OpenFOAM® Performance and Scalability with Cornelis Networks™ OPX on 3rd Generation Intel® Xeon® Processors

Article

Cornelis Networks is the leading independent provider of purpose-built, open-source, scale-out interconnects for high-performance computing, artificial intelligence, and high-performance data analytics. The Cornelis Networks Omni-Path Express™ (OPX) high-performance fabric delivers class-leading throughput, latency, and scalability allowing customers to deploy solutions which enable faster time to solution and improved workload scalability combined with leading price/performance.

To highlight the performance and price-performance capabilities of the Omni-Path fabric, this paper compares the performance of an industry standard OpenFOAM® benchmark case on clusters interconnected by both Cornelis OPX and NVIDIA® InfiniBand HDR fabrics.

OpenFOAM is a suite of solvers, pre- and post-processing utilities developed to solve modelling and simulation problems in the domain of CFD. In this paper, the industry standard motorbike tutorial case visualized in Figure 1[1] is meshed at a resolution of 20M and 42M cells and is used to demonstrate how the fabric affects application run time and scalability.

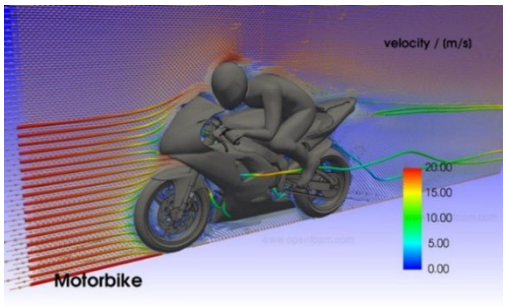

Figure 1. Visualization of OpenFOAM Motorbike.[1]

These simulations model a low speed (incompressible) air flow around a motorcycle and rider. The simulation solves for a steady-state solution which is commonly used in many scientific fields where metrics such as skin friction and drag are a primary design factor. While smaller workloads can be performed in a single compute node where CPU performance and memory bandwidth are key, larger models spanning multiple nodes require a high-performance and cost-effective fabric.

Cornelis OPX is designed specifically for high-performance, parallel computing environments. It is built utilizing a unique link-layer architecture and a highly optimized OFI libfabric provider[2] delivering higher message rates and lower latencies than competing interconnects and with leadership price/performance.

Scalability

Figure 2 compares the performance scalability of the benchmark using up to 16 Intel Xeon Scalable Platinum 8358 dual-socket nodes, for a total of 1024 cores, connected with Cornelis OPX fabric using a single rail operating at 100Gbps and the same nodes connected with an NVIDIA InfiniBand HDR fabric operating at 200Gbps. Intel MPI 2021.8 is used for both fabrics, with UCX version 1.15.0 from

[1] OpenFOAM runTimePostProcessing visualization: https://www.openfoam.com/news/main-news/openfoam-v3.0/post-processing
[2] https://ofiwg.github.io/libfabric/

the latest HPC-X v2.15 used for the NVIDIA InfiniBand HDR measurements. Cornelis Omni-Path measurements are performed with the OPX provider from libfabric 1.18.0.

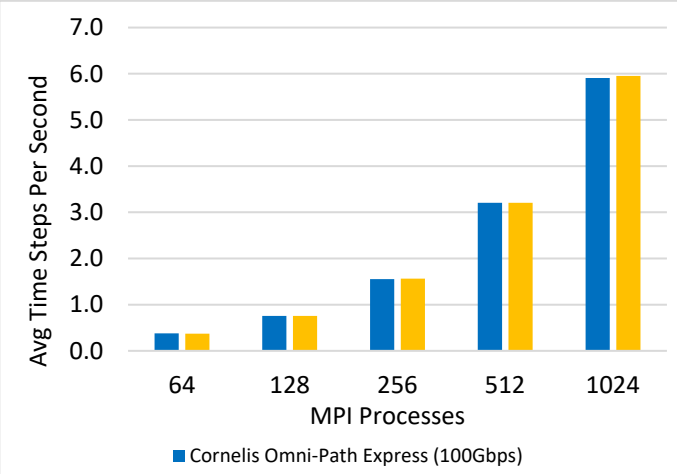

Figure 2. Scalability of the OpenFOAM Motorbike 20M cell model.

The results show that a single 100Gbps rail of Cornelis OPX delivers a comparable level of performance to NVIDIA InfiniBand HDR running at 200Gbps in a 16-node cluster. Each data point was run five times, eliminating the minimum and maximum performance, and averaging the middle three.

Price-Performance

In addition to performance, another important consideration in fabric selection is price. For this second comparison, MSRP pricing[3] was used to build a 16-node cluster consisting of a single edge switch, 16 cables, and 16 host adapters.

Performance is shown in terms of job throughput per day on a fully utilized 16-node cluster normalized by the cost of the fabric.

In Figure 3, the results show that a Cornelis OPX connected cluster delivers an average of 1.55x better job throughput per fabric cost running the Motorbike 20M and 42M test cases compared to the NVIDIA InfiniBand HDR cluster. This means users can obtain peak OpenFOAM performance with a lower budget, or they can deploy more nodes with the same budget to increase computational capacity and/or shorten the time to results.

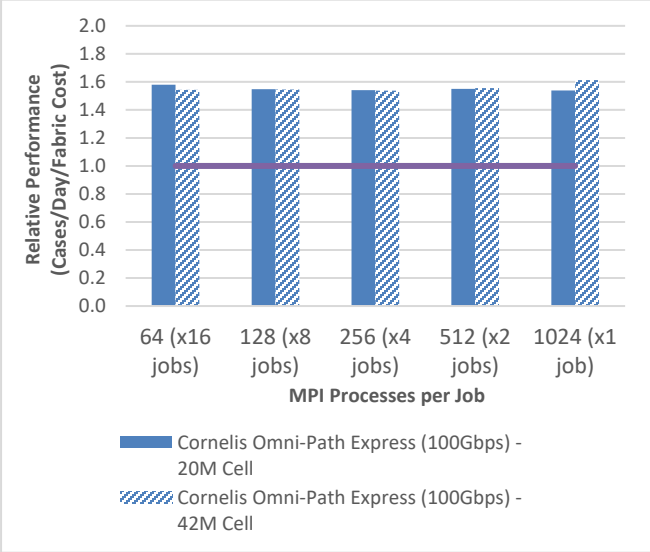

Figure 3. Job Throughput (Cases/Day) normalized by Fabric Cost.

[3] 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 cable100CQQF3020 - $147. Exact pricing may vary depending on vendor and relative performance per cost is subject to change.

In conclusion, the OpenFOAM software combined with Cornelis Networks OPX fabric delivers leadership performance and up to 1.55x better return on investment. Cornelis Networks Omni-Path (100-series) hardware is available now, contact sales@cornelisnetworks.com to get started!

System configuration

Tests performed on 2 socket Intel® Xeon® Scalable Platinum 8358 Processor-based servers. Rocky Linux 8.4 (Green Obsidian). 4.18.0-305.19.1.el8_4.x86_64 kernel. 32x16GB, 256 GB total, 3200 MT/s. BIOS: Hyper-Threading: Disabled. Virtualization Technology: Disabled. Power and Performance Policy: Performance. C-State: C0/C1. C6: Disabled. P-States: Disabled. Turbo Boost: Enabled.
OpenFOAM v22.06 SimpleFoam compiled with gcc 10.2. Example run command: `mpirun -np ${NP} -ppn ${PPN} -f hostfile simpleFoam -parallel  blockMeshDict 100x40x40 (20M) and 130x52x52 (42M), decomposeParDict - scotch decomposition.`

Cornelis Omni-Path Additional run flags: `-genv FI_PROVIDER=opx -genv LD_LIBRARY_PATH=${LIBFABRIC_LIB_PATH}:$LD_LIBRARY_PATH FI_OPX_HFI_SELECT=0`

NVIDIA HDR Additional run flags: `-genv UCX_NET_DEVICES=mlx5_0:1 -genv FI_PROVIDER=mlx -genv I_MPI_COLL_EXTERNAL=0`

Legal Disclaimer