VSC-4 — The Next Step

Herbert Störi

VSC Research Center, TU Wien

VSC-4 (Vienna Scientific Cluster 4) has finally been ordered. After a procurement process that took considerably more time than hoped and expected, the Vienna Company EDV-Design won the contract in a competitive bidding process. The system ordered is a Lenovo system very similar to the SuperMUC NG in the LRZ (Leibniz Supercomputing Centre) in Garching/Germany.

Our new system will consist of 790 water cooled nodes (Lenovo SD650), each with two Intel Skylake Platinum 8174 processors. Each processor has 24 cores with a base frequency of 3.1 GHz. A turbo mode up to 3.9 GHz is available if not all cores are active. Consequently, a node has 48 physical cores. The 700 standard nodes have a main memory of 96 GByte. There will be 78 fat nodes with 384 GByte of main memory and 12 very fat nodes with 768 GByte. Each node is also equipped with an SSD device of 480 GByte, available as temporary storage during the runtime of a job.

The system will be complemented with 10 login nodes and parallel file systems. Compute nodes, login nodes and file system nodes will be interconnected with 100 Gbit/s Omni-Path. The Omni-Path network is a two-level fat tree with an oversubscription of 2:1. There are no islands. An edge switch connects to 32 compute nodes, giving non-blocking access to 1536 cores. At the time of writing this abstract, the details of the file systems are still under discussion but will be disclosed during the talk.

While we do not expect major problems with moving applications to the new system, using it in an efficient and performant way might prove challenging. The large number of 48 cores in one compute node suggests to use shared memory or hybrid parallelization schemes. A node reaches 250 points in the SPECrate 2017 Floating Point benchmark, which is about 5 times the performance of a VSC-3 node (16 cores). Looking at the Linpack benchmark, a single VSC-4 node will reach about 3 TFlop/s, about 10 times the performance of a VSC-3 node at slightly less than 300 GFlop/s. A large part of this floating-point performance will come from the two AVX-512 units per core. Each AVX-512 unit performs 8 fused-multipy-add instructions per cycle, permitting —in an optimized situation— 32 floating-point instructions per cycle, albeit at a slightly reduced processor frequency.

The compute nodes are directly water cooled allowing to use primary cooling water with a temperature in excess of 43 °C, permitting year-round free cooling. Approximately 90 % of the energy will be removed by this high-temperature loop, while 10 % need to be removed by air cooling. This permits a very reasonable energy foot-print.

We expect delivery of the system in March 2019 and full user operation in June 2019. At this point, we abstain from any speculation about the TOP500 placement of the new system.

service@vsc.ac.at
vsc.ac.at