July 2017

Parallel IO in the LFRic Infrastructure

Samantha V. ADAMS [a,1], Olga ABRAMKINA [b] and Yann MEURDESOIF [b] and Mike REZNY [c]

[a] *UK Met Office, Exeter, Devon, UK*
[b] *Institut Pierre Simon LaPlace, France*
[c] *Monash University, Australia*

Abstract. This paper describes an investigation into parallel IO using the XIOS framework with the Met Office LFRic Infrastucture. LFRic is currently in development as the replacement for the Met Office Unified Model to enable weather and climate modelling on HPC platforms of the next decade. This will involve running models on hundreds of thousands of cores, therefore a key aspect of the LFRic project is scalability. At this scale, IO is likely to be problematic so LFRic has been exploring the XIOS parallel IO framework as a solution. The purpose of our preliminary experiments was twofold. Firstly, to assess general scalability of our benchmark model using XIOS for parallel output. Secondly, to explore tuning of this benchmark according to some XIOS and Lustre metrics. We successfully ran jobs on up to 82,944 cores on our Cray XC40 machine and found that XIOS was able to efficiently write to one output file whilst still showing scalability with respect to runtime. An investigation into Lustre file striping parameters showed that, with appropriate striping, we could use fewer IO servers than our original guess and still maintain a good performance.

Keywords. Weather and Climate modelling, Parallel IO, XIOS

1. Introduction

Future High Performance Computers will require applications to run on many thousands of cores and potentially hybrid architectures (combinations of multi-core, many-core and GPGPUs). This raises challenges for both legacy and new codes as maintenance and development will increasingly require two sets of skills: in the scientific domain and the computational science domain. The UK Met Office Unified Model (UM) [1] underpins the majority of its work to deliver weather forecasts every day for a variety of customers including the general public. It also provides the modelling and simulation infrastructure for climate research. This model is now roughly 30 years old and will not be sustainable for the future in terms of taking advantage of future trends in High Performance Computing - the challenge is to be able to run models on a variety of (possibly hybrid) architectures making use of many thousands of cores whilst operating with the lowest power consumption possible. The future demand for higher resolutions and more complex calculations whilst still meeting operational timescales for forecasts will be partic-

[1]Corresponding Author: Samantha V. Adams, UK Met Office, Exeter, Devon, UK; E-mail: samantha.adams@metoffice.gov.uk.