

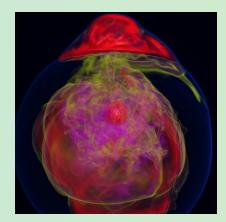
Analyzing Parallel I/O

Julian Kunkel (Georg-August-Universität Göttingen/GWDG) Hendrik Nolte (Georg-August-Universität Göttingen/GWDG) Shane Snyder (Argonne National Laboratory)

November 16th, 2023

Motivation

- An ability to understand parallel I/O performance is critical to ensuring efficient use of HPC resources
 - However, users, system admins, and I/O researchers are often ill-equipped for understanding an increasingly complex HPC I/O ecosystem
- Complex systems
 - Diverse application I/O workloads
 - Deep storage hierarchies utilizing emerging storage hardware
 - I/O middleware to bridge the gap
- In this BoF, we attempt to form a broad community of parallel I/O stakeholders to discuss state-of-the-art in analyzing parallel I/O, with the following goals:
 - Inform community of recent advances in tools/techniques for I/O monitoring
 - Discuss experience/limitations of current approaches
 - Derive a roadmap for future I/O tools to capture, analyze, predict, and tune I/O



Visualization of entropy in Terascale Supernova Initiative application. Image from Kwan-Liu Ma's visualization team at UC Davis.



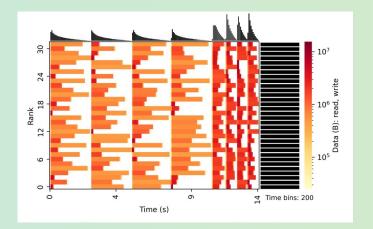
HPE/Cray Aurora system at the ALCF

Agenda

- Talks
 - "The Network Testing Mode in elbencho v3.0", Sven Breuner (VAST)
 - "The HPC IOAnalysis Repository", André Brinkmann (JGU Mainz)
 - "MangolO", Radita Liem (RWTH Aachen)
 - "Understanding Storage Performance using Benchmarking Experiences at GWDG", Hendrik Nolte (GWDG)
- (Audience-driven discussion)

- New heatmap module offers tradeoff between Darshan's traditional condensed data capture and fine-grained tracing with DXT
 - Now available for POSIX, MPI-IO, and STDIO interfaces
- *PyDarshan framework* enables extraction of Darshan log data in popular Python formats (e.g., pandas dataframes) and forms foundation of a newly developed Darshan job summary tool
 - Available via PyPI: pip install darshan
- (In progress) New *DAOS modules* providing in-depth characterization of various DAOS APIs
 - DAOS file system (DFS) as well as native object APIs (key-val, array)

- New heatmap module offers tradeoff between Darshan's traditional condensed data capture and fine-grained tracing with DXT
 - Now available for POSIX, MPI-IO, and STDIO interfaces
- PyDarshan framework enables extraction of Darshan log data in popular Python formats (e.g., pandas dataframes) and forms foundation of a newly developed Darshan job summary tool
 - Available via PyPI: pip install darshan
- (In progress) New *DAOS modules* providing in-depth characterization of various DAOS APIs
 - DAOS file system (DFS) as well as native object APIs (key-val, array)



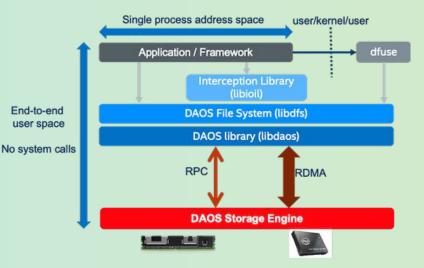
Heatmaps showcase I/O intensity across time (x-axis) and ranks (y-axis). Comparing across interfaces can offer insights into I/O behavior (e.g., collective I/O)

- New heatmap module offers tradeoff between Darshan's traditional condensed data capture and fine-grained tracing with DXT
 - Now available for POSIX, MPI-IO, and STDIO interfaces
- *PyDarshan framework* enables extraction of Darshan log data in popular Python formats (e.g., pandas dataframes) and forms foundation of a newly developed Darshan job summary tool
 - Available via PyPI: pip install darshan
- (In progress) New *DAOS modules* providing in-depth characterization of various DAOS APIs
 - DAOS file system (DFS) as well as native object APIs (key-val, array)

Job Summ	nary			
Job ID	12345			
User ID	6789			
# Processes	512			
Run time (s)	727.0000	27.0000		
Start Time	2022-03-02 13:52:46			
End Time	2022-03-02 14:04:5	2		
Command Line	./E3SM-I0/build/sr	./E3SM-I0/build/src/e3sm_io E3SM-I0-inputs/i_case_1344p.nc -k -o can_l_out.nc -a pnetcdf -x canonical -r 200		
Darshan L	og Informati			
Log Filename		e3sm_io_heatmap_only.darshan		
Runtime Library Version		3.3.1		
Log Format Version		3.21		
POSIX (ver=4) Module Data		20.34 KiB		
MPI-IO (ver=3) Module Data		0.49 KiB		
PNETCDF_FILE (ver=2) Module Data		0.11 KiB		
LUSTRE (ver=1) Module Data		9.36 KiB		
STDIO (ver=2) Module Data		0.08 KIB		
APMPI (ver=1) Module Data		136.94 KiB		
HEATMAP (ve	er=1) Module Data	318.85 KiB		
I/O Sum	mary			
Heat Map:	HEATMAP I	MPIIO	Heat Map: HEATMAP POSIX	
ما المرابد.		Nije.	ավերկ կարությունների հետություն	
10			5.22	
435 -		e	406	
		100 H	336 ***********************************	
306 - 7 204 -				

New PyDarshan-based job summary tool generates detailed HTML reports of job I/O behavior

- New heatmap module offers tradeoff between Darshan's traditional condensed data capture and fine-grained tracing with DXT
 - Now available for POSIX, MPI-IO, and STDIO interfaces
- PyDarshan framework enables extraction of Darshan log data in popular Python formats (e.g., pandas dataframes) and forms foundation of a newly developed Darshan job summary tool
 - Available via PyPI: pip install darshan
- (In progress) New *DAOS modules* providing in-depth characterization of various DAOS APIs
 - DAOS file system (DFS) as well as native object APIs (key-val, array)



Various access methods for DAOS users.

Figure courtesy of Intel

Thanks to all for attending!

BoF website will be updated with speaker materials, notes, etc.: <u>https://hps.vi4io.org/events/2023/sc-analyzing-io</u>