

Paper title

Using Long Vector Extensions for MPI Reductions

Statement

We investigate the impact of the vectorization of MPI reduction operations, and propose an implementation of predefined MPI reduction operations using vector intrinsics (AVX and SVE) to improve the time-to-solution of the predefined MPI reduction operations. With these optimizations, we achieve higher memory bandwidth and an increased efficiency for local computations, which directly benefit the overall cost of collective reductions and applications based on them.

This paper is an extended version of the conference proceedings [1] that considers not only the case of communication benchmarks, but also application benchmarks on different architectures.

We investigated with the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS), which is a molecular dynamics simulation tool from Sandia National Laboratories. It provides different benchmark datasets representing a range of simulation styles and computational expense for molecular-level interaction forces. In our experimental analysis, we evaluate the performance of our reduction operation with LAMMPS granular flow benchmark using the dataset from chute flow (in.chute.scaled).

We also expanded the study and implementation on a different architecture (Arm aarch64) with novel instruction sets - Arm Scalable Vector Extension (SVE). Experiments conducted with novel processors Arm A64FX, show that SVE optimized implementation significantly reduce completion time for collective communication reductions.

We tested AVX implementation with AMD processors in 4.2. We added section 6 with a HPC benchmark to further demonstrate the efficiency of our design. Also, we added Arm SVE based implementation in Section 3.2, 3.4, and experiment results in 4.3.

Figures 2, 5, 8, 9, 12 are additions that present the new implementation and results with new architecture and instructions. We have updated the abstract, introduction, implementation, experiments, conclusion and title of the paper to reflect the supplementary content.

References

- [1] D. Zhong, Q. Cao, G. Bosilca, J. Dongarra, Using advanced vector extensions avx-512 for mpi reductions, EuroMPI/USA '20, Association for Computing Machinery, New York, NY, USA, 2020, p. 1–10. doi:10.1145/3416315.3416316.
URL <https://doi.org/10.1145/3416315.3416316>