HPI Hardware Update – April 2016

Markus Dreseler,
markus.dreseler@hpi.de

Summary

- 3D XPoint NVM roadmap shows projected system penetration
- HPE sells battery-backed DRAM for "persistent" memory, increasing database performance
- Intel releases midrange Broadwell Server CPUs
- Intel demonstrates CPUs with integrated FPGA
- Oracle allows for open access to Database Accelerator (DAX) coprocessor API
- Intel moves their development cycle from a Tick-Tock-Cycle to Three-Step-Cycle for new processor generations

3D XPoint

Micron released a roadmap looking at the mass production for their 3D XPoint NVRAM for mainstream servers. The most interesting part here is the projected market penetration over the next years. Starting with only 5% in 2018, they expected more than a fourth of the server memory to be powered by 3D XPoint in 2022 [XP1].
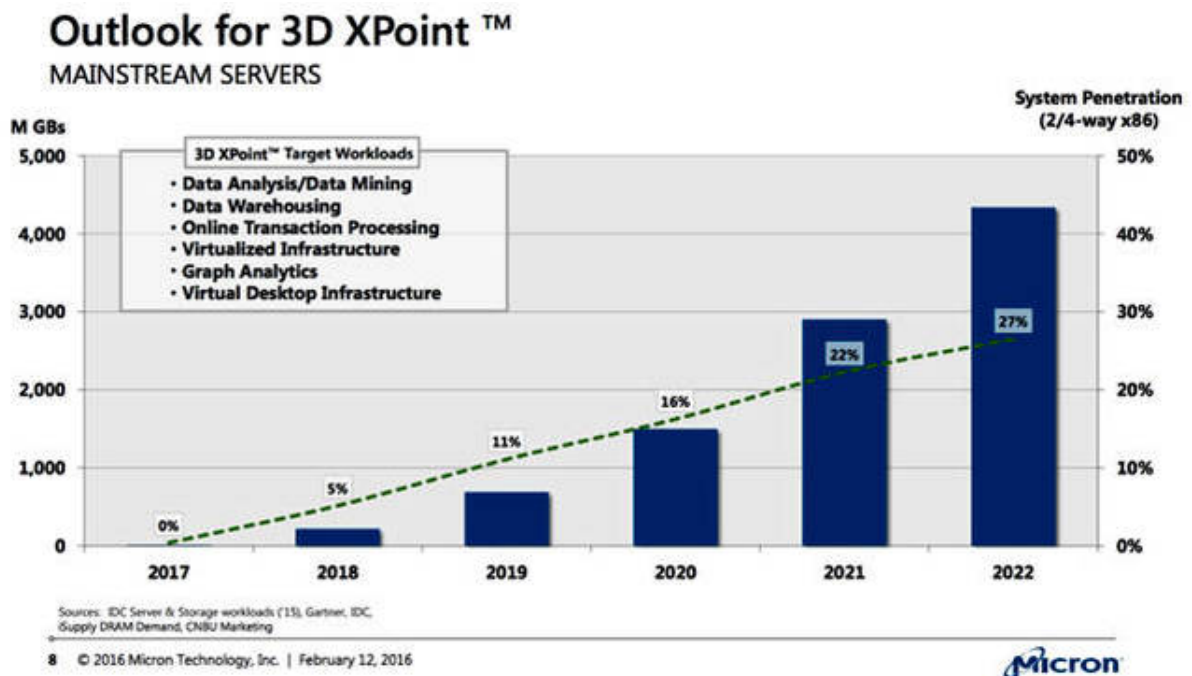

Figure 1: Micron's 3D XPoint roadmap [XP1]

HPE sells powerfail-protected DRAM

HPE adds powerfail-protected DRAM to its Broadwell-based Proliant Gen9 servers. The technology uses an additional battery and has an FPGA for writing the DRAM data back to NAND flash in case of a crash. Currently, 128 GB of memory can be supported in the system. The 8 GB DRAM/NAND combo modules are sup-

plied by Micron and come at a price tag of $899 each (compare to $250-300 for standard DRAM) [HP1]. The DIMMS support the maximum DDR4 speed of 2400 MT/s (19,2 GB/s peak). Larger DIMMs are being developed and 16 GB samples are expected this quarter [HP2,HP3].

With regards to databases, HPE claims up to 2x faster Microsoft SQL Server logging and up to 4x faster cluster replication [HP3]. This appears to be mostly an improvement in terms of latency (from 372 ms to 181 ms), but not so much in terms of throughput (970k to 1080k transactions/s) [HP4].

It is to note that this is not "non-volatile memory" in the strict sense. Compared to actual NVRAM, the common denominator is a protection against power outages. This technology does not bring a higher capacity or better price per GB; however, due to being based on regular DRAM, it has none of the latency and bandwidth disadvantages of NVRAM.

Xeon E5 v4 Broadwell

Intel released their new mid-range server processors Xeon E5 v4, based on the Broadwell architecture. These can be used in either single- or dual-socket servers. These are the first mid-class CPUs on the 14nm architecture [BW1]. Still, they are socket-compatible to previous Haswell processors, which allows reusing existing servers [BW2].

Several improvements have been made to improve the overall performance. Among these are an increased core count (22 from 18) and support for faster DDR4-2400 RAM with a peak transfer rate of 19,2 GB/s. They also include optimizations for certain processor operations that now require a lower number of cycles, improved address predictions, and a larger out-of-order scheduler [BW1,BW3].

Of relevance for optimized database components is a change to the AVX vector operations (used in attribute scans on compressed vectors). With the Broadwell architecture, the execution of AVX operations does not slow down other threads on the same processor any more [BW2,BW4].

These Broadwell processors also introduce an optimistic coherency model optimized for improved local and remote bandwidth at the cost of an increased latency [BW4].

Xeon with built-in FPGAs

Additionally, Intel announced that "later this year", they will ship Broadwell processors that have Field-Programmable Gate Arrays (FPGAs) on the same package. These are hardware components that can be programmed to implement accelerators for certain algorithms, such as compression, encryption, or analytics [FP1].

This would allow developers to offload these tasks to more specialized and optimized hardware, thus increasing the throughput and reducing the CPU load.

It is to note that the current layout has the Xeon processor and the FPGA sitting on the same package (but not on the same die). They are connected using QPI,

the same interconnect that also connects multiple processors. As such, they have regular access to the system's memory and are part of the coherence domain [FP2,FP3]. Additionally, they can have their own memory attached (see Figure 2).
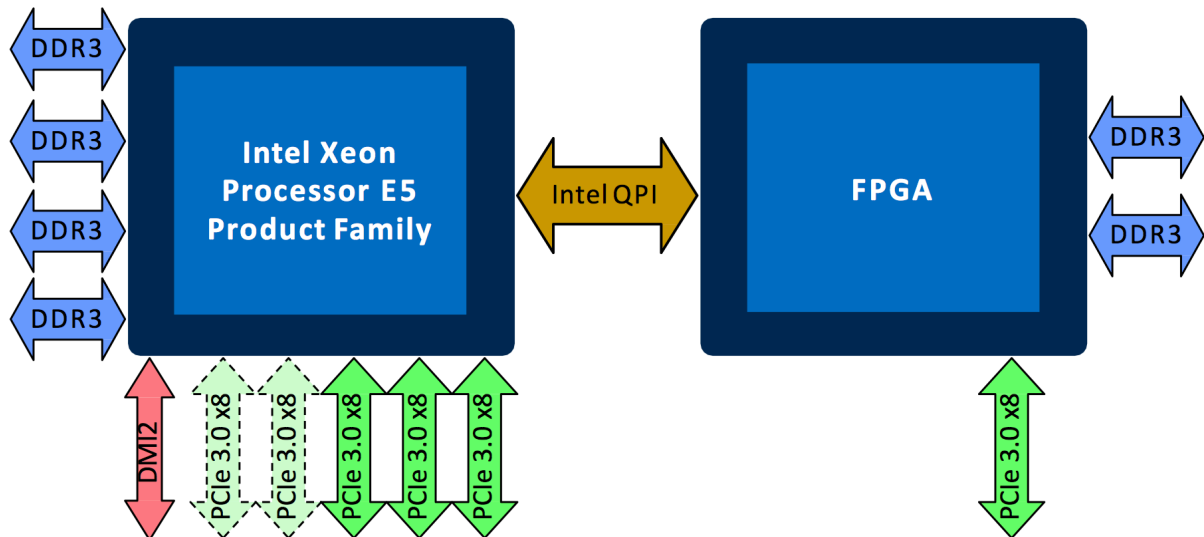

Figure 2: Architecture of a combined Xeon + FPGA processor [FP3]

Offloading more and more work to specialized accelerators appears to be a direction in which more and more technologies are headed. As examples, Oracle uses their Database Accelerators (DAX) to improve DB performance, Mellanox is working on network accelerators that moves more message passing operations to the hardware, and SGI wants to offload similar message handling to their specific hardware.

Newsflash

- Oracle is about to open up the API access to their database accelerators [NF1]. These are currently used in their 12c database to decompress, filter, and analyze data off-CPU. Previous numbers claim a 8.2x performance advantage over Haswell processors that do not have such an accelerator [NF2].

- Since 2007, the development of Intel CPUs have followed a "tick-tock" cycle, in which a "tick" represented a shrinked processor architecture and a "tock" an optimized microarchitecture. After having had to delay the 10nm Cannonlake architecture and introducing Kaby Lake as a third processor manufactured in a 14nm process, rumors were that Intel was about to abandon this cycle. It is now confirmed that they will follow a "process (tick) – architecture (tock) – optimization" cycle in which only every third step comes with a smaller manufacturing process. This accounts for the fact that it becomes more and more difficult to reduce processor sizes [NF3].

References

[XP1] http://files.shareholder.com/downloads/ABEA-45YXOQ/1576449138x0x875021/4BEAA02E-BBC2-402C-A51D-B3B2C6B8C3D4/Winter_Analyst_Day_2016.pdf

[HP1] http://www.eetimes.com/document.asp?doc_id=1329309
[HP2] http://www.theregister.co.uk/2016/03/31/hpe_adds_powerfailprotected_nvdimm/
[HP3] http://www.theregister.co.uk/2016/03/08/persistent_memory_ides_of_march_from_hpe/
[HP4] http://community.hpe.com/t5/Servers-The-Right-Compute/HPE-unleashes-Persistent-Memory-the-performance-of-memory-with/ba-p/6846074

[BW1] http://www.nextplatform.com/2016/03/31/broadwell-brings-xeon-e5s-balanced-performance-bump/
[BW2] http://insidehpc.com/2016/04/why-new-intel-xeon-processors-make-sense-for-hpc-applications/
[BW3] http://www.enterprisetech.com/2016/03/31/new-intel-xeon-processors-benefit-hpc-applications
[BW4] http://www.anandtech.com/show/10158/the-intel-xeon-e5-v4-review

[FP1] http://www.theregister.co.uk/2016/03/14/intel_xeon_fpga/
[FP2] http://www.nextplatform.com/2016/03/14/intel-marrying-fpga-beefy-broadwell-open-compute-future/
[FP3] http://www.slideshare.net/insideHPC/using-xeon-fpga-for-accelerating-hpc-workloads

[NF1] https://www.oracle.com/corporate/pressrelease/free-open-apis-and-developer-kit-for-dax-03082016.html
[NF2] https://blogs.oracle.com/BestPerf/entry/20151025_rte_t7_1
[NF3] http://hexus.net/tech/news/cpu/91499-intel-officially-ends-tick-tock-era/