Using Persistent Memory with your SAP HANA Database

Article

Last year, Lenovo, Intel and SAP worked closely together on launching Intel Optane Persistent Memory. SAP HANA 2 SP03 was the first database to fully utilize Optane's AppDirect mode, allowing the column store of HANA -- which is 95% of your HANA database -- to now reside in Persistent Memory. Lenovo ThinkSystem SR950 servers were used in the development labs, where realistic customer scenarios were tested using SAP Capture & Replay.

On top of the known benefits for faster restart in the case of business continuity requirements, it is possible to store more data of the column store part of SAP HANA in memory when using a combination of DRAM and persistent memory. Additionally, bigger SQL queries can be addressed to the database.

The following comparison shows the possible amount of data being stored in older generation hardware systems (in green) and current capabilities with the second-generation Intel Xeon Scalable processor technology ("Cascade Lake") (in blue). What's important to understand is the ratio between DRAM and persistent memory. The right three bars show possible ratios of 1:1, 1:2 and 1:4.

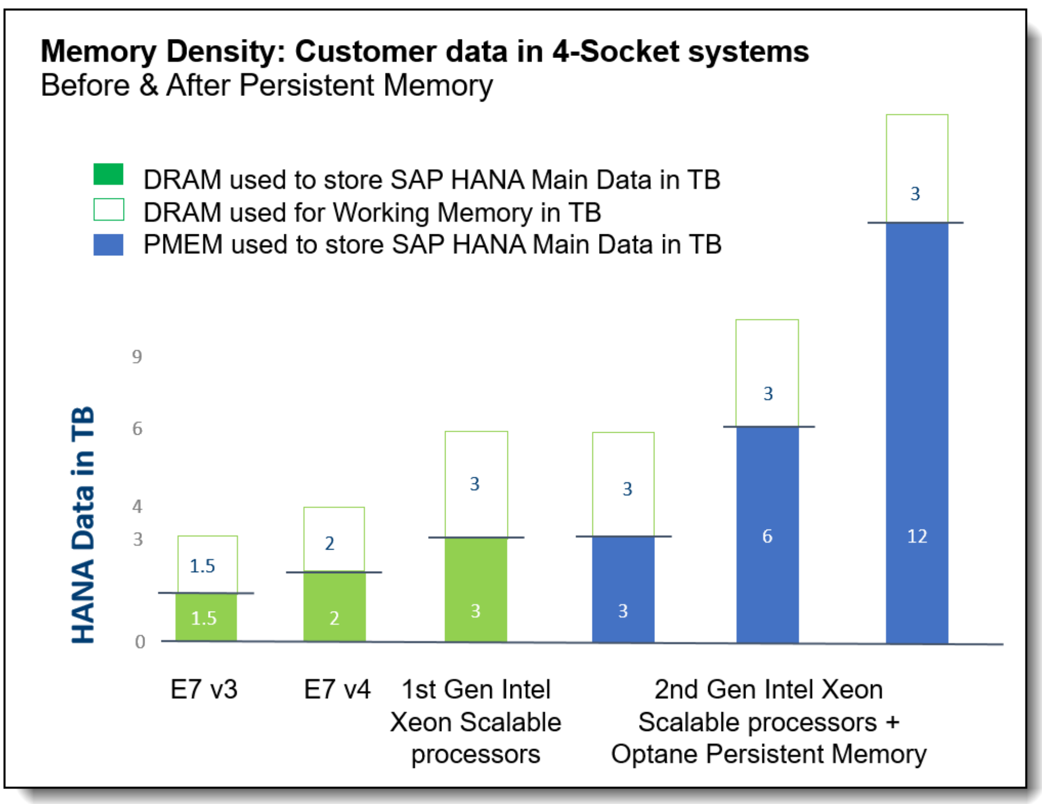

Figure 1. Intel Optane PMem and 2nd Gen Xeon Scalable Processors drive increased memory density

The appropriate ratio will be dependent on the business workload behavior and the layout of the table sizes in the customer environment. For that, an expert sizing exercise is required. Lenovo has a vast amount of experience with sizing SAP customer workloads. A sizing request form can be found on the Lenovo website. More documentation is available on the SAP One Support Launchpad in SAP Notes 2296920 and 2786237 for sizing, and Note 2700084 which covers FAQs. Intel has also published a configuration guide.

The following figure shows an example of the memory DIMM population in a 4-socket server for a transactional workload: on the left side with DRAM only (all 64 GB DIMMs) and on the right side a 1:2 mixture of DRAM and persistent memory. Each second-generation Intel Xeon Scalable processor has 12 DIMM slots out of which 6 are populated with 64 GB DRAM DIMMs and the other 6 are populated with 128 GB persistent memory DIMMs.

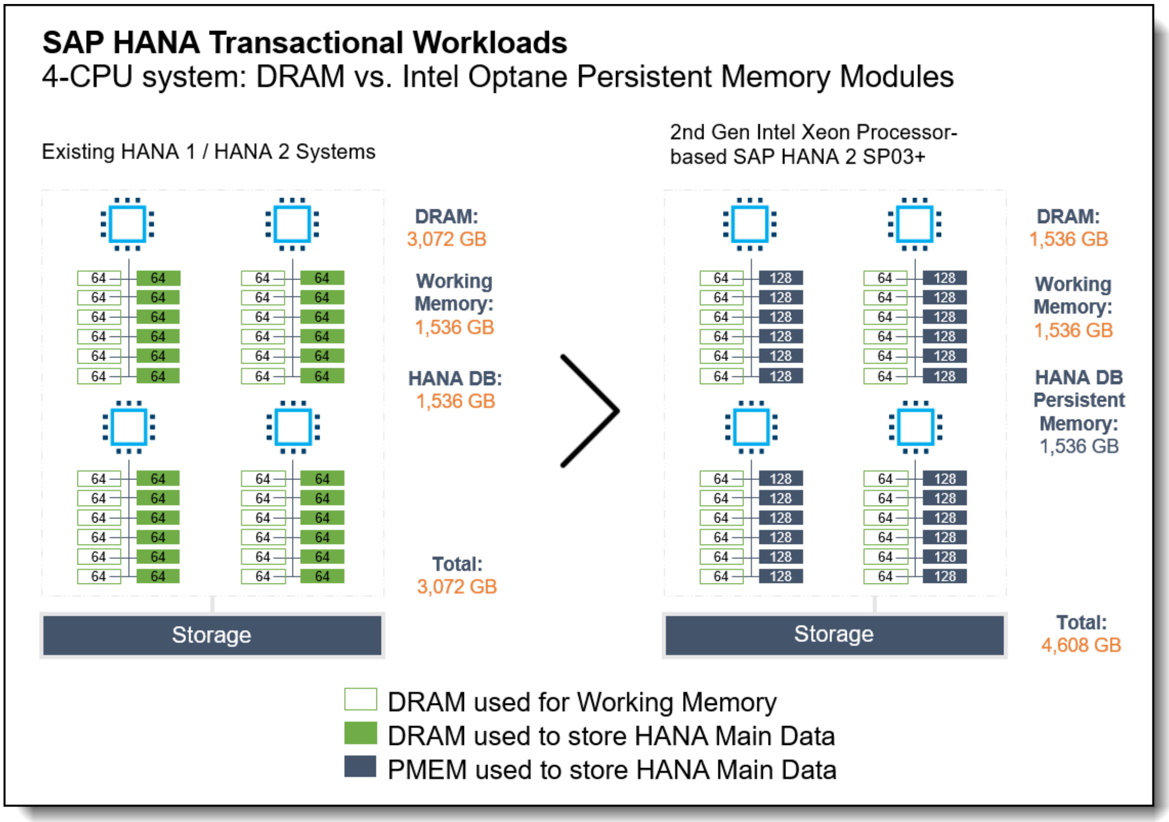

Figure 2. Comparison of DRAM vs. Intel Optane configurations for SAP HANA transactional workloads

Production Workload Tests

For the testing of SAP ECC on SAP HANA, we used the SAP Capture & Replay feature. Other tools like Parasoft or Load Runner are also applicable.

High Level Testing Steps

1. Upgrade HANA Cockpit to SP11 Patch 14 or the most recent version in the SAP Marketplace
2. Upgrade your OS to SLES 15 or higher, RHEL 7.6 or higher
3. Grant privileges to run capture in production
4. Identify peak load based on month end of quarter end peak processing (see the workload heatmap in the figure below)
5. Run a manual CAPTURE session in production with the option to automatically trigger a full backup
6. Perform a REPLAY in the non-production environment where you have the persistent memory setup

7. Review and compare the results

The following figure depicts the heat map based on the SAP Early Watch Alert. We chose a time when the most business users login to the system during month end close in the financial application.

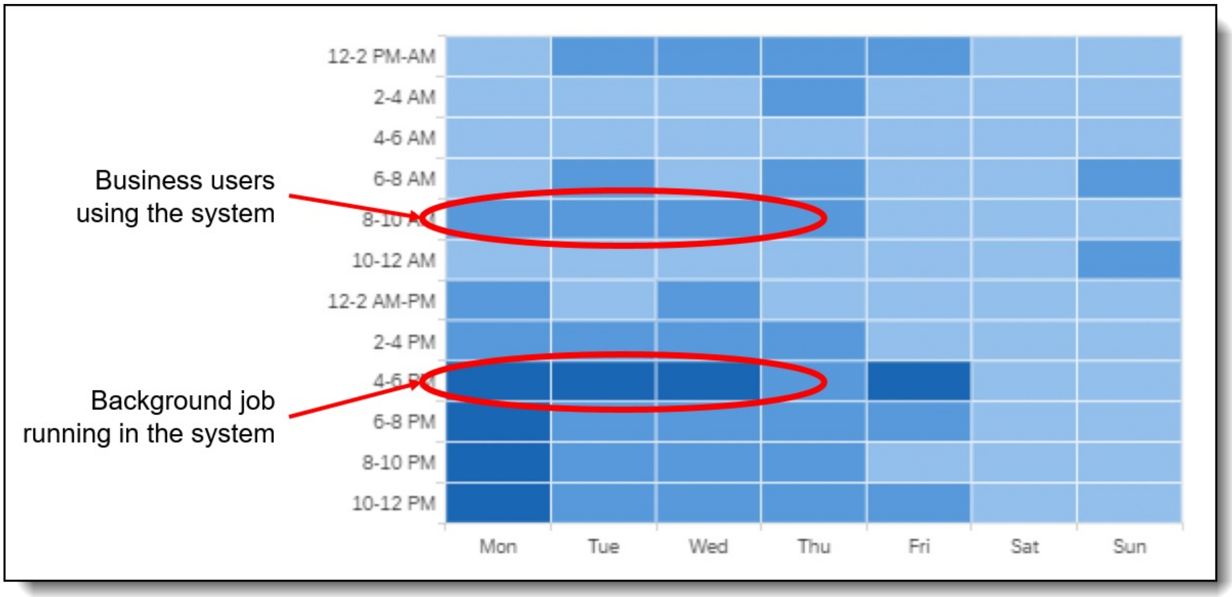

Figure 3. Heat map of workload activity

When the processes are replayed in the non-production environment, quite a few of the statements are skipped since they are a duplicate entry, or you have replayed the workload in an incorrect version of the backup.

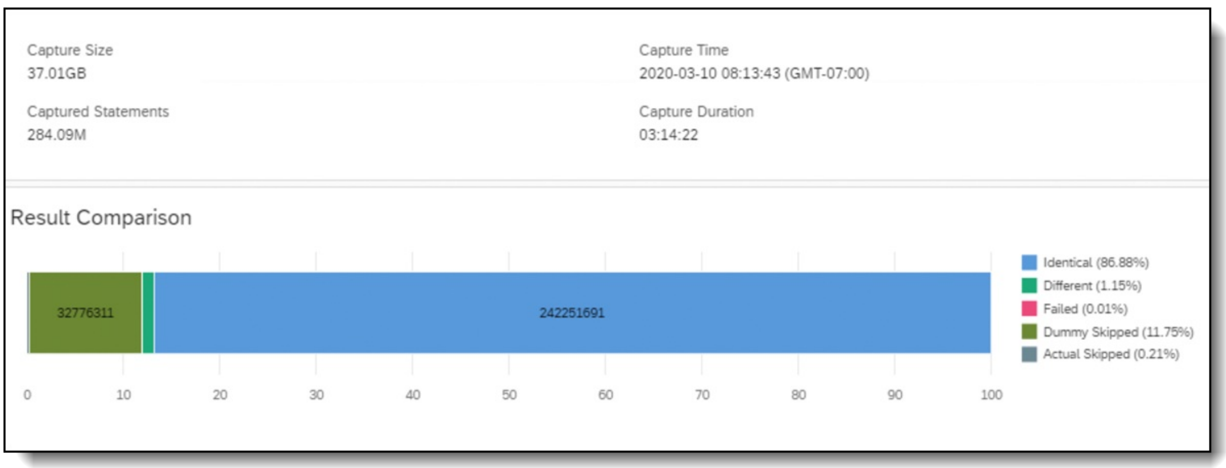

Figure 4. Quality of the Data Capture

When the capture is occurring, we see anywhere between 2-4% increase in CPU & memory resource utilization. We have captured multiple peak windows as well as depicted in the following graphic.

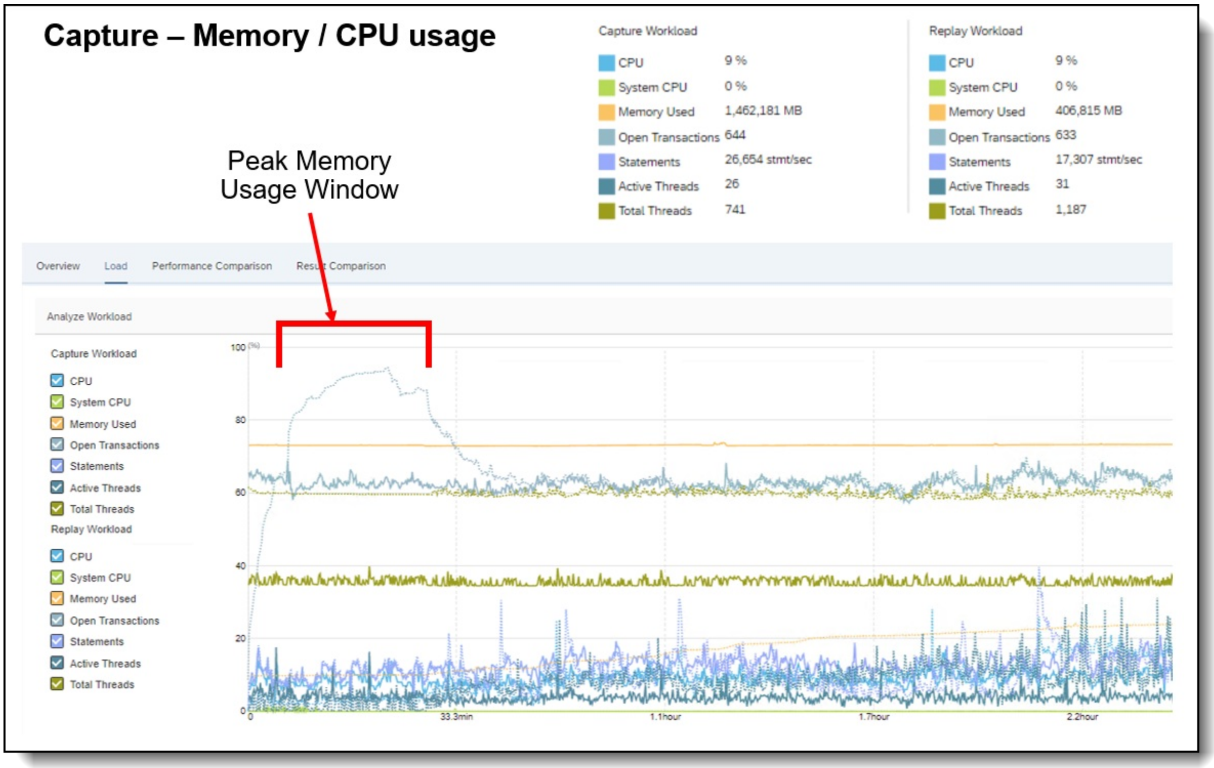

Figure 5. Capture of memory/CPU usage

Result Summary

The production workload capture was running on a Intel Xeon E5 v4 ("Broadwell") server which processed 18,200 SQL statements / second. We saw generational CPU improvement when we replayed it in the Persistent Memory system running on second-generation Intel Xeon Scalable ("Cascade Lake") processors which ran 19,600 SQL statements / second.

99.97% of the SQL statements ran at the same speed or better and in our case they were 8% better. We did not tune any queries to achieve the above results just an Apples to Apples comparison test across two processor generations.

Notices

This document, LP1347, was created or updated on June 18, 2020.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP1347
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP1347.

Trademarks