ExaNeSt: Low-Latency Communication and Acceleration in a liquid-cooled energy-efficient Prototype Rack

Manolis G.H. Katevenis (FORTH, Crete, GR), ExaNeSt Coordinator

Abstract:

ExaNeSt, ExaNoDe, and EcoScale, three "sister" projects, as well as EuroEXA, their follow-up project, work for advancing the state-of-the-art in specific technologies that are needed for exascale computing over an energy-efficient platform based on 64-bit ARM processors and Reconfigurable Accelerators (FPGA's).
ExaNeSt focuses on dense packaging, the interconnection network, storage and data access, and also works with a rich set of real, full HPC Applications that have been ported, optimized, and evaluated on the ARM platform.
ExaNeSt has built an HPC Testbed consisting of very dense "QFDB"
daughter cards plugged onto Mezzanine boards and immersed into liquid-cooled blades; this now contains 96 FPGA's, 1.5 TBytes of DRAM, and 6 TBytes of non-volatile solid-state storage, and runs a full stack of systems software and libraries, HPC job manager, and our rich set of Applications, while its size will soon be doubled.
The nodes are interconnected via a 3-Dimensional Torus network with 10 Gbps links and 185 ns per-hop latency.
The virtualized network interfaces, integrated on-chip with the processing cores, feature a zero-copy, protected, user-level RDMA engine with 1024 channels, packetizers, mailbox queues, and offer resilient communication at half-microsecond one-way, one-hop, user-to-user software latency.
Global storage, with per-job SSD/NVM on-demand temporary parallel file system, is provided by BeeGFS with replication extensions; low-latency memory-mapped storage access path, and Virtual Machine (VM) support have been added.
Selected real, full Applications from materials science, climate forecasting, computational fluid dynamics, astrophysics, neuroscience, and an open database suitable for large-scale data analytics have been ported and optimized for this platform.
Evaluation is on-going and shows a significant advantage over Intel-based platforms, ranging from 30% to half, and sometimes one full order of magnitude reduction in Energy-to-Solution.
The ARM processors that have been used out of necessity (A53, the only ones available in FPGAs in 2016) only offer a very modest floating-point performance, so the time-to-solution is quite longer than on Intel platforms, but the full performance advantages, while retaining energy-efficiency, will appear with next generations of ARM processors.
Additional evaluations of FPGA-accelerated applications, in collaboration with the EcoScale and EuroEXA projects, have demonstrated significant energy advantages at performances that are competitive to Intel and Nvidia GPU platforms.