

Using Computational Storage Devices: OpenMP/MPI and Charliecloud

(LA-UR-21-27950)

Abstract

As the amount of data in the world grows, researchers have sought solutions to process data more quickly. One potential solution that has been explored in a High Performance Computing (HPC) context is using Computational Storage Devices (CSDs) to process data closer to where it is stored. In our previous experiments, the performance of the six Computational Storage Devices (CSDs) using Apache Spark and Hadoop showed limited applications and complex overhead. To address this overhead, we turned to other computational tools such as Python, C++, and MPI. In addition, we benchmarked less computationally intensive tasks like building container images with Charliecloud. Our results show that Spark contributes significantly to the speed of computation on CSDs. We find the CSDs are ineffective for offloading our operations while our host was under stress. Finally, using CSDs for offloading Charliecloud image building is a promising potential solution for HPC users that we recommend is further researched.