Computing at the speed of memory

Accelerating the path to exascale with silicon photonics

Contents

Technical white paper

In an age where information is a key driver of research and development (R&D) and explosive data growth is challenging even the most proficient computers, technological advancement is crucial to success. This major paradigm shift is leading high performance computing (HPC) developers to pursue superior system performance in order to drive economic growth, enable scientific discovery, and solve the world's most daunting issues. However, rising data volumes are overwhelming many of today's HPC applications and services, and the need to quickly extract and enact on critical insights is accelerating the demand for a new generation of supercomputers.

The U.S. Department of Energy (DOE) launched Project PathForward, a principal component of the Exascale Computing Project (ECP) to improve developer productivity and galvanize R&D efforts in the race to exascale. The creation of a sustainable exascale ecosystem will usher in a new realm of innovation which promises to enhance economic leadership, bolster national security, and ensure competitive advantage on a global scale. Today, the fastest supercomputers conduct operations at petascale, or a quadrillion (10^{15}) calculations per second. The goal of the ECP is to build a system capable of executing a quintillion (10^{18}) calculations per second at a benchmark of 25 megawatts. This level of performance will catalyze unprecedented breakthroughs in precision medicine, weather and climate forecasting, additive manufacturing, renewable energy, scientific exploration, and much more.

However, developers are facing significant design hurdles in the quest for exascale. Energy consumption and bandwidth allocation are two central concerns regarding system performance and usability. Transmitting vast quantities of information between disparate parts of the machine or data center requires an extraordinary amount of power, and as data workloads continue to grow, the energy, bandwidth, and memory footprints expand exponentially.

To address these issues, developers are leveraging photonics to streamline communication, accelerate data transfer, and significantly reduce energy consumption. Using fiber optics, photons are transmitted and received by transceivers made up of an electronic chip, a laser, and a lens that focuses light from the laser into the optical fiber, effectively increasing data throughput by a factor of 10. Photonic transmission combined with non-volatile memory (NVM) allows the system to retain information without drawing power.

Computing at the speed of an exaflop/s requires a balance of communication to computation. Data must be communicated at a minimum of 0.01 bytes/s per flop/sfflop, and the system must target 1 to 5 picojoules of energy, or one million millionths (10-12), per bit to efficiently transfer data end to end without spending the entire power budget on the interconnect. This target is fundamentally impossible to reach with electronic interconnects, which utilize nanojoules of energy, 1,000X bigger and rapidly decrease in performance as distance increases. By converting operations from electrical to optical, the system becomes distance-independent, faster, and more resilient.

Streamlining system communication

In today's HPC landscape, memory capacity and bandwidth are struggling to keep pace with the increase in flops, leaving developers scrambling to research, design, test, and deploy robust new solutions and applications. Without question, a chief concern of exascale is implementing the right technologies to modernize, strengthen, and sustain the next interconnect frontier for HPC. Two families of technology are excelling in this technological race.

The first family consists of Vertical Cavity Surface-Emitting Lasers (VCSELs), or miniaturized lasers, which have existed in development and production for over a decade. Like microscopic flashlights, VCSELs produce information by modulating, or turning on and off, different light wavelength (that is, different colors) that correspond to different channels. For the next few years of "early exascale" systems, developers will hone VCSELs technology to reach the appropriate speed, cost, and level of integration, and identify vendors in the supply chain to help create a sustainable production process. The key to empowering VCSELs is finding effective ways to streamline production and deployment. HPC developers are borrowing technologies from other industries, miniaturizing practical components to integrate with electronics (e.g. drivers that convert electrical to optical signals), and packaging them in a way that has a minute physical footprint and can be closely connected with computing elements.

A combination of mechanical, optical, and electrical technologies will allow users to leverage coarse-grained wavelength-division multiplexing (CWDM) to pack multiple optical signals into a single fiber. The signals generate colors that bounce around a system of mirrors and combine in the same fiber, all within the space of a few cubic millimeters. This is small enough for the VCSELs components to surround the compute and switch component sockets and supply optimal bandwidth at the lowest possible cost.

The second family employs silicon photonics, a pioneering technology in which data is transferred between computer chips by a laser, with the optical fiber built directly into semiconductor chips. To circumvent the physical size limitations of VCSELs, in the short term silicon photonics systems can share a larger external laser across multiple transducers, which is a cheaper and more powerful method of transmitting light to all modulators. Research is also under way to integrate lasers directly in the same Silicon photonics chip, which will further reduce cost. Then, silicon photonics microrings within each modulator turn the light on and off to resonate a specific color, increasing the density of wavelengths in the same fiber by a factor of 10 compared to VCSELs.

Additionally, this process enables users to reach target bandwidth at slower individual speeds. That is, instead of driving one color at 100 gigabits and toggling 100,000,000,000X per second, users can harness 10 separate colors toggling at 10 gigabits, which comes with significant power savings and higher signal integrity. Building an exascale system to accommodate both families is the foundation for "computing at the speed of light."

Achieving memory-driven computing

As advancements are made to optimize power usage and bandwidth constraints, memory-driven computing architectures leveraging NVM fabric and photonic communication links will soon enable large-scale supercomputers to process exaflops workloads. This approach places memory, not processing, at the center of the computing platform to deliver optimum scalability, advanced data processing, and persistence.

Memory-driven computing advocates using memory semantics to unify communication and data access across the entire exascale system. This implies that we need to provision memory speed interconnect bandwidth very close to the computing elements. While traditional communication links are measured in tens or hundreds of gigabits per second, memory speed is usually measured in several hundreds of gigabytes per second, so the challenge is to overcome this 10X mismatch. The adoption of optical technologies is vital to achieve the speed and scalability required by these systems.

In a traditional input/output-based network, the network card where the electrical-to-optical conversion takes place is farther from the computing elements, which reduces bandwidth and mandates physically larger and relatively more expensive technology to produce light. As these communication elements migrate closer to the compute socket, the optical converters have to become progressively smaller and more affordable, and despite some physical limitations, the proximity of the optical converters can provide 10X greater bandwidth. Prior to this point, no HPC system could deliver the universal memory semantics necessary to support exascale computing.

Implementing sustainable technologies

Technology is evolving to a point where photonics will play an increasingly active role in HPC solutions. However, the needs of an exascale machine are unlike any system to date, and the relatively small volumes of very high-end systems alone provide few economic incentives for supply chain vendors and technology partners. As a result, developers must find economically viable ways to produce these specialized components—adapting technologies from outside industries and seeking opportunities to repurpose hardware in different domains to avoid the lengthy and expensive R&D process to construct entirely new solutions.

These efforts are a beachhead to develop exascale-specific technologies on an affordable and smaller scale. Once high-end technologies such as VCSELs and microring prototypes have matured, developers can spearhead the supply chain with the most operative low-end technologies to conduct mass development cycles and field experiments. In the pursuit of exascale, large-scale innovation in a conservative product space will be required in order to develop the enabling technologies of exascale.

Why HPE?

In the past several years, Hewlett Packard Enterprise (HPE) has worked to refine photonic transmission to supplement and streamline system communication. In an effort to create a capable exascale machine within the next 5 years, these solutions must quickly surpass the R&D stage of innovation. Science is now translating photonics concepts into mass production prototypes and products, and HPE is centering its resources on converting scientific breakthroughs into economic output, reducing the physical footprint, and securing a supply chain for exascale.

HPE is striving to examine these cutting-edge applications end-to-end to exploit every possible efficiency, from assembly and packaging to scalable frameworks. Integrating photonics into existing solutions and product lines is the culminating step to building a highly composable, cost-effective, and distance-independent system. These principal concerns are driving HPE toward a hybrid IT environment in order to communicate at the speed of memory, and aligning optical technology as close as possible to the computing elements promises to conserve substantial amounts of energy and achieve groundbreaking compute performance.

To catalyze these technological advancements, Hewlett Packard Labs undertook an exceedingly ambitious redesign of system architecture to develop the world's first memory-driven computer. The Machine incorporates the scalability and high speeds of universal memory through a photonics fabric, delivering seamless connectivity, increased security, and unfettered data access. Open system interconnects like Gen-Z are empowering HPE to enhance existing solution architectures and enable future innovation while delivering untapped levels of performance, software efficiency, and competitive advantage.

Conclusion

The HPE approach to exascale is geared to establishing an integrated computing infrastructure that reimagines the capabilities and limitations of today's systems. By revolutionizing the use of photonics, HPE developers are charting a course to the future of HPC to help us foster a vibrant innovation ecosystem and harness insight from massive datasets with the full capacity of an exascale system. HPE stands at the forefront of the next generation of supercomputers, and the solutions to the greatest challenges of today and tomorrow.

Learn more at

https://www.labs.hpe.com/next-next/energy

4AA4-xxxxENW, May 2017