Integrating I/O Measurement into Performance Optimisation and Productivity (POP) Metrics

Radita Tapaning Hesti Liem
High Performane Computing
RWTH Aachen University
Aachen, Germany
liem@itc.rwth-aachen.de

I. Introduction

Growing degrees of parallelism within the scientific community increases the importance of the I/O optimization of the HPC application. There are tremendous challenges from the intertwined variables that make measurement difficult. For example, measuring I/O computation time within shared file systems needs to consider cluster workloads, filesystem type, and the chosen programming model. In this work, we do an initial assessment to integrate I/O as a part of the new metric for the Performance Optimization Productivity (POP) project.

POP [1] is a Center of Excellence that provides service to analyze parallel codes for academia and industry within the European Union to promote best practice in parallel programming. The existing metrics arrange components affecting performance in a way to make it easy to read and understand. The general efficiency metric is from the product of parallel efficiency and computation efficiency (IPC). Parallel efficiency is coming from the product of load balance with and communication metrics [2]. The new I/O metric is expected to fit into this format.

II. Motivation

I/O performance can be affected by many factors such as filesystem choice, design decision on the I/O access pattern, or the I/O interfaces. The difference should be observable by measuring an application with the current POP metrics on the different environment that represents components affecting I/O performance.

III. Methodology

The current work will be limited to three components: application programming model, filesystems, and different I/O access patterns to see their impact on the existing POP metrics. The tests will be conducted on several real-world and benchmark applications to represent the variety of cases

Individual I/O measurement using differents measurement tools are conducted in the beginning to give a general overview of the application followed by POP analysis to produce the numbers. Based on both result, we expect to see patterns on which areas different filesystem, programming model, I/O pattern may affect POP building blocks component or whether it's necessary to create another stand-alone metrics from this observation

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreements No 676553 and 824080

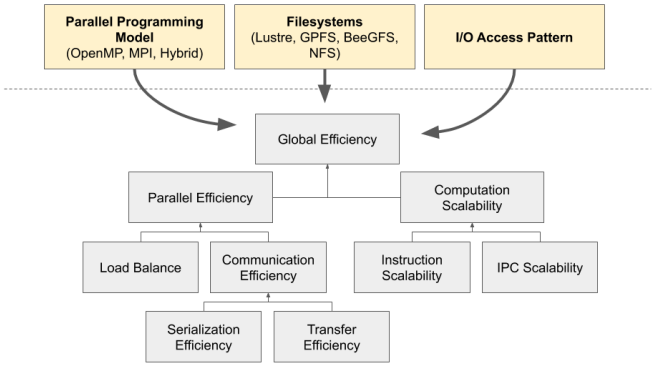

Fig. 1. Illustration on different possible components that can affect the existing POP metrics and represent I/O related impact

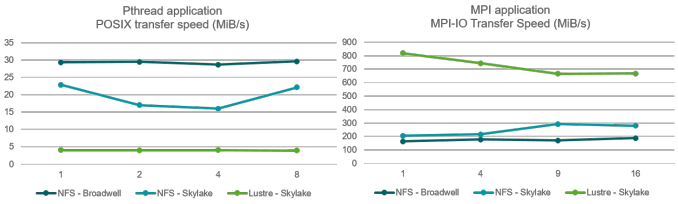

Fig. 2. Two application with different parallel programming model in different filesystems show completely different pattern when its transfer speed measured with Darshan

References

[1] "Performance Optimisation and Productivity," Performance Optimisation and Productivity — A Centre of Excellence in HPC. [Online]. Available: http://www.pop-coe.eu/. [Accessed: 28-Oct-2019].

[2] "Performance Optimisation and Productivity," POP Standard Metrics for Parallel Performance Analysis — Performance Optimisation and Productivity. [Online]. Available: https://pop-coe.eu/node/69. [Accessed: 28-Oct-2019].