



Memory performance on HP Z840/Z640/Z440 Workstations

Performance impact of moving to larger DIMM sizes and populating fewer memory channels

Table of contents

Introduction	2
Memory performance in the HP Z840/Z640/Z440 Workstation.....	2
Application/Benchmark performance comparisons	4
HP Z440 comparison results.....	4
HP Z840 comparison results.....	6
Results discussion.....	7
Conclusion.....	8



Introduction

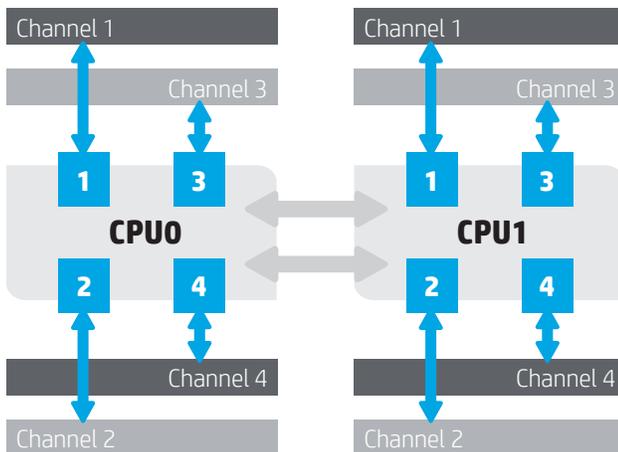
This paper looks at the performance impact of moving to larger DIMM sizes in conjunction with populating fewer memory channels, for specific configurations in HP Z840/Z640/Z440 Workstations.

As a general rule, memory performance is optimized when all memory channels are populated in a system. Moving from a balanced memory configuration (all memory channels populated) to an unbalanced memory configuration (not all memory channels populated) will result in sub-optimal memory performance. This performance impact will depend upon the application, workload and the specific configuration being considered.

Memory performance in the HP Z840/Z640/Z440 Workstation

Each processor in the HP Z840, HP Z640 and HP Z440 Workstation supports four DDR4 memory channels. Each memory channel provides a certain bandwidth capability, so by default the number of memory channels populated impacts the achievable memory bandwidth in a system. In addition to providing the greatest memory bandwidth capability, populating all memory channels (a balanced memory configuration) also allows the greatest interleaving of memory accesses among the channels.

Figure 1. Memory channels



Note:

In a dual processor NUMA configuration (default for Z840/Z640), the interleaving of memory accesses among the channels is on a per processor basis. With NUMA disabled, the interleaving of memory accesses among the channels would be across both processors.

Figure 2. Memory bandwidth comparison
Channels populated vs. cores accessing memory

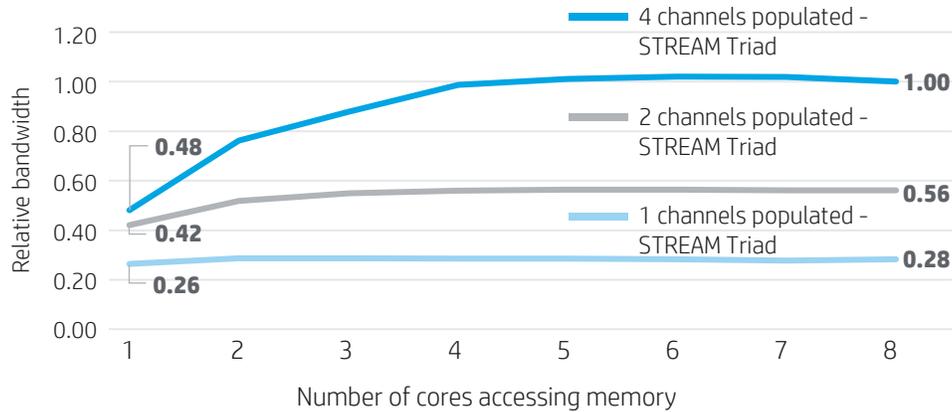


Figure 2 uses a STREAM subtest (Triad), to show how system memory bandwidth is impacted by the number of memory channels loaded. For those not familiar with STREAM, it is a synthetic benchmark that measures the sustainable memory bandwidth in a system. Figure 2 also shows how maximum memory bandwidth for a particular memory channel loading is impacted by the number of cores accessing memory. Looking at the case of all the memory channels loaded, it is apparent that the available memory bandwidth is not saturated until multiple cores are accessing memory simultaneously.

Reviewing this relationship a general rule can be inferred. Single or lightly threaded applications that do not require or use significant memory bandwidth will experience a smaller performance impact moving to larger memory DIMMs, in conjunction with populating fewer memory channels. Threaded applications that do require or use significant memory bandwidth will see a more significant performance impact moving to larger memory DIMMs, in conjunction with populating fewer memory channels.

Application/Benchmark performance comparisons

Application/benchmark performance comparisons were done on specific HP Z840/Z440 memory configurations, evaluating the performance impact of moving to larger DIMM sizes in conjunction with populating fewer memory channels. The comparisons look at moving from 4 GB DIMM configurations to 8 GB DIMM configurations, keeping the total memory capacity equivalent. Similar performance results would be anticipated moving from 8 GB DIMM configurations to 16 GB DIMM configurations, again keeping the total memory capacity equivalent.

The HP Z840 comparison results are representative of a dual processor HP Z640 configuration, while the HP Z440 comparison results are representative of a single processor configuration in either an HP Z640 or HP Z840 system.

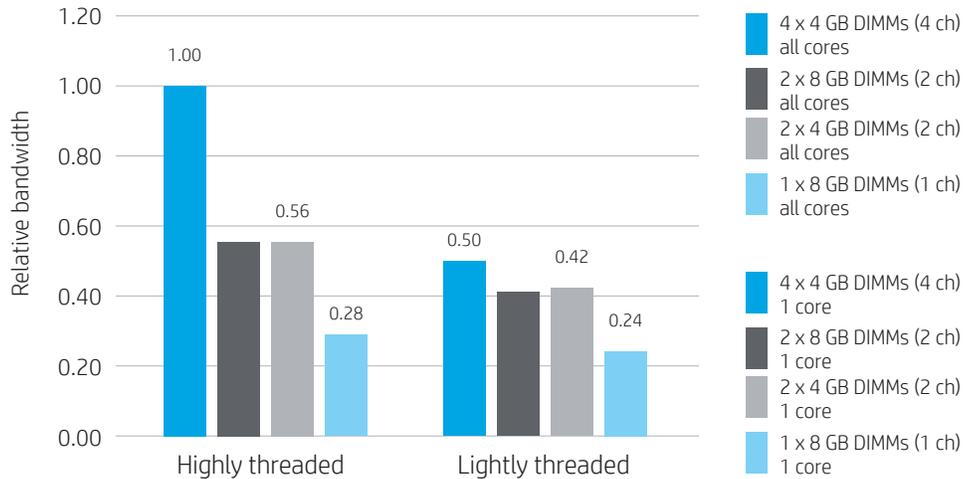
Table 1. Memory configurations evaluated on the HP Z440 and Z840

System	Processor	Total memory	Baseline configuration	Comparison configuration
HP Z440	Intel® Xeon® E5-1680v3	16 GB	4 x 4 GB DDR4-2133 MHz (4 channels loaded - balanced configuration)	2 x 8 GB DDR4-2133 MHz (2 channels loaded - unbalanced configuration)
HP Z440	Intel® Xeon® E5-1680v3	8 GB	2 x 4 GB DDR4-2133 MHz (2 channels loaded - unbalanced configuration)	1 x 8 GB DDR4-2133 MHz (1 channel loaded - unbalanced configuration)
HP Z840	Intel® Xeon® E5-2687W	32 GB	8 x 4 GB DDR4-2133 MHz (4 channels/proc loaded - balanced configuration)	4 x 8 GB DDR4-2133 MHz (2 channels/proc loaded - unbalanced configuration)
HP Z840	Intel® Xeon® E5-2687W	16 GB	4 x 4 GB DDR4-2133 MHz (2 channels/proc loaded - unbalanced configuration)	2 x 8 GB DDR4-2133 MHz (1 channel/proc loaded - unbalanced configuration)

As a reference point, figure 3 shows the memory bandwidth relationship between a balanced configuration (all memory channels loaded) and an unbalanced configuration (not all memory channels loaded), both in the lightly threaded case and in the highly threaded case.

The results across the applications and workloads tested show some cases where the performance impact of memory channel loading is small and other cases where the performance impact is similar to that shown in figure 3.

Figure 3. HP Z440 memory performance Stream (Triad)



HP Z440 comparison results

The Comparison 1 results shown in figures 4, 5, and 6 below are for 16 GB configurations with 4 channels versus 2 channels loaded. The Comparison 2 results are for 8 GB configurations with 2 channels versus 1 channel loaded. Results are normalized to the 16 GB, 4 channels loaded configuration. This is done to highlight the performance impact due to channel loading.

In the below comparisons, result variation from 2-3% is considered run-to-run variation and is not considered a true performance variation.

Figure 4. HP Z440 Product Development

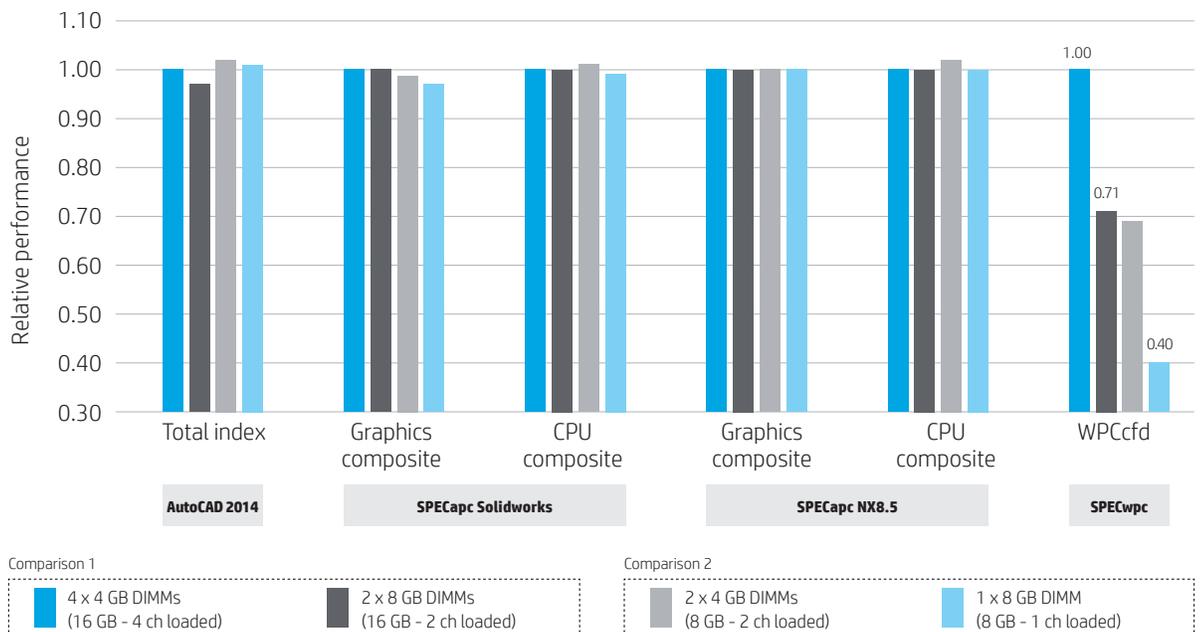


Figure 5. HP Z440 Media and Entertainment

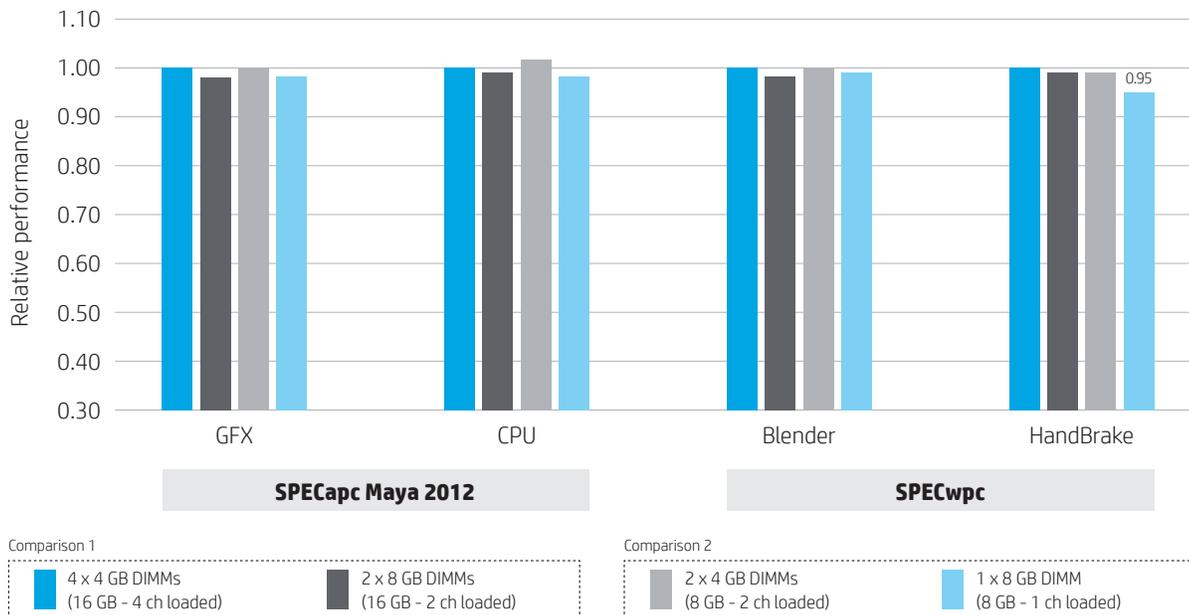
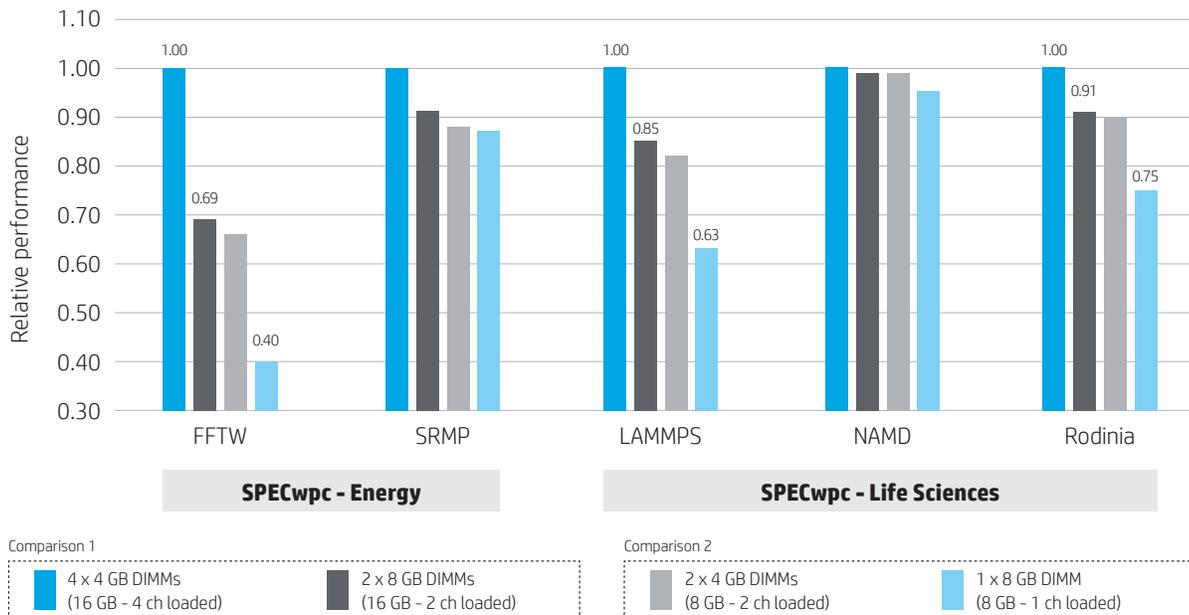


Figure 6. HP Z440 Energy and Life Sciences



HP Z840 comparison results

The Comparison 1 results in figures 7, 8, and 9 below are for 32 GB configurations with 4 channels versus 2 channels loaded per processor. The Comparison 2 results below are for 16 GB configurations with 2 channels versus 1 channel loaded per processor. Results are normalized to the 32 GB, 4 channels loaded per processor configuration. This is done to highlight the performance impact due to channel loading.

Again, in the below comparisons, result variation from 2-3% is considered run-to-run variation and is not considered a true performance variation.

Figure 7. HP Z840 Product Development

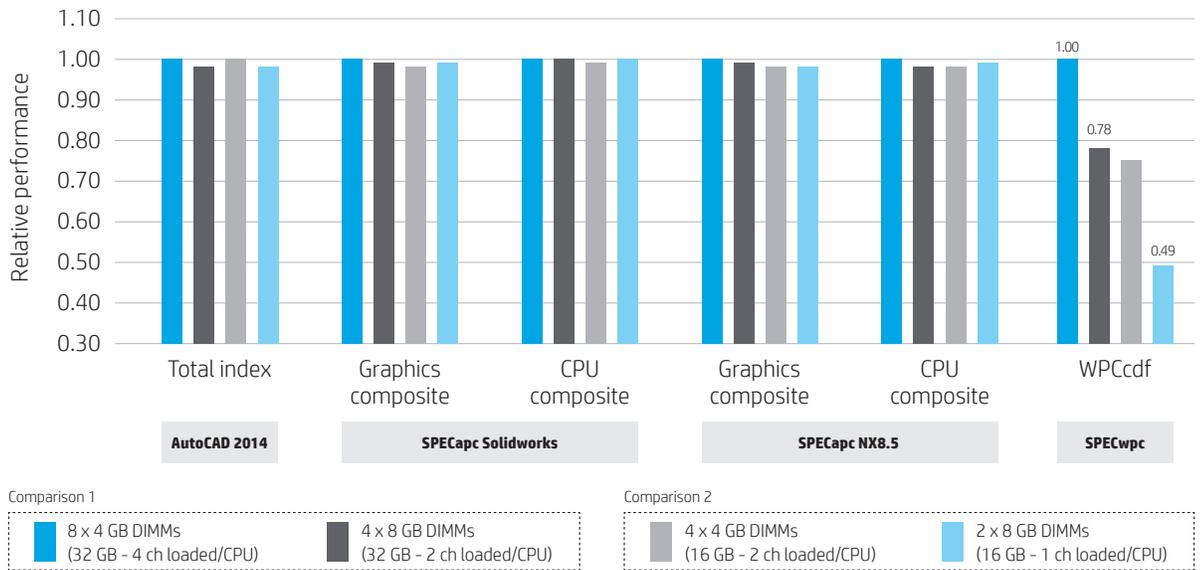


Figure 8. HP Z840 Media and Entertainment

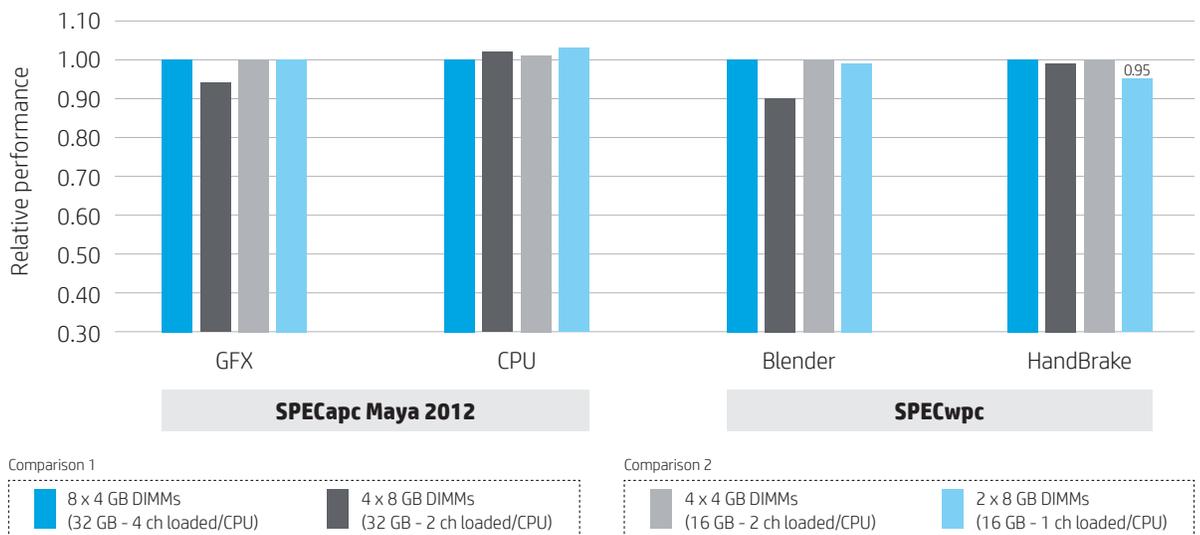
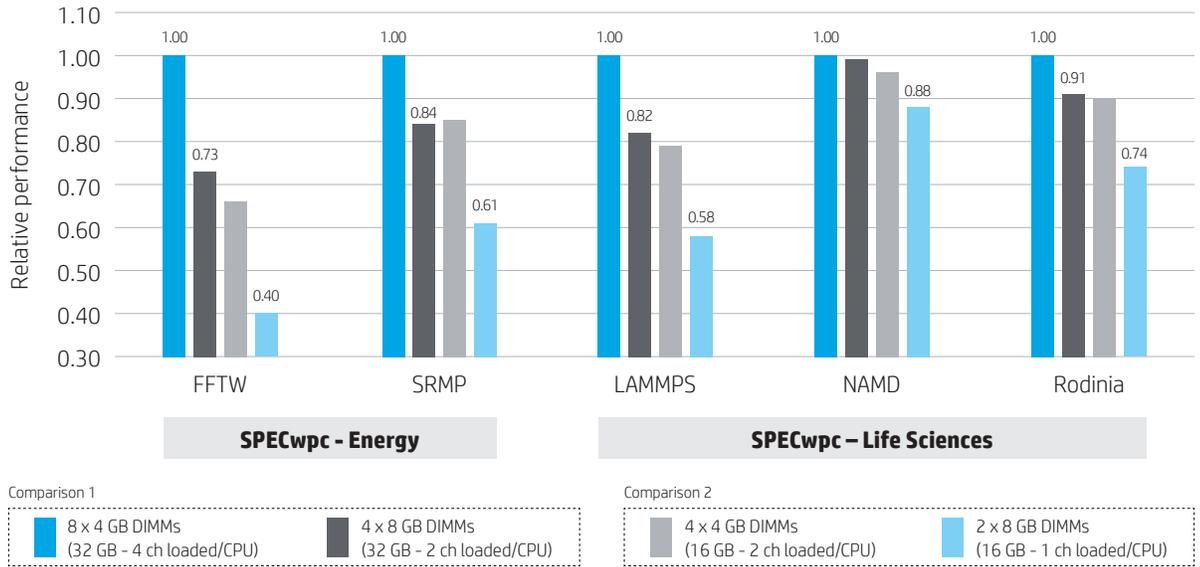


Figure 9. HP Z840 Energy and Life Sciences



Results discussion

The comparison results shown above hold to the general rule: single or lightly threaded applications and workloads that do not require or use significant memory bandwidth will experience a smaller performance impact moving to larger memory DIMMs, in conjunction with populating fewer memory channels. This can be seen in some of the Product Development results, namely Autodesk® AutoCAD, Solidworks and NX. These applications are lightly threaded and the particular workloads used for these runs use very little of the available memory bandwidth, so any performance difference seen due to the memory channel loading would be predicted to be small.

Of the Media and Entertainment applications and workloads evaluated, HandBrake is one that showed sustained use of all the available cores and consistent, although not high, use of memory bandwidth. You can see that HandBrake performance begins to be noticeably impacted by memory channel loading when there is only a single memory channel loaded per processor.

In contrast, the WPCcfd results seen under the Product Development results show a significant performance impact moving to larger memory DIMMs while populating fewer memory channels. This application and workload has sustained use of all the available cores and uses significant memory bandwidth. This can also be seen in the FFTW, SRMP, LAMMPS and Rodinia results shown under the Energy and Life Sciences benchmarks.

Conclusion

For the best memory performance, it is necessary to populate all memory channels on the processors. However, there are applications and workloads where the performance impact from populating fewer memory channels will be less apparent.

As a general rule, single or lightly threaded applications that do not require or use significant memory bandwidth will experience a smaller performance impact moving to larger memory DIMMs, in conjunction with populating fewer memory channels. Threaded applications that do require or use significant memory bandwidth will see a more significant impact moving to larger memory DIMMs, in conjunction with populating fewer memory channels.

Resources, contacts, or additional links

hp.com/go/whitepapers

Learn more at

hp.com/zworkstations

Sign up for updates
hp.com/go/getupdated



Share with colleagues



Rate this document

© Copyright 2015 Hewlett-Packard Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

Intel and Xeon are trademarks of Intel Corporation in the U.S. and other countries. Autodesk and AutoCAD are registered trademarks or trademarks of Autodesk, Inc., and/or its subsidiaries and/or affiliates in the USA and other countries.

4AA5-6052ENW, January 2015

