Accepted Manuscript

Transferring a petabyte in a day

Rajkumar Kettimuthu, Zhengchun Liu, David Wheeler, Ian Foster, Katrin Heitmann, Franck Cappello

PII: S0167-739X(18)30228-0
DOI: https://doi.org/10.1016/j.future.2018.05.051
Reference: FUTURE 4226

To appear in: *Future Generation Computer Systems*

Received date : 31 January 2018
Revised date : 15 April 2018
Accepted date : 22 May 2018

Transferring a Petabyte in a Day[1]

Rajkumar Kettimuthu[a], Zhengchun Liu[a,*], David Wheeler[d], Ian Foster[a,b], Katrin Heitmann[a,c], Franck Cappello[a]

[a]*Mathematics & Computer Science Division, Argonne National Laboratory, Lemont, IL 60439, USA*
[b]*Department of Computer Science, University of Chicago, Chicago, IL 60637, USA*
[c]*High Energy Physics Division, Argonne National Laboratory, Lemont, IL 60439, USA*
[d]*National Center for Supercomputing Applications, University Of Illinois at Urbana-Champaign, Urbana, IL 61801, USA*

Abstract

Extreme-scale simulations and experiments can generate large amounts of data, whose volume can exceed the compute and/or storage capacity at the simulation or experimental facility. With the emergence of ultra-high-speed networks, researchers are considering pipelined approaches in which data are passed to a remote facility for analysis. Here we examine an extreme-scale cosmology simulation that, when run on a large fraction of a leadership computer, generates data at a rate of one petabyte per elapsed day. Writing those data to disk is inefficient and impractical, and in situ analysis poses its own difficulties. Thus we implement a pipeline in which data are generated on one supercomputer and then transferred, as they are generated, to a remote supercomputer for analysis. We use the Swift scripting language to instantiate this pipeline across Argonne National Laboratory and the National Center for Supercomputing Applications, which are connected by a 100 Gb/s network; and we demonstrate that by using the Globus transfer service we can achieve a sustained rate of 93 Gb/s over a 24-hour period, thus attaining our performance goal of one petabyte moved in 24 hours. This paper describes the methods used and summarizes the lessons learned in this demonstration.

Keywords: Wide area data transfer, GridFTP, Large data transfer, Cosmology workflow, Pipeline

1. Introduction

Extreme-scale scientific simulations and experiments can generate much more data than can be stored and analyzed efficiently at a single site. For example, a single trillion-particle simulation with the Hardware/Hybrid Accelerated Cosmology Code (HACC) [1] generates 20 PiB of raw data (500 snapshots, each 40 TiB), which is more than petascale systems such as the Mira system at the Argonne Leadership Computing Facility (ALCF) and the Blue Waters system at the National Center for Supercomputing Applications (NCSA) can store in their file systems. Moreover, as scientific instruments are optimized for specific objectives, both the computational infrastructure and the codes become more specialized as we reach the end of Moore's law. For example, one version of the HACC is optimized for the Mira supercomputer, on which it can scale to millions of cores, while the Blue Waters supercomputer is an excellent system for data analysis, because of its large memory (1.5 PiB) and 4000+ GPU accelerators. To demonstrate how we overcame the storage limitations and enabled the coordinated use of these two specialized systems, we conducted a pipelined remote analysis of HACC simulation data, as shown in Figure 1.

[1]one PiB = 2^{15} bytes

☆Corresponding author: Zhengchun Liu, Bldg. 240, Argonne National Lab., 9700 S. Cass Avenue, Lemont, IL 60439, USA
Email addresses: kettimut@anl.gov (Rajkumar Kettimuthu), zhengchun.liu@anl.gov (Zhengchun Liu), dwheeler@illinois.edu (David Wheeler), foster@anl.gov (Ian Foster), heitmann@anl.gov (Katrin Heitmann), cappello@mcs.anl.gov (Franck Cappello)