large data sets and mixed workloads.

According to both Intel and SAP, SAP HANA is the first major DBMS to optimize a data platform with the App Direct mode in Intel Optane DC PMM, delivering the full potential of capabilities. As a result, SAP HANA will be able to have higher memory capacity, persistent data storage in-memory, and overall reduced TCO because of simplified infrastructure and data management strategies.

More memory capacity means more data can be held in memory versus other persistent storage, including data that is frequently used (hot) and data that might be less frequently used (warm). This enables the ability to scale up versus scale out for larger data sets, which can help improve performance by eliminating network overhead and the complexity of managing scale-out scenarios. This results in simpler data tiering strategies, reduced infrastructure needs, and potential for improved performance on large data sets.

Persistent memory storage leads to significantly less time loading data from traditional persistent storage into memory (e.g., SDD). Today, SAP HANA doesn't need to wait until data is loaded into memory at start-up, and this advancement will enable the most optimal operations on data almost instantaneously. This enables rapid restart and recovery times immediately with the full performance benefits of SAP HANA.

SAP has worked closely with Intel for many years to ensure that SAP HANA is fully optimized, by leveraging App Direct mode, to exploit all the advantages that Intel Optane DC persistent memory has to offer. To take advantage of this new technology, companies will need to upgrade to the latest hardware from Intel and minimally running SAP HANA 2.3 or greater. Any application running on SAP HANA will then be able to take advantage of this technology. The good news is that no changes are required for the applications; they will work just as they are.

In short, the benefits offered by SAP HANA in conjunction with Intel Optane DC persistent memory are as follows:

- Increased memory capacity leading to faster responses to business queries
- Enabling advanced ML and predictive analytics in real time at a lower cost
- Simplified and increased scalability of solutions built on SAP HANA
- Continuous in-memory availability and performance due to in-memory persistency
- Lower TCO through hardware cost reductions, simplified data storage hierarchy, simplified technology landscape, and easier data life-cycle management

FUTURE OUTLOOK

It is reasonable to expect that many vendors will take advantage of Intel Optane DC persistent memory and will find a variety of new ways to take advantage of its capabilities, possibly including dramatic changes in database core architectures. SAP HANA is likely to continue to evolve in this way as well. In addition, it is reasonable to expect that SAP S/4HANA and other SAP business solutions will continue to be enhanced to take further advantage of high-speed, low-cost embedded analytics that are made possible through the PMEM capabilities of this technology.

CHALLENGES/OPPORTUNITIES

Many competing technologies are likely to emerge to take advantage of PMEM technology, offering challenges to SAP in both the database and applications software markets. At the same time, it seems likely that other memory chip manufacturers will develop their own PMEM capabilities to challenge Intel. Intel and SAP continue to innovate aggressively as they have for many years, not only with incremental changes to their current technologies but also with revolutionary new designs that push the boundaries of what is possible in processing large amounts of data rapidly, consistently, and robustly.

CONCLUSION

Many limitations exist in database management systems as a result of the fact that main memory is impermanent and limited. Compromises are necessary, including the use of flash and disk storage to augment DRAM in managing modern enterprise databases. PMEM promises to sweep away these limitations, allowing databases to operate at a scale and speed never before contemplated. Intel has advanced such a transformative technology with Intel Optane DC persistent memory. SAP has moved first to take full advantage of the capabilities afforded by this revolutionary new main memory capability in SAP HANA.

The terms of engagement in the database world have changed fundamentally. Everyone who designs, builds, and uses databases and their applications must be prepared for radical changes that will make their use of data more powerful, more cost effective, and more impactful on the enterprise.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-community.com
www.idc.com