

General I/O Characteristics of HPC Applications on Multiple Peta-scale Supercomputers.

Huong Luu, William Gropp, Marianne Winslett
University of Illinois at Urbana - Champaign

In this project, we study the common I/O characteristics of thousands of HPC applications running on multiple peta-scale supercomputers at the Argonne Leadership Computing Facility (ALCF), the National Energy Research Scientific Computing Center (NERSC) and the National Center for Supercomputing Applications (NCSA). By collecting data from multiple platform, we enable the study of I/O requirements and implementation choices of applications running on different sized systems across multiple science domains. We study data from 3 latest supercomputers at these facilities: Mira (ALCF), Edison (NERSC) and Blue Waters (NCSA). The data set contains 9 month of data from each platform and a total of nearly 1.5 million jobs. This study continues our work in [1] with new data from Blue Waters and different analyses.

First, we study the choice of I/O interface per job or application that includes text I/O, POSIX I/O and MPI-IO. A job is considered to use text I/O (such as `fprintf`, `fscanf`) if it does not use any other library to perform I/O. A job is considered to use POSIX I/O if none of its files is accessed by MPI-IO. Job that uses high-level I/O library which calls to MPI-IO is also considered using MPI-IO. As we can see in Figure 1, the adoption rate of MPI-IO on Mira (GPFS file system) is better than on Edison and Blue Waters (Lustre file system). This result suggests further investigation of MPI-IO performance on Lustre file system.

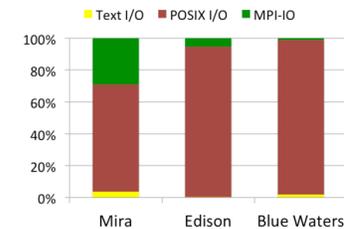


Figure 1: Choices of I/O interface

Second, we study the application's I/O requirement and achievement by considering the maximum I/O throughput and maximum amount of data read/written per application across all its runs. On Blue Waters, 60% of applications never ever in any of its runs achieved more than 1 GB/s throughput. On Mira and Edison, the percentage are 56% and 75% accordingly. On the other hand, 55% of applications on Blue Waters does not read/write more than 1

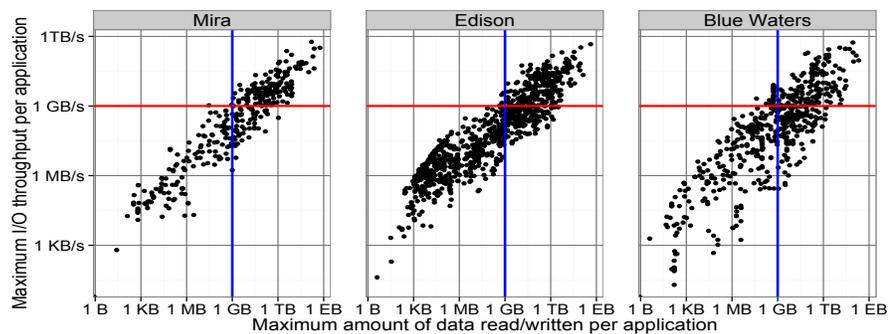


Figure 2: Maximum throughput vs. amount of data per application

GB of data in any of its runs. The percentages on Mira and Edison are 44% and 60%. Even though applications with small amount of data would not be able to achieve high I/O throughput, they should not spend too much time in I/O either.

Lastly, we create some questions to ask for community feedback to help us improve the interpretation of this data.

Reference:

1. Luu et al. A Multiplatform Study of I/O Behavior on Petascale Supercomputers. In *Proceedings of the 24th International Symposium on High-Performance Parallel and Distributed Computing (HPDC '15)*.