

## Deliver High-Performance Networking on Ethernet

### Intel high-performance networking supports demanding workloads on clusters using a high-bandwidth, scalable Ethernet fabric.

The explosion of ChatGPT has escalated interest in artificial intelligence (AI), and companies everywhere are investigating how AI can be used throughout their business. To keep up with this growth, enterprise IT teams must build new high-performance clusters to deliver the computing power they need. However, AI and other scalable workloads, such as high-performance computing (HPC), require high-performance networking. IT teams face fresh challenges trying to deliver the AI computing power the business requires in possibly an unfamiliar clustered environment.

Companies will make numerous budget decisions when building a new cluster. The budget must cover the new server nodes and the supporting infrastructure, such as rack server solutions, software licensing, and fabric costs. Proprietary fabrics supporting AI and HPC application workloads, such as InfiniBand, can consume much of the cluster budget. This fabric cost might reduce the number of nodes that can be incorporated. In addition, these proprietary fabrics require specialized knowledge to implement, integrate with other data center services, and maintain. This can drive up costs because these complicated fabrics require additional IT training and are time-consuming to learn and difficult to debug.

One solution might be to use what's readily available across networks that can scale for AI and HPC. Ethernet is a well-known, common interconnect. Quoting the Ethernet Alliance, "Ethernet is Everywhere." It's already used in AI and HPC clusters, where cost is a key criterion when deciding what fabric to use. But while Ethernet has the advantage of ubiquity, standard Ethernet alone cannot scale to meet AI/HPC requirements.

### Enabling Ethernet in a clustered and distributed computing environment

Intel's High-Performance Networking (HPN) with Ethernet solution tackles these challenges directly by delivering Ethernet anywhere a clustered, distributed compute environment is required: AI, HPC, HPC in the cloud, and across some enterprise applications. Because HPN uses standard Ethernet components, it is less expensive than proprietary solutions and requires no additional training. The solution also helps achieve application performance comparable to proprietary fabrics when building for the most common cluster sizes.

HPN includes a software stack similar to other high-performance fabrics. It extends the capabilities of Remote Direct Memory Access (RDMA) over Converged Ethernet (RoCE) v2 to enable Ethernet to become a fabric for HPC/AI.

## Intel® Ethernet Fabric Suite

The next generation of mature fabric software ported to Ethernet, Intel® Ethernet Fabric Suite, pulls together the HPN solution. Evolved from Intel® Omni-Path Architecture, this software is tuned to optimal performance with OpenFabrics Interfaces (OFI) to scale and provide the required function.

The Intel Ethernet Fabric Suite is fully integrated into Intel® oneAPI Collective Communications Library (oneCCL), an open, cross-industry, standards-based, unified, multiarchitecture, multi-vendor programming model that delivers a common developer experience across accelerator architectures.

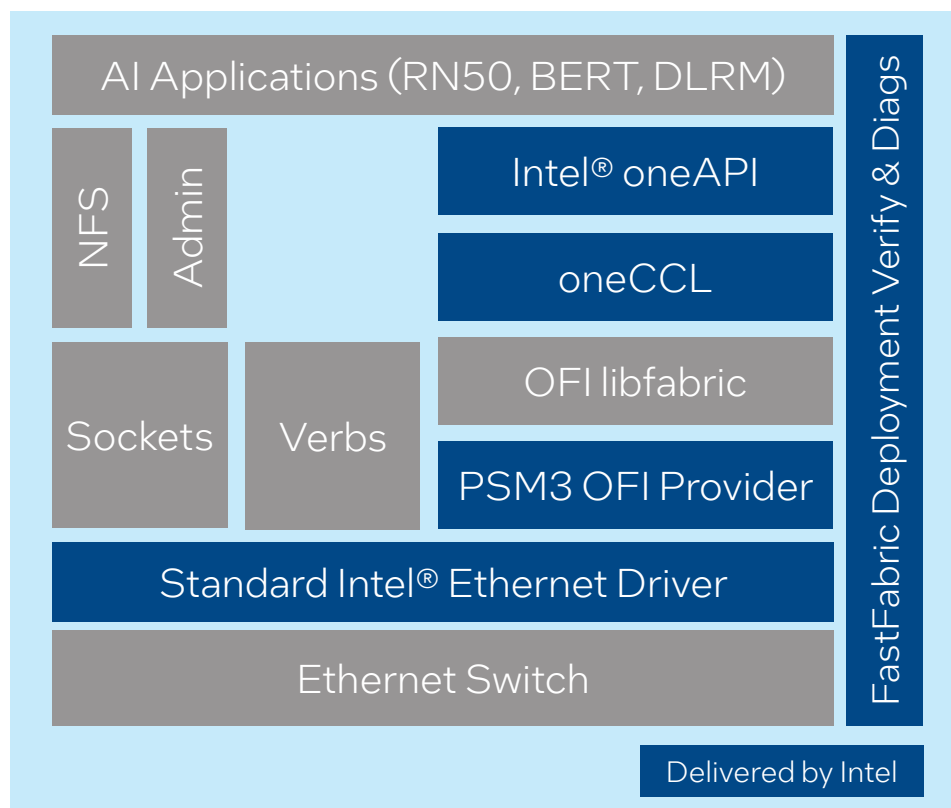


Figure 1. Intel Ethernet Fabric Suite is open-source software implementing the OpenFabrics Interfaces (OFI)

## Intel® Ethernet 800 Series

The Intel® Ethernet 800 Series is a family of controllers and network adapters capable of providing high bandwidth with RoCE v2 and support for PCIe 4.0. The latest-generation network adapter can deliver bandwidth up to 200 gigabit Ethernet (GbE) in a single PCIe 4.0 x16 slot to meet bandwidth-intensive application requirements.

## Solution benefits

The Intel HPN with Ethernet solution is the result of Intel's leadership in providing continuous innovation for more than 40 years in Ethernet solutions. Combining Intel Ethernet with Intel® Xeon® Scalable processors enables cost-effective AI for many inference usages, as well as small training models where GPUs are not needed to deliver effective time-to-train metrics. These efforts have resulted in a solution with numerous benefits.

## Application performance

Application performance is the most important HPC/ AI performance metric. A recent performance test demonstrated that Intel HPN with Ethernet results were comparable to those of a proprietary fabric on Pytorch, resnet 50 (see Figure 2).<sup>1</sup>

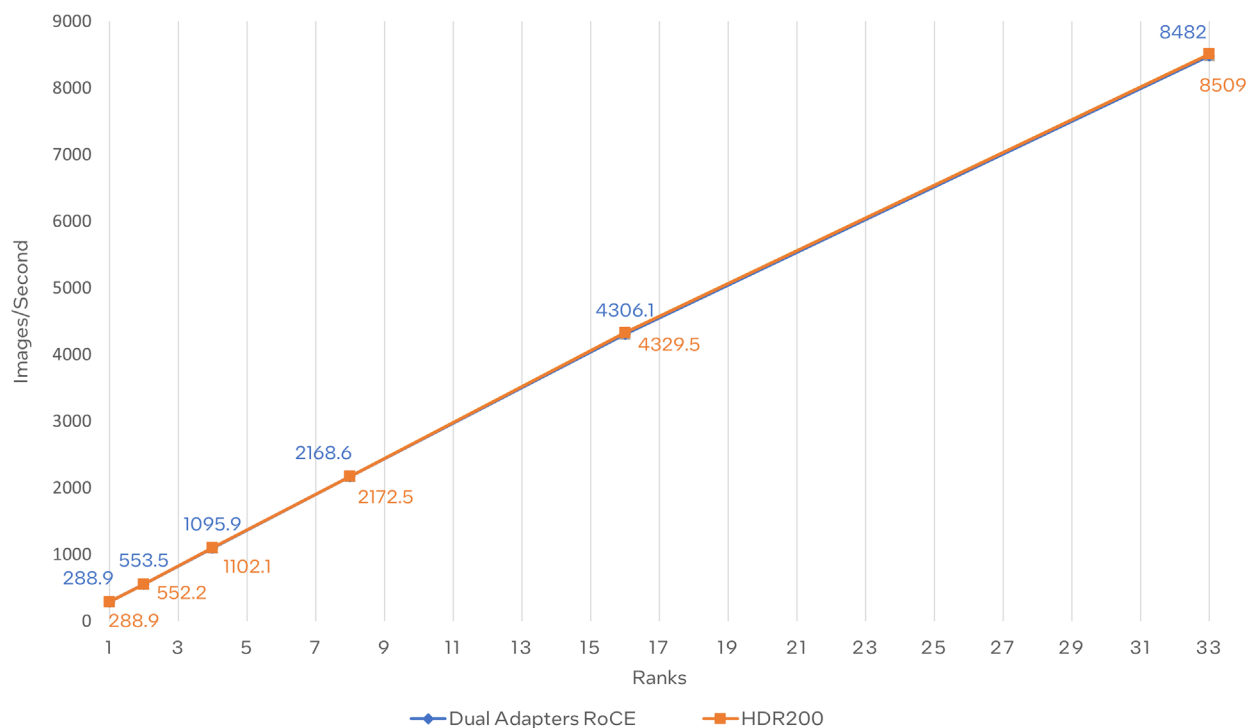


Figure 2. Comparing Intel Ethernet with InfiniBand HDR performance PyTorch, resnet 50-V1.5 (synthetic image data)

## Scalability

The verbs interface, used traditionally with Ethernet, does not normally scale well for HPC and AI applications. At the heart of the Intel HPN with Ethernet solution is the Performance Scaled Messaging (PSM) layer. PSM addresses scalability challenges such as communications, stack memory footprint, latency at scale, latency jitter, and resiliency at scale. The solution, which incorporates the third generation of PSM (PSM3), enables Ethernet to scale with increasing node counts.

## Lower total cost of ownership (TCO)

With comparable application performance, the Intel HPN with Ethernet solution provides the required performance for less upfront costs than proprietary fabrics. This savings allows the purchase of additional servers for even higher cluster performance and lower overall costs than proprietary fabrics. Depending on the version of InfiniBand, Ethernet can be up to 2x less expensive per port. This savings is partly due to the number of ports per switch. Ethernet switches with 128 ports are available, enabling less complex cluster configurations. As a result, the number of required switches is smaller when the cluster size increases. Fewer switches mean a smaller data center footprint, fewer cables, and lower power consumption.

## No code changes

Because Intel HPN with Ethernet is compatible with existing message-passing interface (MPI) applications and supports OpenFabrics Alliance (OFA) interfaces, applications work immediately. The solution also offers strong verbs applications performance for file systems, reliable socket performance, and compatibility with existing applications using OFA verbs or sockets.

## Support for intelligent networking

Intel HPN with Ethernet supports HPC/AI clusters built on high-performing Intel® Xeon® Scalable processors. It helps ease deployment and deliver optimal performance on Intel-based clusters by supporting numerous tools and specifications, such as:

- Intel MPI Library: Delivers flexible, efficient, and scalable cluster messaging on the leading MPI implementation for Intel® architecture-based systems.
- FastFabric: Enables quick fabric deployment, verifying fabric configuration and setup. This tool can also be used to diagnose fabric problems.

## Learn more

Intel high-performance networking (HPN) with Ethernet delivers a reliable out-of-the-box experience and proven interoperability for current and future networking infrastructure.

Contact your Intel representative or visit [intel.com/ethernet](https://intel.com/ethernet).

1. Performance results are based on testing by Intel as of February 2021. See configuration disclosure for details. Performance varies by use, configuration, and other factors. Learn more at [www.intel.com/PerformanceIndex](https://www.intel.com/PerformanceIndex). Configuration: Tests performed on 2-socket Intel Xeon Platinum 8170 processor at 2.10 GHz. Intel Hyper-Threading Technology (Intel HT Technology) enabled. Intel Turbo Boost Technology enabled with Intel P-State driver. Red Hat Enterprise Linux 8.1 (Ootpa) 4.18.0-147.el8.x86\_64 kernel. 12 x DDR4, 196,608 MB, 2,666 megatransfers per second (MT/s). irdma version 1.3.19. ice version 1.3.2. CVL firmware-version: 2.15 0x800049c3 1.2789.0. Ethernet switch: Arista DCS-7170-32CD-F, 4.22.IFX-CLI. PFC enabled on priority 0. CX-5 InfiniBand Mellanox SB7800 Switch-IB2. CX-5: MLNX\_OFED\_LINUX-5.1-2.3.7.1. Intel MPI 2019.10, FI\_PROVIDER=psm3 and mlx (UCX) for Mellanox. HPN: Intel Ethernet Fabric Suite 11.0.0.162. EDR InfiniBand: Mellanox OFED with UCX. Contact HPN.org for application testing details.

Performance varies by use, configuration and other factors. Learn more at [www.intel.com/PerformanceIndex](https://www.intel.com/PerformanceIndex).

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software or service activation.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

© Intel Corporation. Intel, the Intel logo and Xeon are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others.