

14G with Skylake – how much better for HPC?

Dell EMC HPC Innovation Lab, September 2017

This document describes the features of the latest Dell EMC servers and compares performance to previous generation systems for a variety of HPC applications.

New servers and Skylake

Dell EMC recently announced the [14th generation PowerEdge server portfolio](#) (14G) which supports the latest generation Intel Scalable Processor Family (the micro-architecture that is code named “Skylake”). In addition to the latest processor and accelerator/GPU support, 14G includes several other technology additions. These servers have enhanced systems management and security features via iDRAC9, can support up to 24 NVMe drives per server, include support for NVDIMM and future 3D XPoint memory technologies, allow options for direct contact liquid cooling within the server, etc.

This document is focused on the performance gains available with the latest generation Intel Skylake CPUs in Dell EMC 14G platforms. The Skylake CPU (SKL) supports six DDR4 memory channels per socket with memory modules that can run at up to 2667 MT/s. SKL provides up to 28 cores per socket with TDP up to 205 W per socket. Intel has introduced AVX512 instructions on SKL, and this doubles the floating point capabilities of SKL over previous generation Xeons. The processor now supports 512 bit registers, and the Platinum 8100 and Gold 6100 SKL CPU models have two fuse multiply add (FMA) units, each of which can execute 8 double precision calculations per cycle. With two floating point operations per FMA instruction, Skylake can execute 32 FLOP/cycle, double the previous generation Xeon which was 16 FLOP/cycle. Note that some models of the SKL CPU like the Gold 5100, Silver 4100 and Bronze 3100 CPUs have one FMA unit, giving 16 FLOP/cycle. As [before](#), the CPU frequency will depend on whether the code has a high density of AVX2 or AVX512 instructions. An application will run at faster CPU clock speeds when running non-AVX codes, and will run slowest with a high density of AVX512 instructions.

Other changes in SKL include 48 PCIe lanes per socket vs. 40 lanes previously, and a new interconnect called UPI (Ultra Path Interconnect) between the sockets, replacing the previous QPI interconnect. UPI can operate at up to 10.4 GT/s, faster than the 9.6 GT/s with QPI. There are other architectural changes like a larger L2 cache for the cores, a non-inclusive L3 cache, a new uncore interconnect, distributed home agent, optimized turbo bins, per core P-states, etc. Architectural changes in the silicon lead to new tuning options in the BIOS and one of these, Sub NUMA Clustering, is discussed in [this](#) blog.

We focus on measuring full system performance and compare 14G compute centric performance to the previous generation Dell EMC platforms. Some of the performance improvements are due to faster memory, some due to AVX512, some due to additional cores and some due to the combination of all the Intel micro-architecture enhancements. We show results for up to six generations of Intel processors and four generations of Dell EMC servers. The storage system and I/O is not a significant portion in these tests.

The shorthand used in the graphs below is explained here.

- 11G WSM – Dell EMC 11th generation servers with support for Intel Xeon 5600 series processors, micro-architecture code named Westmere (WSM).
- 12G SB – Dell EMC 12th generation servers with support for Intel Xeon 2600 series processors, micro-architecture code named Sandy Bridge (SB).
- 12G IVB – Dell EMC 12th generation servers with support for Intel Xeon 2600 v2 series processors, micro-architecture code named Ivy Bridge (IVB).
- 13G HSW – Dell EMC 13th generation servers with support for Intel Xeon 2600 v3 series processors, micro-architecture code named Haswell (HSW).
- 13G BDW – Dell EMC 13th generation servers with support for Intel Xeon 2600 v4 series processors, code named Broadwell (BDW).
- 14G SKL – Dell EMC 14th generation servers with support for Intel Xeon Scalable Processor Family, micro-architecture code named Skylake (SKL).

System benchmark performance – STREAM and HPL

We start our performance study with two standard system benchmarks, [STREAM](#) to measure the system memory bandwidth and [HPL](#) to measure compute performance.

Figure 1 shows the memory bandwidth capability of a two socket (2S) Skylake based system for two CPU models, the mid-tier Gold 6150 and the highest end Platinum 8180. It also plots system memory bandwidth for previous generations going as far back as the Dell EMC 11G servers with Westmere (Intel Xeon X5600). In all tests, memory was configured ideally for the platform under test. That is, each data point in Figure 1 represents the best memory bandwidth for that generation of CPU using an ideal memory configuration for HPC for that CPU micro-architecture.

On a dual socket Skylake server with six memory channels per socket, a balanced memory configuration dictates that all memory channels be populated and populated identically. This provides configuration choices of 12 DIMMs with one DIMM per channel (DPC) or 24 DIMMs with 2 DPC. 24 DIMMs is the maximum number of memory slots in a 2S Skylake server.

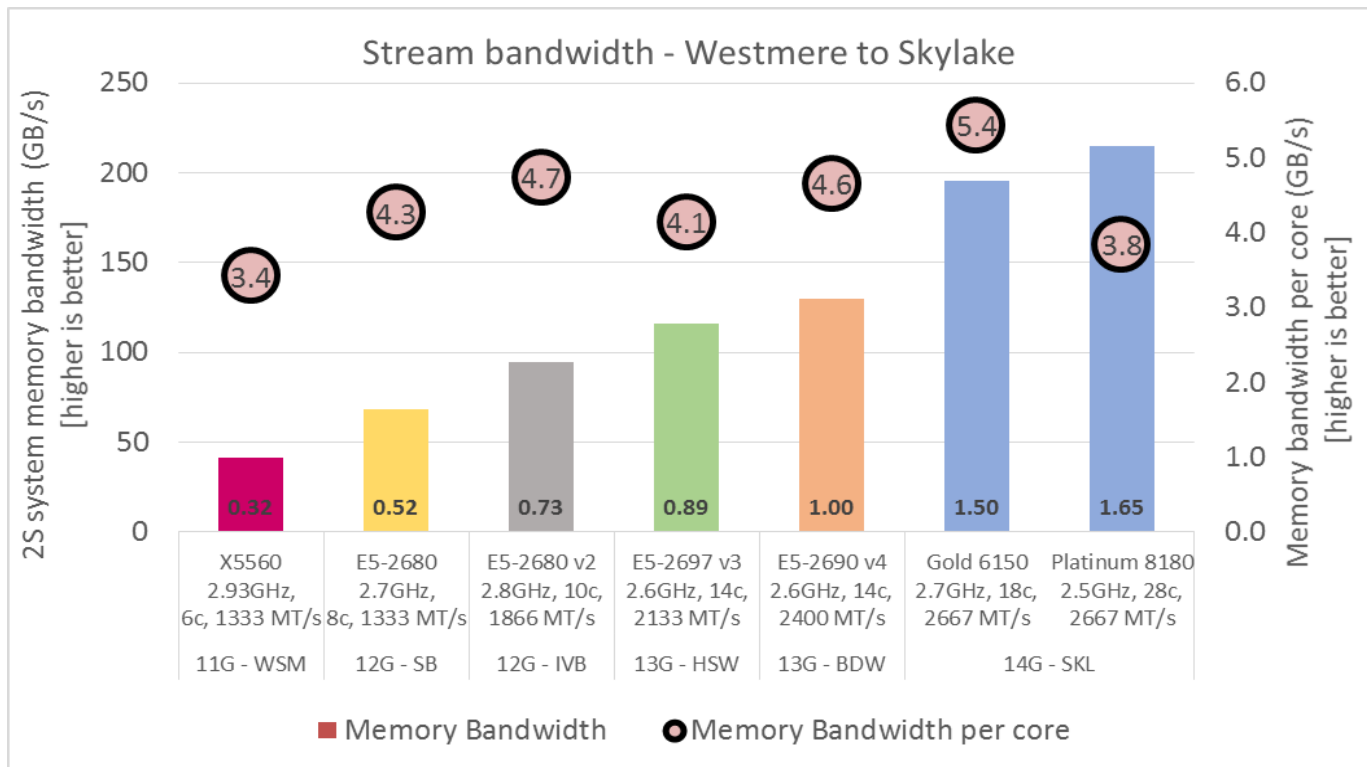


Figure 1 - Stream performance over generations

The height of the bars in Figure 1 gives the memory bandwidth of a 2S system, the data labels show performance relative to BDW and the circles show memory bandwidth per core. The measured system memory bandwidth has increased by 50-65% on SKL compared to the previous generation BDW, mainly due to the increased number of memory channels (50% more channels than BDW) and the faster memory speed (11% faster memory than BDW). Looking at memory bandwidth per core, the SKL 6150 provides a comfortable 5.4 GB/s per core. The 8180 CPU has the maximum number of SKL cores and even with 214 GB/s in the system, the memory bandwidth per core drops below 4 GB/s to 3.8 GB/s/core, which might not be ideal for some HPC applications.



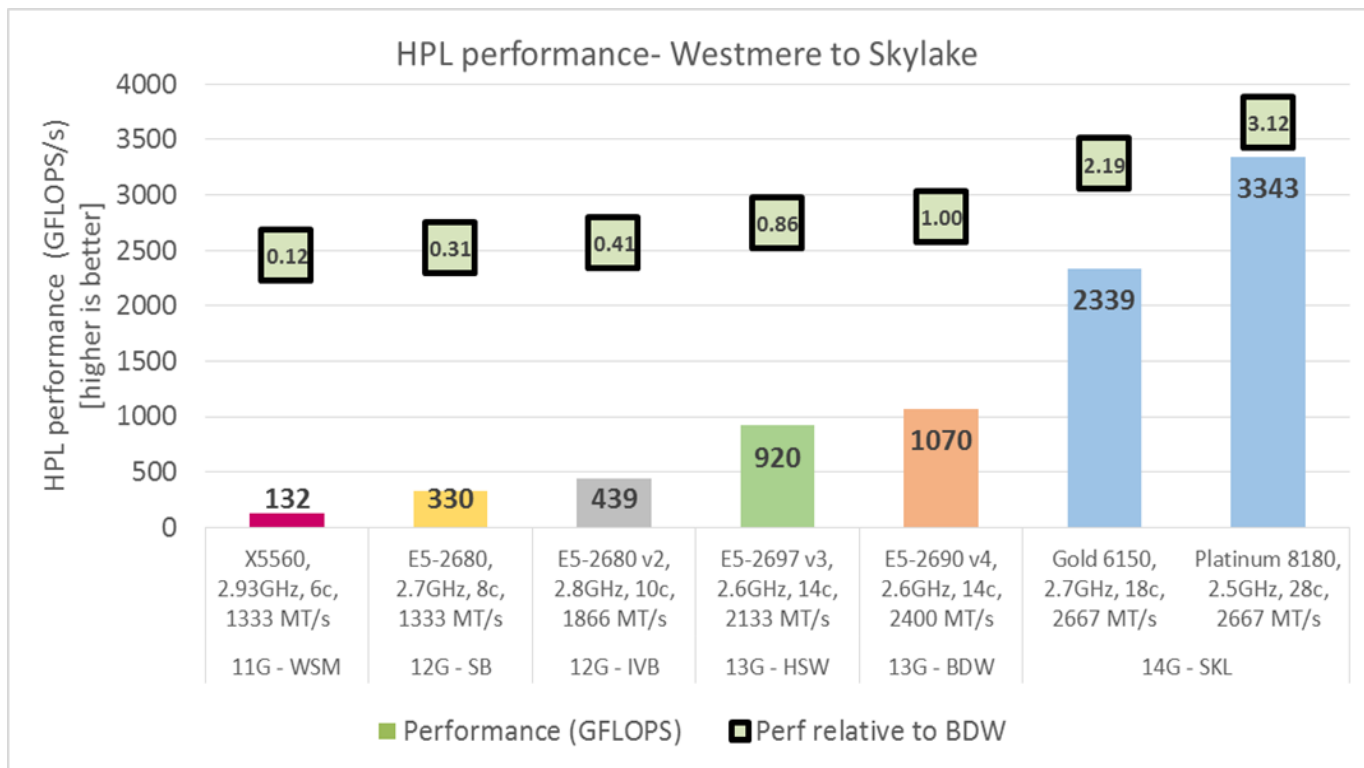


Figure 2 - HPL performance over generations

Moving on to HPL, Figure 2 shows the compute performance of a server over the last six generations. Similar to the Stream results, data is plotted from Intel Xeon X5600 series (Westmere) to Intel Scalable Processor Family (Skylake) and shows the performance of the system when all cores are in use. The bars show HPL performance for a 2S system and the squares show performance relative to BDW. A single Skylake dual socket system provides 3.3 TFLOPS of measured performance! Skylake systems perform 2x to 3x faster than the previous generation servers, and this is mainly due to AVX512 instructions and the associated increase in floating point operations per cycle. HPL performance also depends on the number of cores, and the increased number of cores is the second reason for increased performance with the top bin SKL models.

Weather modelling – WRF

[WRF](#) performance for two benchmark datasets, Conus 12km and Conus 2.5km is plotted in Figure 3. Conus 12km results are plotted for six generations and Conus 2.5km data for four generations as we started running tests with Conus 2.5km in the Ivy Bridge timeframe. Performance on SKL is about 50% faster than Broadwell with the Gold 6150 processor, and ~70% faster with the highest end Platinum 8180 processor. WRF is well parallelized and uses AVX instructions, and the additional cores helps Skylake performance. SKL also provides more memory bandwidth, which is useful for WRF. For example, comparing the 18 core 6150 processor to 14 core E5-2690 v4, the SKL processor has 29% more cores, about the same clock speed and 11% faster memory with equivalent memory bandwidth per core, and WRF performs 50% faster. SKL efficiencies are better than BDW in terms of better turbo residency. On the 8180, the SKL has 100% more cores than BDW, but the clock speed is lower. Although the memory bandwidth per core on 8180 is lower than the particular BDW SKU tested, SKL still records a 70% performance improvement over BDW.

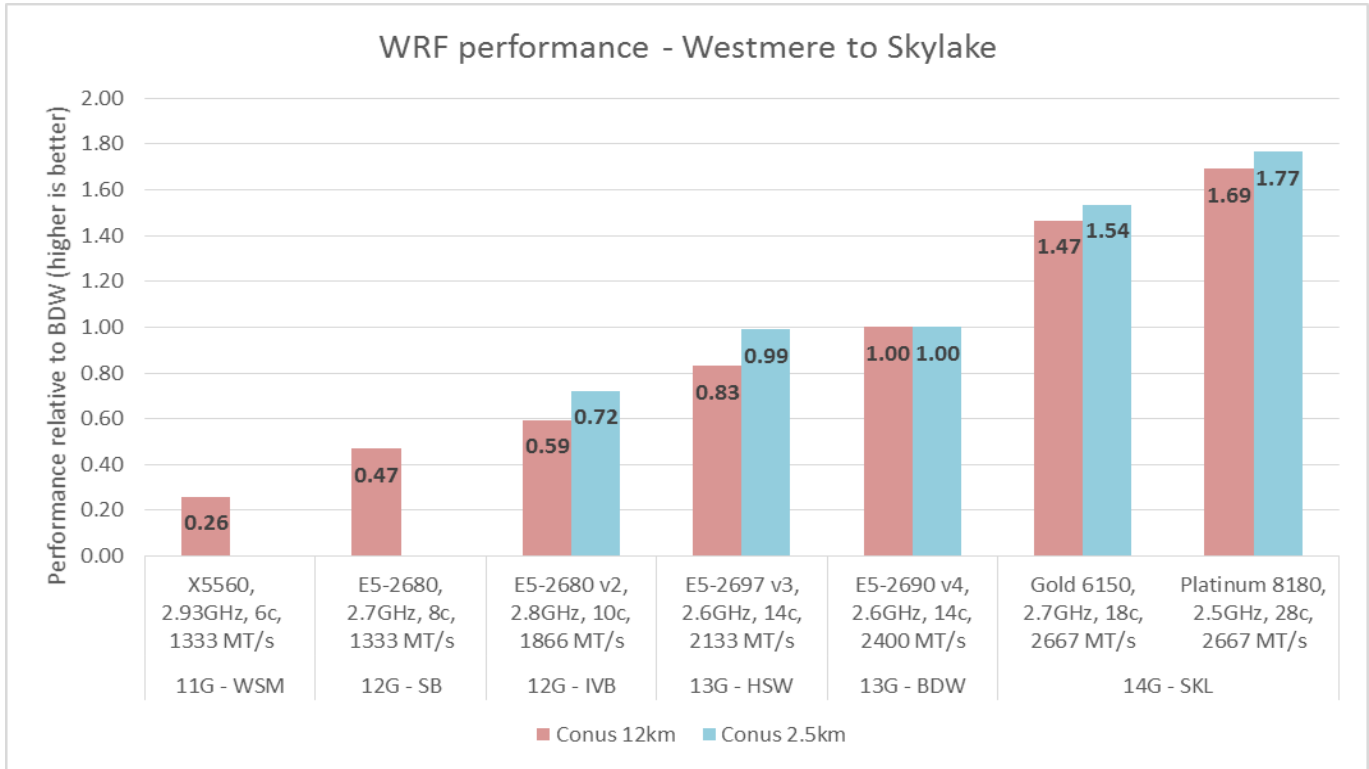


Figure 3 - WRF performance over generations

Life Sciences – BWA Pipeline

Alignment of raw reads to a reference genome is one of the key steps in next generation sequencing data analysis, and BWA is a software package for mapping low-divergent sequences against a large reference genome, such as the human genome. BWA performance for 416 million reads sequencing data is shown in Figure 4. Skylake Gold 6148 and Platinum 8160 are compared to the previous generation E5-2697 v4 (Broadwell) on dual socket systems. The height of the bars in the figure represent relative speed-ups against single core Broadwell performance and show that all three CPUs scale well in terms of the number of cores used for BWA.

This graph shows that, for the same number of cores, SKL is ~10% faster than BDW. Since SKL systems have more cores than BDW for the same rack space, the throughput on SKL systems can improved much more than 10%; this is explained below.

The results of throughput tests are shown in Figure 5. The throughput tests are performed on a single compute node. The E5-2697 v4 system uses Lustre as storage, while the 14G Skylake systems are equipped with local disks (4 SAS drives in RAID 0 configuration). The performance difference between Lustre and local storage is not a big factor for the small input size here. In order to obtain maximum throughput of each system, the number of samples per system is increased and processed concurrently. Hence, the maximum throughput is determined by the highest number of samples that can be processed simultaneously on each system. The height of bars represents the number of billion reads per hour, and the red solid line shows the normalized maximum throughput achieved by each system based on the maximum throughput of E5-2697 v4. Since higher number of cores are available for Skylake, higher throughput from Skylake is expected. Efficiency of Skylake allows Gold 6148 to process 26% more data with only 11% more cores.



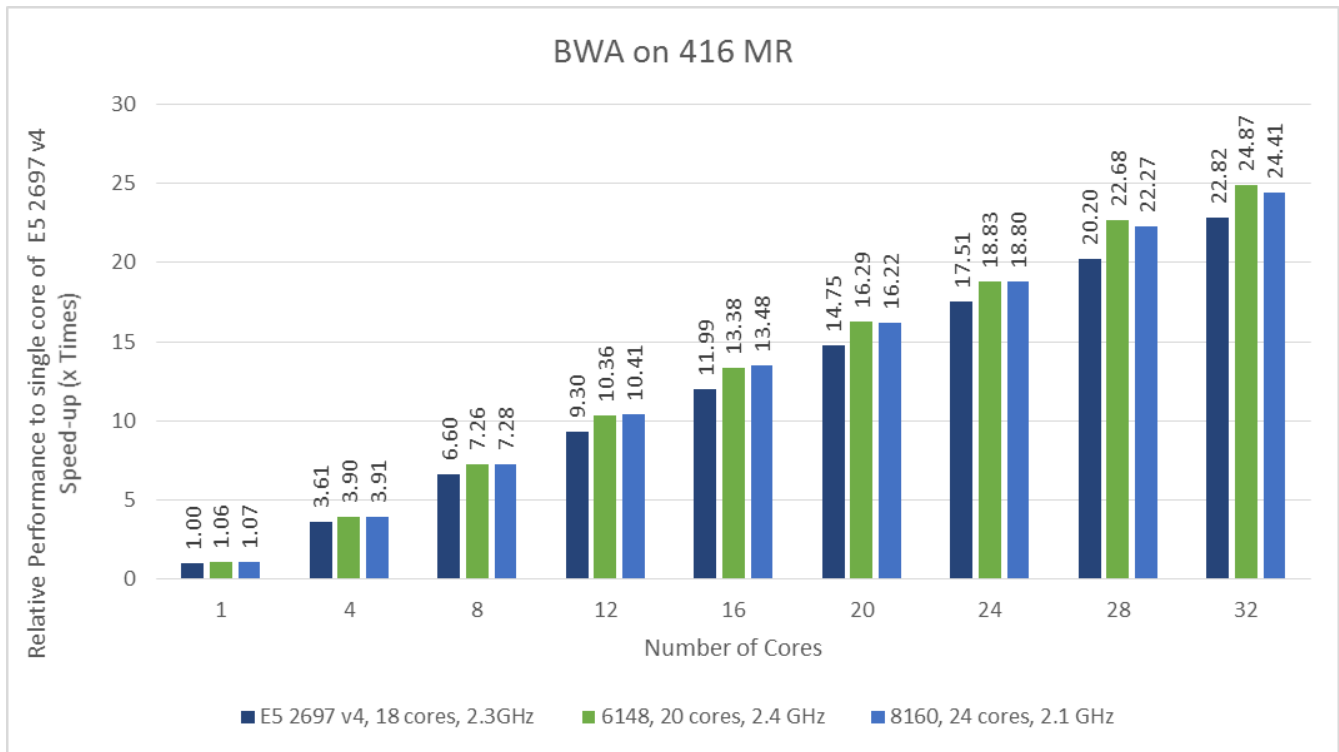


Figure 4 - BWA performance over generations. The height of the bar represents relative performances against a single core of E5 2697 v4.

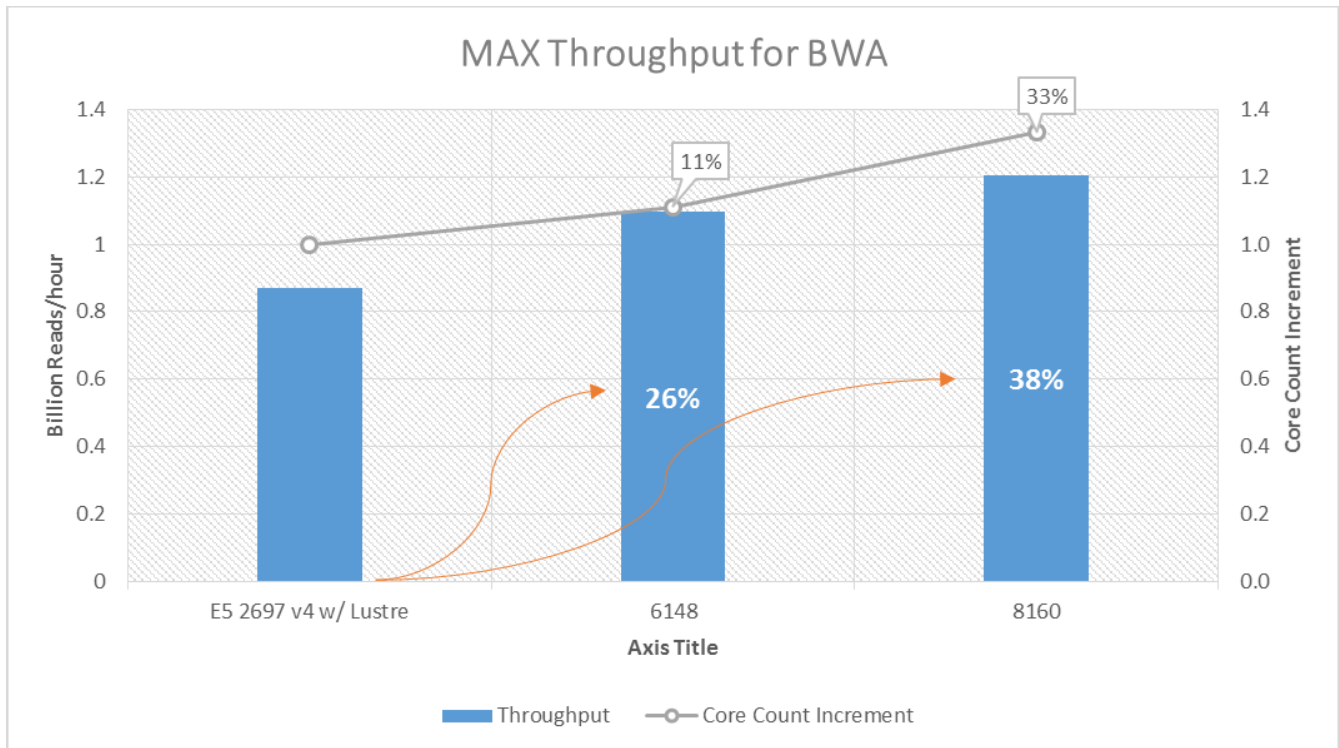


Figure 5 - BWA throughput over generations



Manufacturing applications – ANSYS Fluent, STAR-CCM+ and LS-DYNA

ANSYS® Fluent® and STAR-CCM+® software are computational fluid dynamics applications commonly used for research and product development. LS-DYNA® is a general-purpose finite element software in use by multiple industries. The system performance for these three applications, across four processor generations, is presented here. Standard benchmark datasets are used for the performance comparisons.

The performance for ANSYS® Fluent® for the aircraft_wing_14m dataset across four processor generations is plotted in Figure 6. In this figure, the height of the bars and the label represents the 2S system performance relative to the 13G Intel Xeon E5-2697A v4 result. For this application and dataset, the tested Skylake processors range from 23% to 62% faster than the 13G Broadwell system.

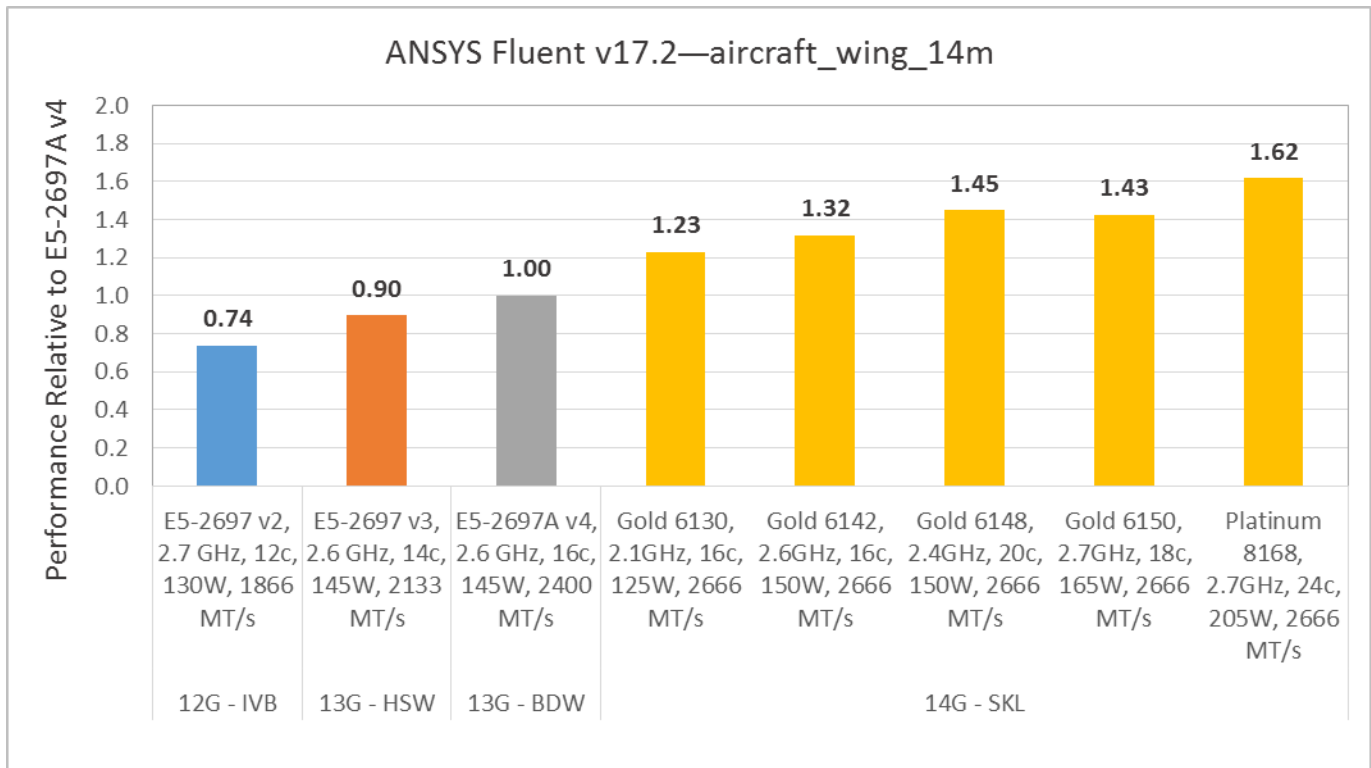


Figure 6 - Fluent performance over generations

The performance for STAR-CCM+® for the VtmUhoodFanHeatx68m dataset across four processor generations is plotted in Figure 7. Similar to the Fluent results, in this figure, the height of the bars and the label represents the 2S system performance relative to the 13G Intel Xeon E5-2697A v4 result. For this application and dataset, the tested Skylake processors range from 10% to 58% faster than the 13G Broadwell system.



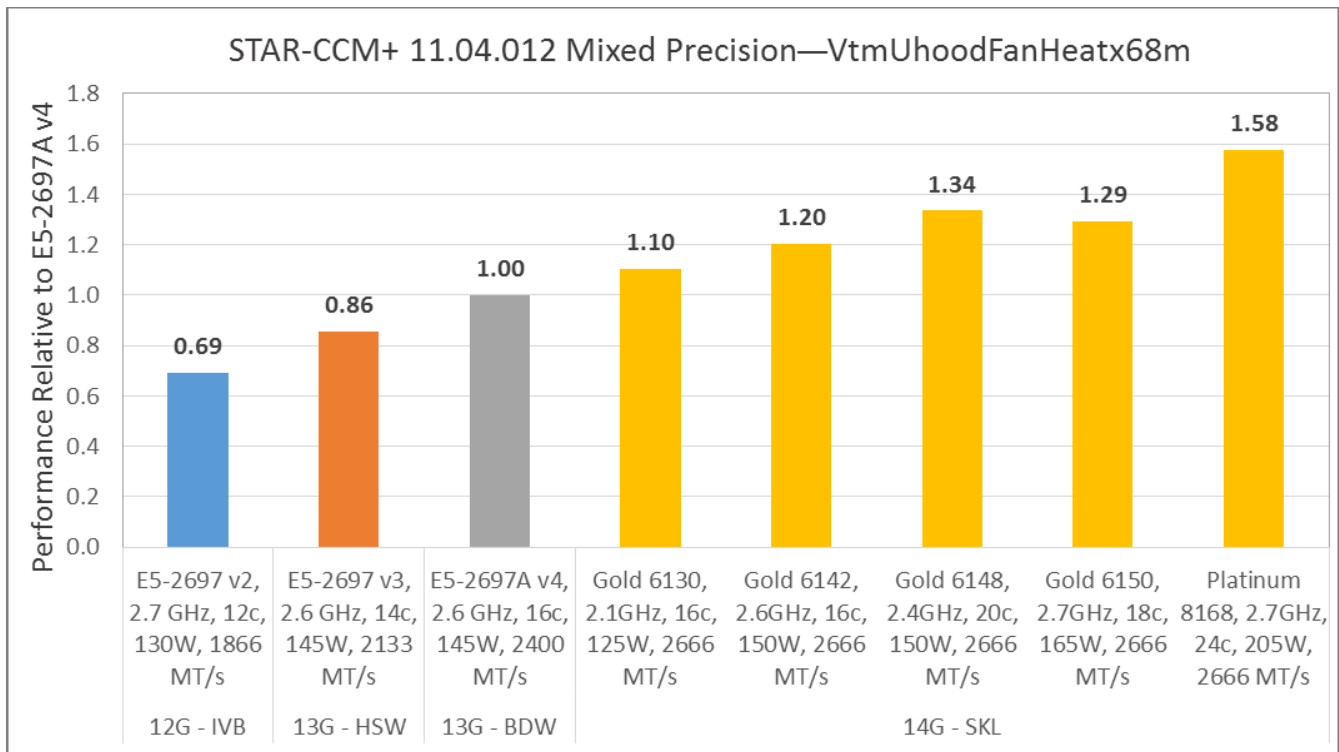


Figure 7 - STAR-CCM+ performance over generations

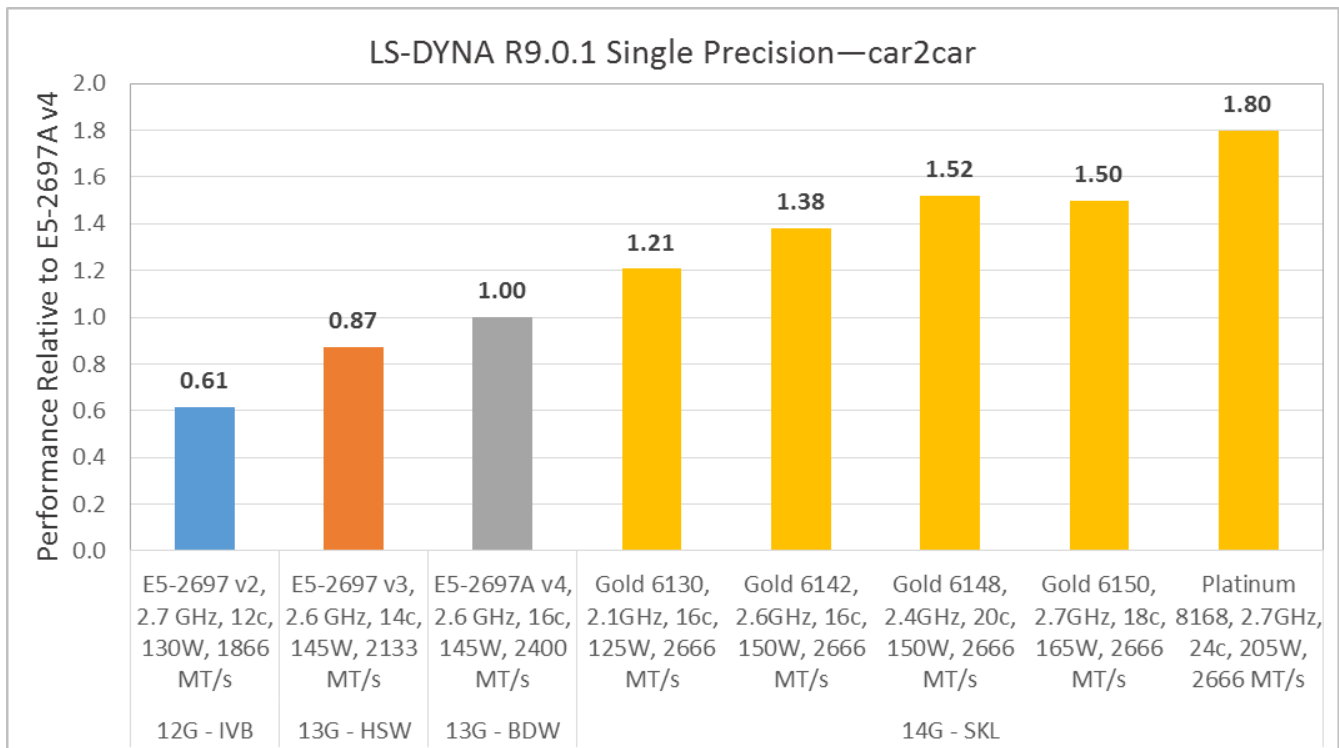


Figure 8 - LS-DYNA performance over generations



The performance for LS-DYNA® for the Car2Car dataset across four processor generations is plotted in Figure 8. In this figure, the height of the bars and the label represents the 2S system performance relative to the 13G Intel Xeon E5-2697A v4 result. The LS-DYNA AVX2 binary was used for Haswell through Skylake results, while the SSE2 binary was used for Ivy Bridge, since Ivy Bridge does not support the AVX2 instruction set. For this application and dataset, the tested Skylake processors range from 21% to 80% faster than the 13G Broadwell system.

For the tested digital manufacturing applications, 14G servers with Skylake can offer a significant performance advantage relative to Broadwell systems, ranging from 20% to 40% for processors similar core counts.

Conclusion

We expect the Gold series of processors to be ideal for most HPC use cases, providing a good balance between number of cores, memory bandwidth per core, compute capability, CPU power, and performance per core, as well as value in terms of price-for-performance.

On manufacturing applications, 14G servers provide 2x performance improvement over 12G (Ivy Bridge) based systems when using Xeon Gold CPUs, and up to 3x performance improvement over 12G (Ivy Bridge) when using Platinum CPUs.

Comparing to the previous generation 13G (Broadwell) systems, 14G with Gold CPUs provides 30-50% performance improvement on manufacturing applications and WRF, and 25% performance improvement on BWA. With Platinum CPUs, 14G is 40% to 80% better than 13G (Broadwell) for WRF, BWA and manufacturing applications.

The performance improvements with HPL are significantly more due to the introduction of AVX2 and AVX512; 14G is 2x-3x faster than 13G (Broadwell) and up to 7.5x faster than 12G (Ivy Bridge).

14G and Skylake publications to date from the HPC Innovation lab are listed below, additional applications and cluster level studies will be added in the coming months.

1. [BIOS characterization for HPC with Intel Skylake processor](#)
2. [NAMD Performance Analysis on Skylake Architecture](#)
3. [LAMMPS Four Node Comparative Performance Analysis on Skylake Processors](#)
4. [Deep Learning Inference on P40 vs P4 with Skylake](#)
5. [Performance study of four Socket PowerEdge R940 Server with Intel Skylake processors](#)
6. [HPCG Performance study with Intel Skylake processors](#)
7. [Dell EMC HPC Systems - SKY is the limit](#)

