RAPIDS — Resource and Application Productivity through computation, Information, and Data Science

Application Engagement & Community Outreach: Tiger Teams; Liaisons; Outreach

Data Understanding | Platform Readiness | Scientific Data Management

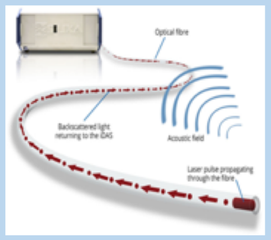
Scientific Data Management: Supporting Scientific Discoveries Through Efficient I/O

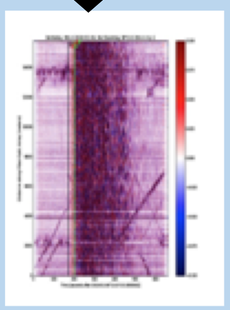
S. Klasky, J. Wu, B. Dong, S. Byna, N. Fortner, B. Geveci, R. Latham, W. Liao, K. Mehta, N. Podhorszki, K. Huck, A. Sim, P. Subedi, P. Davis, M. Parashar

Application Example: Extracting Earthquakes Signals from Dark Fiber Data

Fiber optic cables not being used for data communication (AKA, dark fiber) have been used to collect petabytes of data about ground motion. This data set requires extensive compute power to extract signals for earthquakes, water levels, and other geophysical phenomena.

The use case below shows how RAPIDS technologies are used to reduce the execution time needed for analyzing a particular data set from weeks to seconds





Local Similarity Calculation

Master/Target channel:
Neighbor channel:
Time Window:
Time Shift Range:

(n-5) th channel
n th channel
(n+5) th channel

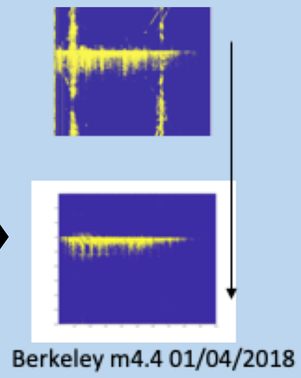
$$S_{n,n+5}(t) = \max_{s\in[-L,L]} \mathrm{corr}(x_n([t-M,t+M]),\ x_{n+5}([t+s-M,t+s+M]))$$

$$S_{n,n-5}(t) = \max_{s\in[-L,L]} \mathrm{corr}(x_n([t-M,t+M]),\ x_{n-5}([t+s-M,t+s+M]))$$

$$S_n(t) = \tfrac{1}{2}(S_{n,n-5}(t) + S_{n,n+5}(t))$$

Xin Xing, etc., "Automated Parallel Data Processing Engine with Application to Large-Scale Feature Extraction", *MLHPC,* 2018,

Matlab (perform with 12 workers)	Many weeks
Python (multithread with 64 threads)	2 hours
C++ with ArrayUDF (MPI with 500 processes)	20 sec.



Berkeley m4.4 01/04/2018

RAPIDS Technology: ArrayUDF

ArrayUDF consolidate common repeated programming efforts involving data partition, data communication, caching, transformation and so on into a single system that supports scientific analysis operations as user-defined functions (UDF)

Common HPC Data Analyses

For each operation *P* **Do**
Develop *P*'s :
- Data management ⇢ *Redundant* X
- Expression execution ⇢ *Diverse*
- Other components: ⇢ *Redundant* X
  parallel,
  communication
  cache,
  etc.

End For

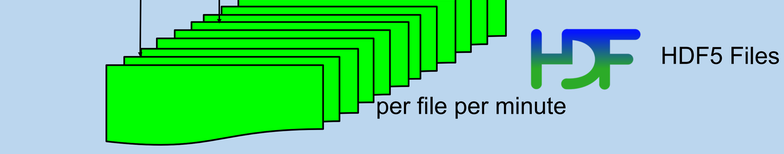
User-defined Functions (UDF)

Operation expression 1 — *Diverse*

UDF API

- Data management
- Generic exec. engine
- Other components: parallel, comm., cache, etc.

One single Shared and optimized middleware

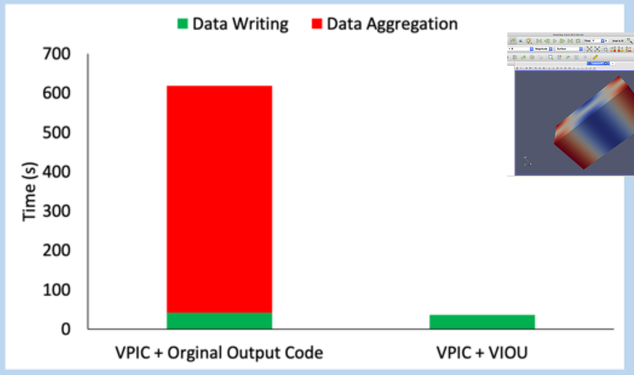


RAPIDS technology: HDF5 Virtual Dataset make accesses to a large number of HDF5 files more convenient

Application Example: VPIC - Vector Particle-In-Cell (VPIC), a particle-in-cell simulation code for modeling kinetic plasmas

VIOU, a VPIC I/O utility

- Structured data organization in HDF5
  - Use n-to-1 I/O pattern to replace n-to-n I/O pattern in field data dump
  - Support multidimensional data
- Fast I/O
  - Merge small I/O operations into large and contagious I/O operation
- XDMF metadata based visualization
- Open Source
  https://bitbucket.org/dbin_sdm/viou/src/master/



" ... the new output saves about 25 to 30 percent of CPU time and improves "time to science" by something like 2 days …. right now I am very happy! "

-- Kilian, Patrick Frank Heiner, physicist at LANL

RAPIDS Technology: Understanding I/O Performance

developed a machine-learning-based I/O performance modeling approach that is robust to HPC system state changes (e.g., hardware degradation, hardware replacement, software upgrades).

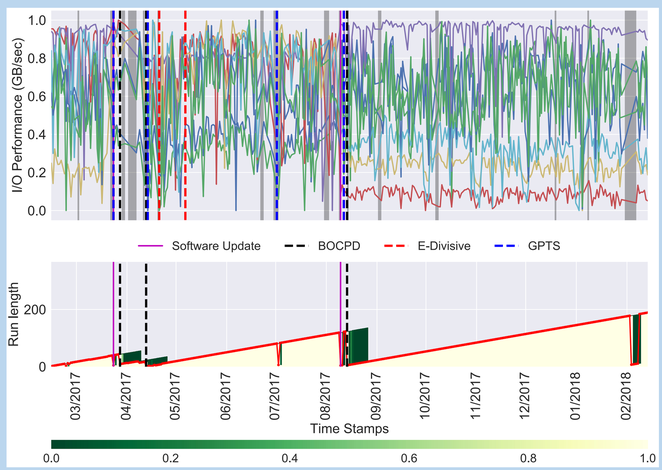
Significance and Impact

Hardware and software changes that affect I/O performance in HPC systems are common but no effective methods to cope up those changes. Our approach automatically finds those changes and adapts the performance model, which can potentially improve the system utilization and application scheduling.

Research Details

- Online Bayesian detection to automatically identify the location of events that lead to changes in near-real time
- Moment-matching transformation that converts the training data collected before the change to be useful for retraining.
- Approach demonstrated on I/O performance data obtained on Lustre file system at NERSC.

S. Madireddy, P. Balaprakash, P. Carns, R. Latham, G. K. Lockwood, R. Ross, S. Snyder, and S. Wild. Adaptive Learning for Concept Drift in Application Performance Modeling, Preprint, ANL/MCS-P9132-0918, 2019.

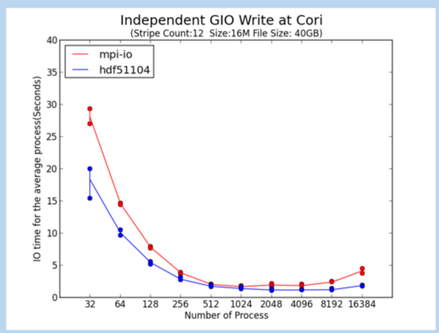


Online method that monitors the change in the I/O performance of an application and adapt the model to these changes

We use application I/O performance data collected on Cori, a production supercomputing system at NERSC, to demonstrate the effectiveness of our approach. The results show that our robust models obtain ***significant reduction in prediction error---from 20.13% to 8.28%*** when the proposed approaches were used in I/O performance modeling.

Application Example: HACC

Exploring the possibility of using HDF5 in HACC,
Found performance did not scale as well as expected
Culprit appeared to be underlying I/O pattern: writing to non-contiguous (by process) blocks
Re-implemented with different data layout and found the expected high performance (see figure)
Proposed a new API routine to explicitly control the allocation order of data chunks to allow high performance with original data layout



The performance information is automatically gathered through **Darshan**

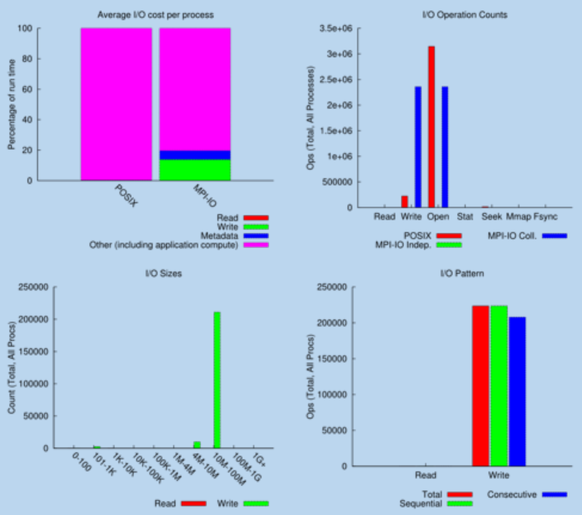
Each job instrumented with Darshan produces a single characterization log file

Darshan command line utilities are used to analyze these log files

Example: Darshan-job-summary.pl produces a 3-page PDF file summarizing various aspects of I/O performance

The figure on the right shows the I/O behavior of a 786,432 process turbulence simulation (production run) on the Mira system at ANL

Application is write intensive and benefits greatly from collective buffering



Argonne NATIONAL LABORATORY · BROOKHAVEN NATIONAL LABORATORY · BERKELEY LAB Lawrence Berkeley National Laboratory · The HDF Group · University of Delaware · Kitware · Lawrence Livermore National Laboratory · Los Alamos National Laboratory EST. 1943 · Northwestern University · Oak Ridge National Laboratory · The Ohio State University · University of Oregon · Rutgers · The University of Utah · U.S. Department of Energy · Office of Science