



## Cornelis Networks Omni-Path Delivers 53% Better Price-Performance than NVIDIA InfiniBand HDR when Running SeisSol on AMD EPYC Processors

### White Paper

---

This article serves to complement a recent Cornelis Networks™ study<sup>1</sup> in the scalability of SeisSol<sup>2</sup> to include a price-performance comparison between Cornelis Networks<sup>3</sup> Omni-Path™ (100Gbps) and NVIDIA®<sup>4</sup> InfiniBand HDR (200Gbps) on AMD EPYC™ 7713 64 core processors. Readers are encouraged to review the referenced paper as it contains simulation details, relevant performance metrics, helpful build tips, and a Slurm submission script. This paper shows that with Cornelis Omni-Path, users can achieve up to 1.53X higher price-performance when running SeisSol when compared to NVIDIA InfiniBand HDR.

### Simulation Details

SeisSol is a high-performance, open-source seismic wave simulation software package used in geophysics academia and may be applied in the oil and gas industries. TPV33 simulations<sup>5</sup> are run with SeisSol v1.0.1 on 64 core dual socket AMD EPYC 7713 nodes. Each node contains 16 NUMA nodes, with 8 cores per NUMA. NVIDIA InfiniBand HDR simulations use UCX 1.15.0, while Cornelis Omni-Path uses the highly optimized OPX 1.18.1 libfabric provider. SeisSol is compiled using the Intel Classic Compiler-2021.9.0,<sup>6</sup> and openmpi-4.1.4<sup>7</sup> as the MPI wrapper. The simulation is partitioned into 8 cores per rank and 16 ranks per node. The domain is partitioned into .58M, 2.2M, and 7.7M elements, each of which are simulated five times using 1, 2, 4, 8, and 16 nodes.

### Scalability and Price-Performance

CPU utilization (Hardware GFLOPS (HWFLOPS)) and job throughput (cases/day) are tabulated for each simulation and results are averaged across node count, mesh granularity, and fabric provider. By convention, the fabric is differentiated by color, and mesh granularity is differentiated by transparency in the following figures.

Figure 1 shows that simulations using Cornelis Omni-Path have up to a 67% lower run-to-run variability (at 16 nodes) than NVIDIA InfiniBand HDR. Run-to-run variation is defined as the relative standard deviation (RSD) of the five runs. On average across all node counts and mesh sizes, NVIDIA InfiniBand HDR has a higher run to run variability by 1.72X. Moreover, increasing the mesh resolution lowers the %RSD, similar to previous observations.

---

<sup>1</sup> <https://www.cornelisnetworks.com/seissol-performance-and-scalability-with-cornelis-networks-omni-path-on-3rd-generation-intel-xeon-scalable-processors/>

<sup>2</sup> <https://seissol.org/>

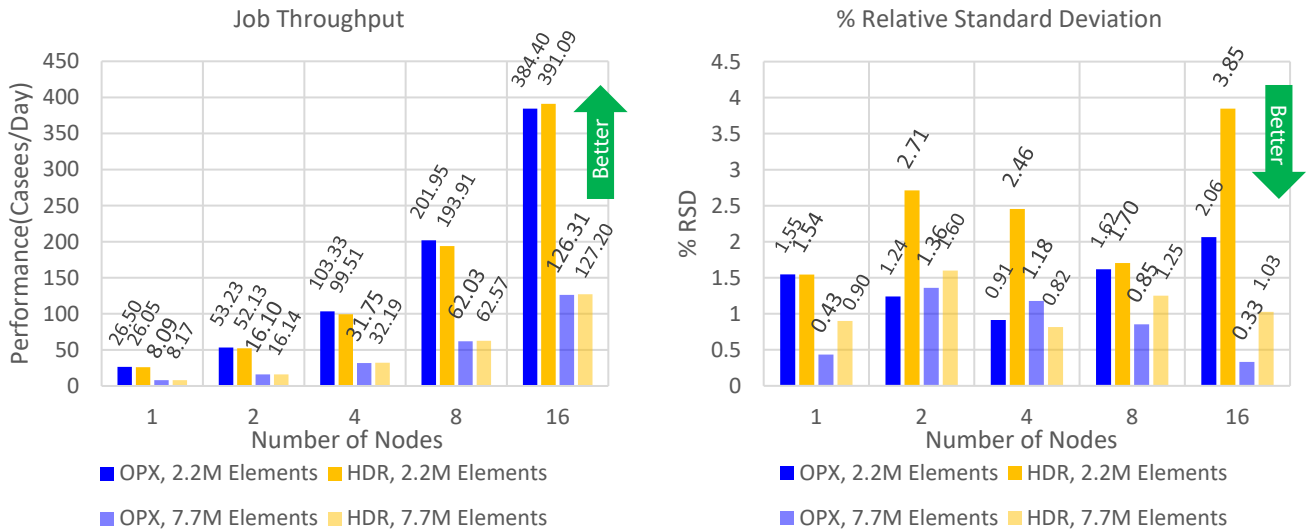
<sup>3</sup> <https://www.cornelisnetworks.com>

<sup>4</sup> <https://network.nvidia.com/files/doc-2020/wp-introducing-200g-hdr-infiniband-solutions.pdf>

<sup>5</sup> <https://strike.scec.org/cvws/tpv33docs.html>

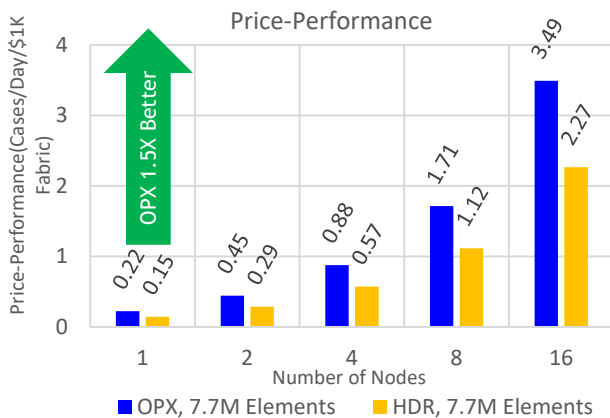
<sup>6</sup> <https://www.intel.com/content/www/us/en/docs/cpp-compiler/developer-guide-reference/2021-10/overview.html>

<sup>7</sup> <https://www.open-mpi.org/software/ompi/v4.1/>



**Figure 1.** Job throughput(left) and relative standard deviation of job throughput (right) for TPV33 simulations using 2.2 (solid) and 7.7 (transparent) million elements for 1, 2, 4, 8, and 16 nodes on NVIDIA InfiniBand HDR (yellow) and Cornelis Omni-Path 100 Gbps (blue).

For job throughput shown in Figure 1, the number of cases per day between Cornelis Omni-Path and NVIDIA InfiniBand HDR are within a standard deviation of each other across all nodes and mesh resolutions. Any performance differences between fabrics are therefore statistically negligible. Simulations perform with upwards of 90% scaling efficiency which implies efficient usage of computational resources.



**Figure 2.** Price-performance for TPV33 simulations using 7.7M elements for 1, 2, 4, 8, and 16 node simulations on Cornelis Omni-Path 100Gbps (blue) and NVIDIA InfiniBand HDR (yellow).

An important consideration when running HPC applications is the cost of equipment. For a price-performance comparison between NVIDIA InfiniBand HDR and Cornelis Omni-Path, fabric costs to build a 16-node cluster using both technologies are obtained from public sources.<sup>8</sup> Job throughput for 7.7M simulations<sup>9</sup> is normalized by these costs and price-performance is shown in Figure 2. Users can expect to achieve the same job throughput on AMD EPYC 7713 nodes using a 16-node Cornelis Omni-Path cluster for 64% of the fabric cost compared to an equivalent

<sup>8</sup> MSRP Pricing obtained on 7/11/2023 from <https://store.nvidia.com/en-us/networking/store>. Mellanox MCX653105A-HDAT \$1628 per adapter. Mellanox MQM8700-HS2F managed HDR switch, \$25555. MCP1650-H002E26 2M copper cable – \$281. Cornelis Omni-Path Express MSRP pricing as of 7/11/2023. Cornelis 100HFA016LSN 100Gb HFI \$880 per adapter. Cornelis Omni-Path Edge Switch 100 Series 48 port Managed switch 100SWE48QF2 – \$19750. Cornelis Networks Omni-Path QSFP 2M copper cable 100CQQF3020 – \$147. Exact pricing may vary depending on vendor and relative performance per cost is subject to change.

<sup>9</sup> Since performance is similar between fabrics for a given resolution, the price-performance does not change, and other resolutions are omitted for brevity.



cluster using NVIDIA InfiniBand HDR, yielding 1.53X higher performance per fabric investment.

## Conclusions

This paper expands a recent study to AMD EPYC 7713 processors for the SeisSol seismic wave HPC application. It includes comparisons of performance and price-performance between Cornelis Omni-Path and NVIDIA InfiniBand. Up to a scale of 16 nodes, it has been shown that while the absolute performance of the two fabric technologies is similar, users will experience more consistent, predictable performance while running over Cornelis Omni-Path. Also, the price-performance of Cornelis Omni-Path far exceeds NVIDIA InfiniBand, meaning that budget can be saved or spent in other areas to improve computational capacity including more compute nodes, memory, storage, or software licenses.

Cornelis Omni-Path 100 is ready to ship with minimal lead time; contact your vendor of choice today to start experiencing HPC with Cornelis Omni-Path 100!

## System Configuration

Simulations performed on dual socket AMD EPYC 7713 64-Core Processors. Rocky Linux 8.4(Green Obsidian). 4.18.0-305.19.1el8\_4.x86\_64 kernel. 32x16GB, 256 GB total, 3200 MT/s. BIOS: Logical processor: Disabled. Virtualization Technology: disabled. NUMA nodes per socket 4. CCXAsNumaDomain: Enabled. ProcTurboMode: Enabled. ProcPwrPerf: Max Perf. ProcCStates: Disabled

NVIDIA InfiniBand HDR: 200Gbps MCX653105A-HDAT adapters connected with a MQM8700-HS2F switch. MLNX\_OFED\_LINUX-5.4-2.4.1.3, OpenMPI 4.1.5rc2 and UCX as provided by hpcx-v2.16-gcc-mlnx\_ofed-redhat8-cuda11-gdrCOPY2-nccl2.18-x86\_64

Cornelis Omni-Path: Host Fabric Interface Adapter 100 Series 100HFA016LSN, Edge Switch 100-Series 48-port 100SWE48QF2. OPXS 10.11.1.1.1, OpenMPI 4.1.4.

SeisSol v1.0.1 compiled using Intel Classical Compiler – 2021.9.0, openmpi-4.1.4. Example mpirun command, where omp are the number of cores per rank, and hostfile.txt contains the simulation nodes:

```
mpirun -np (process per node)*(number of nodes) --map-by ppr: node:pe=omp -x  
OMP_NUM_THREADS=(omp-1) -mca btl self, vader <fabric flags> --hostfile hostfile.txt  
SeisSol_Release_drome_4_elastic parameters.par
```

Fabric-specific flags:

Cornelis Omni-Path: -x FI\_PROVIDER=opx -mca mtl ofi -x FI\_OPX\_HFI\_SELECT=0

NVIDIA InfiniBand HDR: -x UCX\_NET\_DEVICES=mlx5\_0:1 -mca pml ucx -mca coll\_hcoll\_enable 0  
(performance with HCOLL enabled was lower)

## Legal Disclaimer

You may not use or facilitate the use of this document in connection with any infringement or other legal analysis concerning Cornelis Networks products described herein. You agree to grant Cornelis



Networks a non-exclusive, royalty-free license to any patent claim thereafter drafted which includes subject matter disclosed herein.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

All product plans and roadmaps are subject to change without notice.

The products described may contain design defects or errors known as errata which may cause the product to deviate from published specifications. Current characterized errata are available on request.

Cornelis Networks technologies may require enabled hardware, software, or service activation.

No product or component can be absolutely secure.

Your costs and results may vary.

Cornelis, Cornelis Networks, Omni-Path, Omni-Path Express, and the Cornelis Networks logo belong to Cornelis Networks, Inc. Other names and brands may be claimed as the property of others.

Copyright © 2024, Cornelis Networks, Inc. All rights reserved.