ADIOS: Storage and in situ I/O:

Accelerating Scientific Knowledge Discovery with the Adaptable Input Output System

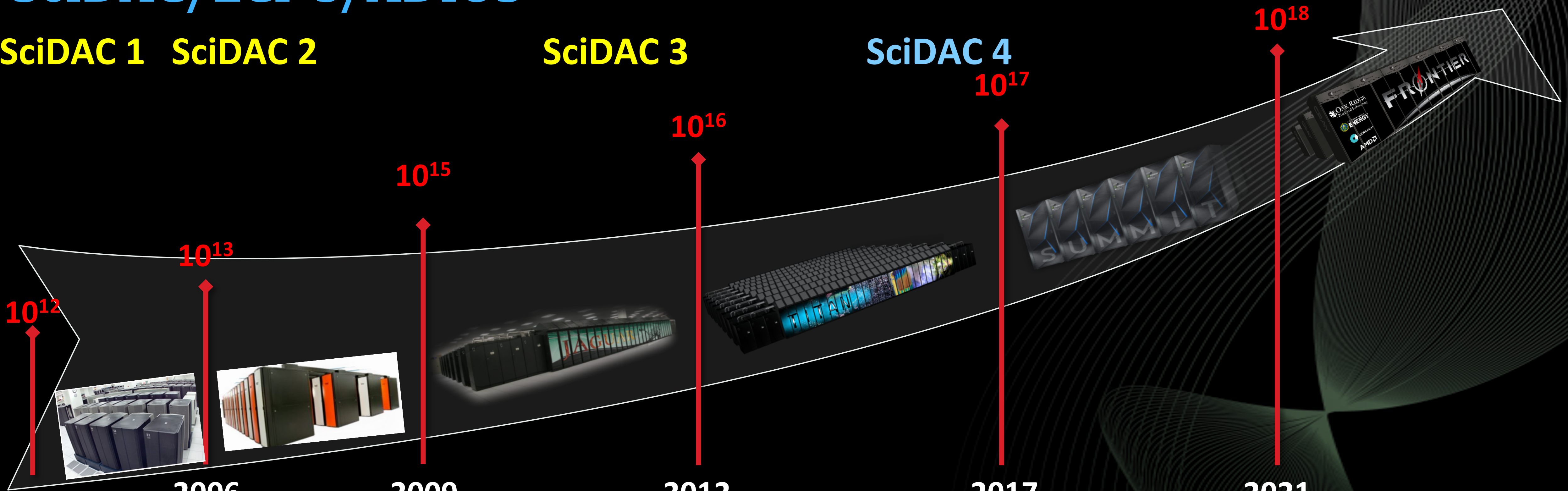
Presenter: S. Klasky[123],
With contributions from: N. Podhorszki[1] , A. Gainaru[1],M. Wolf[1], C. Atkins[8], William Godoy[1], Matthew Wolf[1] , Ruonan Wang[1], Chuck Atkins[8], Greg Eisenhauer[3], Junmin Gu[7], Philip Davis[5], W. Godoy[1], Jong Choi[1], Kai Germaschewski [10], Kevin Huck [11], Axel Huebl[7], Mark Kim[1], James Kress[1], Tahsin Kurc[1], Jeremy Logan[1], Kshitij Mehta[1], Franz Poeschel[8], Eric Suchyta[1], Keichi Takahashi, Lipeng Wan[1], Pradeep Subedi, Mark Ainsworth[19], Berk Geveci [8], Ben Whitney[1]

[1] Oak Ridge National Laboratory, Computer Science and Mathematics Division
[2] University of Tennessee, Knoxville, Department of Electrical Engineering and Computer Science
[3]Georgia Tech, School of Computer Science
[4]New Jersey Institute of Technology
[5]Rutgers University
[6]National Science Foundation, OAC
[7] Lawrence Berkeley National Laboratory
[8] Kitware
[9] Brown University
[10] University of New Hampshire
[11] University of Oregon

Community BOF, 3/30/2021

SciDAC/LCF’s/ADIOS

SciDAC 1 SciDAC 2 SciDAC 3 SciDAC 4



2001 2006 2009 2012 2017 2021

Seaborg: 20 TF **Jaguar:** 25 TF **Jaguar 2:** 2.3 PF **Titan:** 27 PF **Summit:** 200 PF **Frontier 1500 PF**

ADIOS timeline

BP self-describing file format ADIOS 0.01

Data Staging ADIOS 0.1

Research writing performance

Research reading performance

1.2 Data Staging

ADIOS 1.0 released

1.4 weak code coupling

1.6 index, compression

1.8 Query /Indexing

1.10 Recovery/Ease of use

1.11 compression

1.12 Burst buffer

2.2 New framework in C++

2.3 Summit optimizations

SciDAC funding in purple boxes

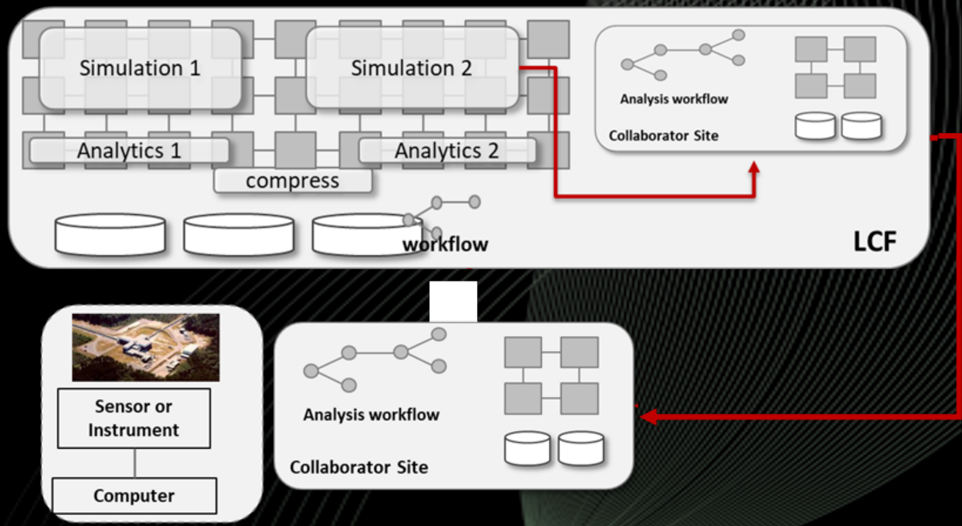
ECP funding

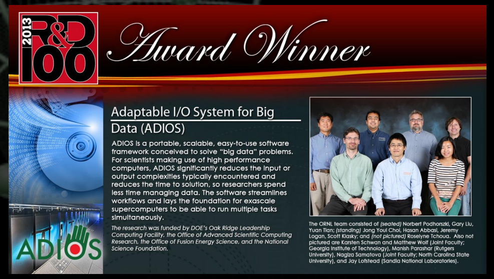
ASCR Research funding

OLCF funding

ADIOS: High-Performance Publisher/Subscriber I/O framework

- An abstraction to allow for high-performance I/O to/from storage and for in situ processing
- Utilizes a publish/subscribe mechanism with self-describing data
- Optimized I/O engines for C/R, strong/loose in situ coupling WAN data streaming, and in-memory object storage
    - Fast Writing/Reading: BP4
    - DAOS optimizations: BP-DAOS
    - Write from GPU: BP-GPUDirect
    - Compatibility with HDF5: HDF5
    - Weak Code Coupling: SST
    - Tight Code Coupling: SSC
    - WAN streaming: DataMan
    - In memory object store: DataSpaces
    - Works with ECP reduction libs: MGARD, SZ, ZFP
- Typical for applications to achieve > 1 TB/s on Summit





Sustainability: is a primary goal of the ADIOS project

- Nightly testing
  - Testing on many different platforms
- Continuous Integration
  - Only allow tested code to be merged
  - Almost 2,000 tests for each commit
- Static and dynamic analysis reports
  - Compile-time and run-time analysis
- Code coverage
  - Level of testing
- External testing
  - Allow feedback from user projects

Nightly testing on target HPC platforms

Nightly

Site	Build Name	Update Revision	Configure Error	Configure Warn	Build Error	Build Warn	Test Not Run	Test Fail	Test Pass
aaargh.kitware.com	Linux-EL7_GCC7	237f1b	0	0	0	0	0	0	95
aaargh.kitware.com	Linux-EL7_Intel17	237f1b	0	0	0	0	0	0	95
aaargh.kitware.com	Linux-EL7_Intel18	237f1b	0	0	0	0	0	0	95
aaargh.kitware.com	Linux-EL7_GCC7_MPICH	237f1b	0	0	0	0	1	0	203
aaargh.kit		237f1b	0	0	0	0	1	0	203
cori.nersc.gov	Linux-CrayCLE6-KNL_GCC_CrayMPICH	1b	0	0	0	0	0	61	123
cori.nersc.gov	Linux-CrayCLE6-KNL_Intel_CrayMPICH	237f1b	0	0	0	0	0	61	123
summitdev.ccs.ornl.gov	Linux-EL7-PPC64LE_GCC-7.1.0_NoMPI		0	0	0	0	0	6	88

age report

Coverage started on Friday, September 28 2018

Coverage Summary	
Total Coverage	72.75
Tested lines	10703
Untested lines	4010
Files Covered	166 of 166
Files Satisfactorily Covered	122
Files Unsatisfactorily Covered	44

Coverage Legend
Satisfactory coverage
Unsatisfactory coverage
Dangerously low coverage

Show coverage over time

Directories (32) | No Executable Code (0) | Zero (11) | Low (11) | Medium (29) | Satisfactory (74) | Complete (41) | All (166)   Show Filters

Show 25 entries

Directory	Status	Percentage	Lines not covered	Priority	Labels
source/adios2/toolkit/aggregator/mpi	Satisfactory	93.14%	12/175	None	
source/adios2/engine/insitumpi	Satisfactory	93.08%	54/780	None	
source/adios2/engine/bp3	Satisfactory	91.45%	26/304	None	
source/adios2/toolkit/sst/dp	Satisfactory	87.86%	34/280	None	
source/adios2/toolkit/format/bp3	Satisfactory	85.29%	350/2380	None	
source/adios2/toolkit/sst/cp	Satisfactory	84.33%	376/2400	None	
source/adios2	Satisfactory	83.33%	6/36	None	
source/adios2/toolkit/format/dataman	Satisfactory	82.52%	50/286	None	
source/adios2/engine/dataman	Satisfactory	82.25%	41/231	None	
source/adios2/toolkit/transport/wan	Satisfactory	81.05%	29/153	None	
source/adios2/toolkit/transportman/dataman	Satisfactory	80.95%	24/126	None	
source/adios2/engine/sst	Satisfactory	80.71%	65/337	None	
bindings/CXX11/cxx11	Satisfactory	79.37%	26/126	None	
source/adios2/core	Satisfactory	74.46%	238/932	None	
source/adios2/toolkit/transportman	Satisfactory	74.03%	40/154	None	
source/adios2/engine/hdf5	Satisfactory	71.52%	47/165	None	
source/adios2/operator/callback	Satisfactory	71.43%	4/14	None	
bindings/Python	Satisfactory	70.83%	112/384	None	
source/adios2/toolkit/format	Satisfactory	70.00%	3/10	None	
source/adios2/toolkit/transport/file	Medium	63.10%	124/336	None	
bindings/Fortran/f2c	Medium	62.60%	135/361	None	
source/adios2/helper	Medium	59.69%	472/1171	None	
source/utils/bpls2	Medium	58.01%	451/1074	None	
source/adios2/toolkit/transport	Medium	56.67%	26/60	None	
source/adios2/toolkit/interop/hdf5	Satisfactory	56.39%	297/681	None	

Showing 1 to 25 of 32 entries   First Previous 1 2 Next Last

Add more commits by pushing to the **sst-bp-compression-tests** branch on **JasonRuonanWang/ADIOS2**.

All checks have passed
13 successful checks   Hide all checks

- **Codacy/PR Quality Review** — Up to standards. A positive pull request. Details
- **cdash** — Build and test results available on CDash   Required   Details
- **ci/circleci: el7** — Your tests passed on CircleCI!   Required   Details
- **ci/circleci: el7-gnu7** — Your tests passed on CircleCI!   Required   Details
- **ci/circleci: el7-gnu7-openmpi** — Your tests passed on CircleCI!   Required   Details

This branch has no conflicts with the base branch
Merging can be performed automatically.

Merge pull request   or view command line instructions.

ADIOS Approach: "How"

- I/O calls are of declarative nature in ADIOS
  - which process writes what: add a local array into a global space (virtually)
  - adios_close() indicates that the user is done declaring all pieces that go into the particular dataset in that timestep
- I/O strategy is separated from the user code
  - aggregation, number of sub-files, target file-system hacks, and final file format not expressed at the code level
- This allows users to choose the best method available on a system without modifying the source code
- This allows developers
  - to create a new method that's immediately available to applications
  - to push data to other applications, remote systems or cloud storage instead of a local filesystem

W. Godoy, N. Podhorszki, R. Wang, et al, ADIOS 2: The Adaptable Input Output System. A framework for High-Performance Data Management, SoftwareX, 2020.

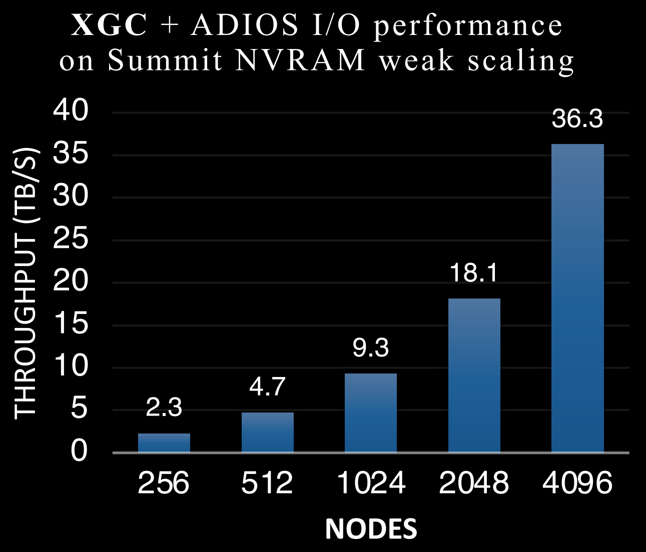
Creating I/O abstractions to accelerate I/O to storage

- One change in the code or input file, to specify the engine

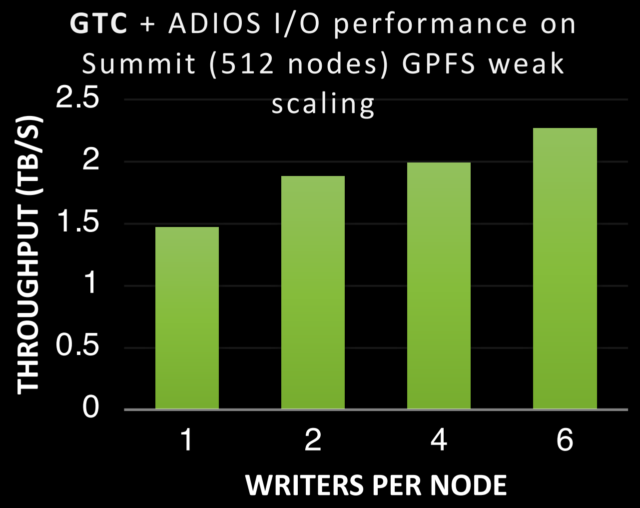
```
adios2::Engine writer = io.Open("analysis.bp",
adios2::Mode::Write);

writer.BeginStep()

writer.Put(varT, T.data());

writer.EndStep()

writer.Close()


adios2::Engine reader = io.Open("analysis.bp",
adios2::Mode::Read);

reader.BeginStep()

adios2::Variable<double> T =
reader.InquireVariable("Temperature");

std:vector<double> t;

reader.Get(varT, t);

reader.EndStep()

reader.Close()
```

The APIs are identical for code coupling

36.3 TB/s to NVRAM



2.3 TB/s to GPFS



ADIOS performance results (measured by the app teams/not us)

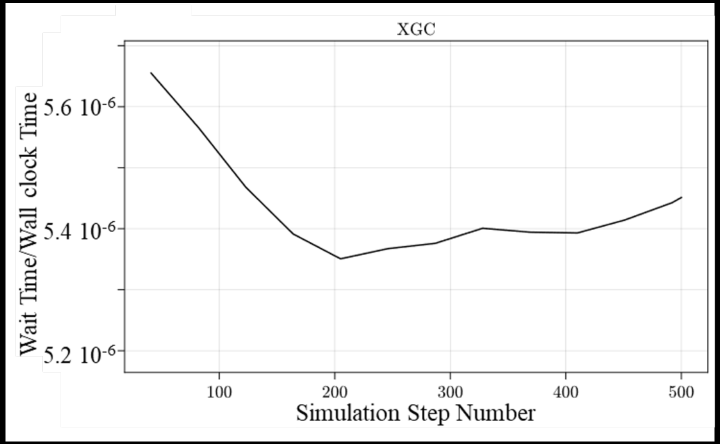
WDMApp

https://github.com/PrincetonUniversity/XGC-Devel

From the WDMApp annual ECP review

XGC on 512 Summit nodes
GENE on 6 Summit nodes
N_m = 9,640,480 vertices
N_p = 8,922 particles/vertex
Timestep = 61.4 seconds.

XGC wait time during charge coupling



Contact: Amitava Bhattacharjee (PPPL)

E3SM-MMF

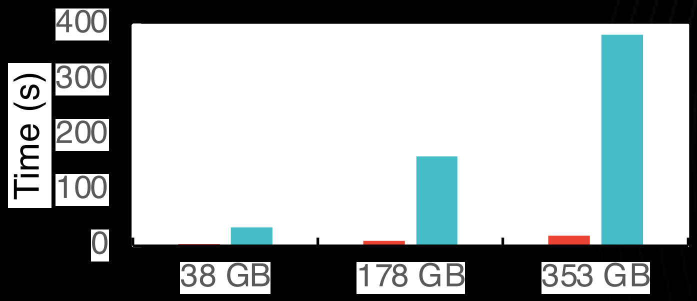
https://github.com/E3SM-Project/scorpio/tree/master

ADIOS 2.x Port is integrated into master SCORPIO

SCREAM project evaluated it on their own and found 4-5x improvement in IO using ADIOS for TBs of data

New I-Case stresses IO

I-Case benchmark on Summit



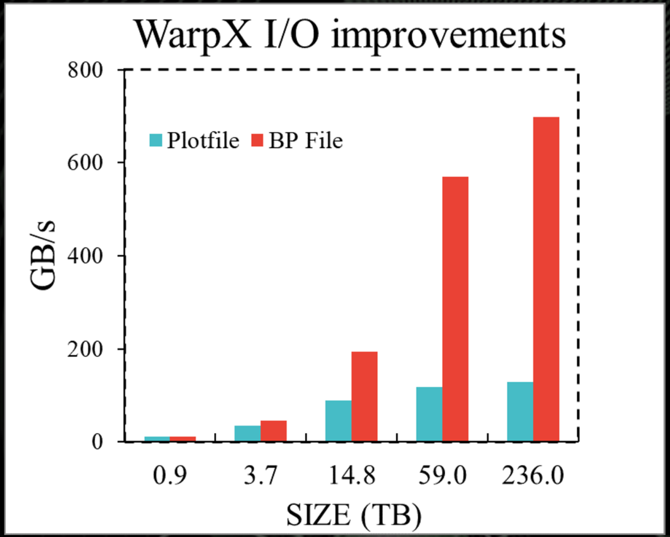
Simulating 1 day, 5 days and 10 days
Writing data every simulated hour
Run on Summit, 1344 MPI processes

Contact: Mark Taylor (SNL), I-Case Peter Thornton (ORNL), SCREAM Peter Caldwell (LLNL)

WarpX

BP4 improved append performance for ADIOS and now applications can see the benefits of that

WarpX and in general, OpenPMD users can get high throughput



WarpX on Summit, weak scaling
6 GPUs, up to 256 nodes
ADIOS vs original AMReX Plot files

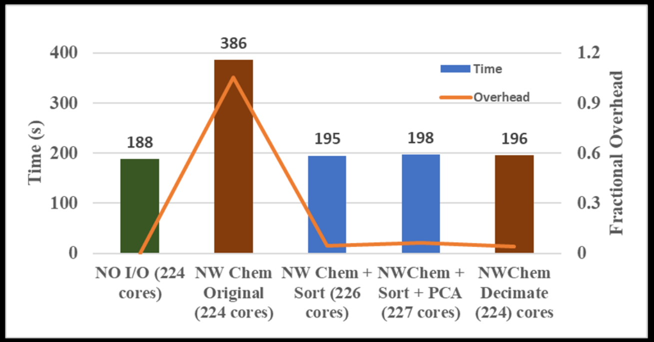
Contact: Jan-Luc Vay, Axel Huebl (LBNL)

ADIOS performance results (measured by the app teams/not us)

NWChem

In situ sorting of atom trajectories can save 50% of runtime

Motion correction with PCA analysis (pbdR script) in situ



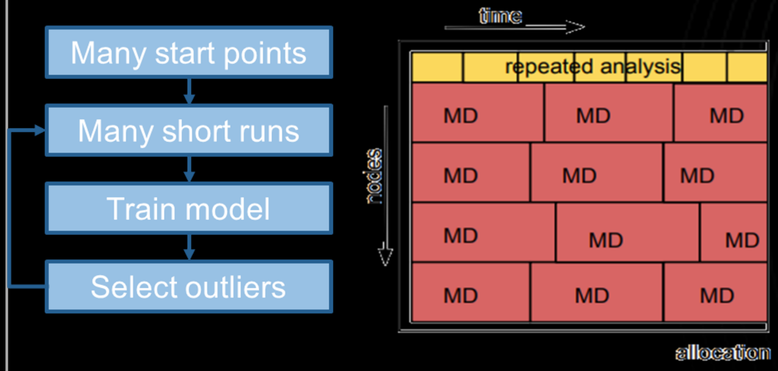
Contact: Tjerk Straatsma (ORNL)

CANDLE/DeepDriveMD

CODAR collaboration for accelerating sampling of macromolecule potential energy surface via online coupling

Many concurrent MD runs + online training + inference (outlier search)

ADIOS for async collection of MD results to training allows for continuous simulation running and training



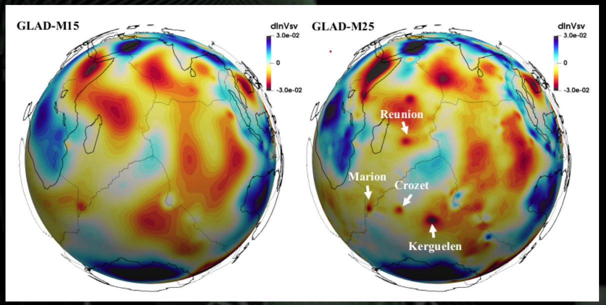
Contact: Arvind Ramanathan, Igor Yakushin (ANL)

Specfem3D_globe

The Adaptable Seismic Data Format (ASDF) was developed that leverages the Adaptable I/O System (ADIOS) parallel library.

It allows for recording, reproducing, and analyzing data on large-scale supercomputers

1.5 PB of data is produced in every workflow step, which is fully processed later in adjoint simulation
https://www.olcf.ornl.gov/2019/07/05/tromp-titan/



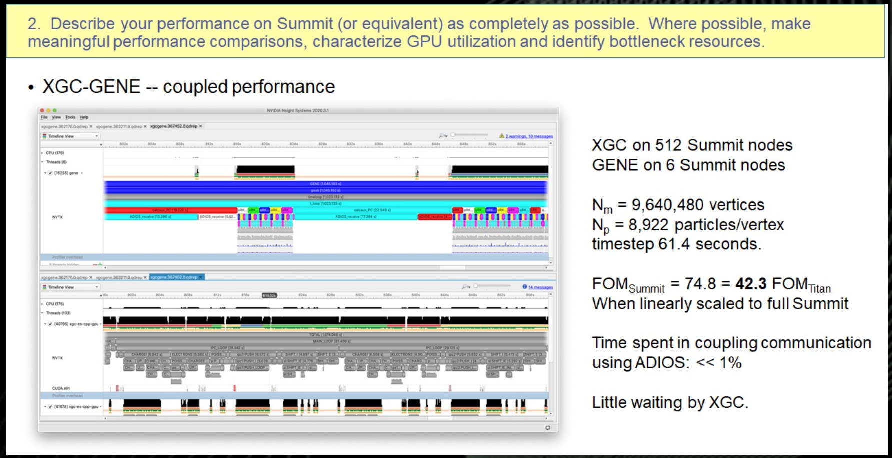
Global adjoint tomography—model GLAD-M25, Geophysical Journal International, Volume 223, Issue 1, October 2020, Pages 1–21, https://doi.org/10.1093/gji/ggaa253

Contact: Jeroen Tromp, Princeton University

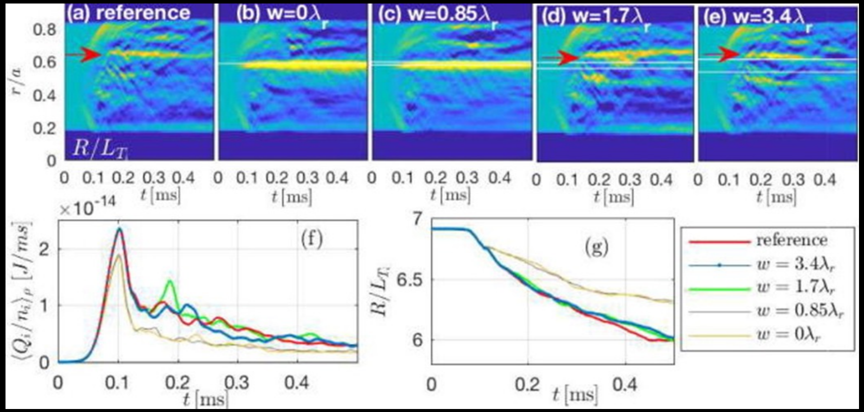
2.2.2.05 ADSE12-WDMApp: High-Fidelity Whole Device Modeling of Magnetically Confined Fusion Plasmas

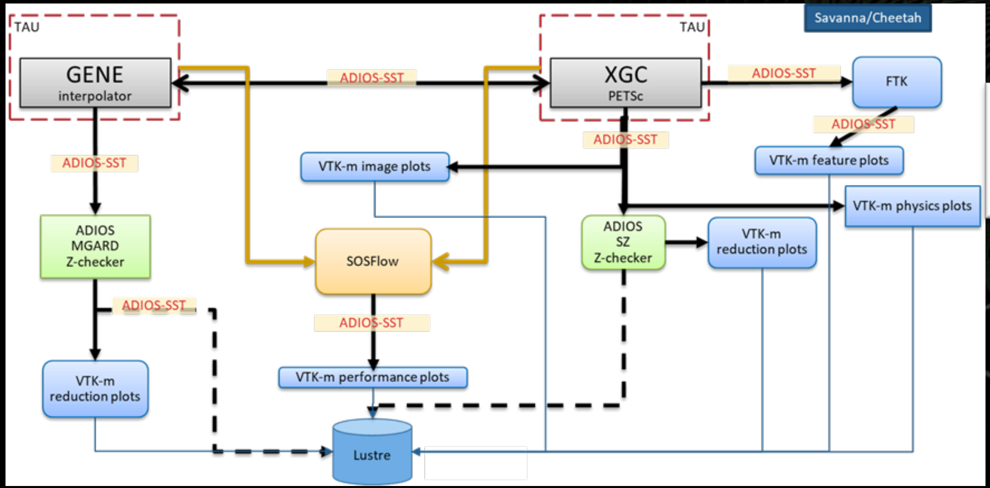
PI: Amitava Bhattacharjee, PPPL,
C. S. Chang, PPPL

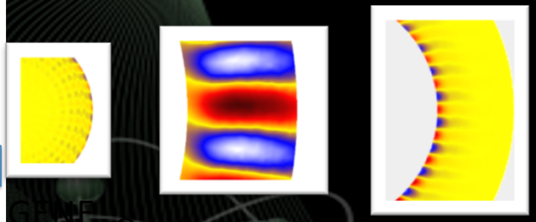
- Different physics solved in different physical regions of detector (spatial coupling)
- Core simulation: **GENE**
  Edge simulation: **XGC**
  Separate teams, **separate codes**
- Recently demonstrated first-ever successful kinetic coupling of this kind
- Data Generated by one coupled simulation is predicted to be > 10 PB/day on Summit



From FY21 WDMApp Review







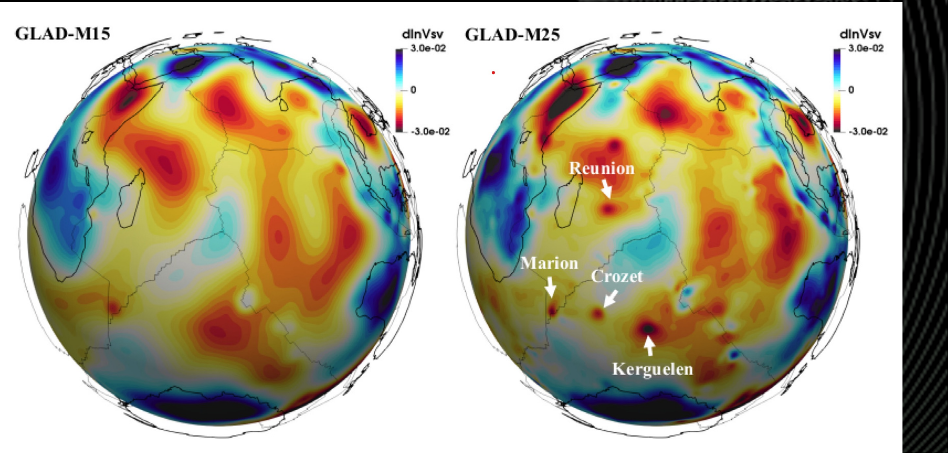
Dominski, J., et al. "Spatial coupling of gyrokinetic simulations, a generalized scheme based on first-principles." *Physics of Plasmas* 28.2 (2021): 022301.
Merlo, G., et al. "First coupled GENE–XGC microturbulence simulations." *Physics of Plasmas* 28.1 (2021): 012303.
Cheng, Junyi, et al. "Spatial core-edge coupling of the particle-in-cell gyrokinetic codes GEM and XGC." *Physics of Plasmas* 27.12 (2020): 122510.

Results: Seismic Tomography Workflow (PBs of data/run)

PI: Jeroen Tromp, Princeton

Scientific Achievement

- Most detailed **3-D model of Earth**’s interior showing the entire globe from the surface to the core–mantle boundary, a depth of 1,800 miles



Map views at 250 km depth of vertically polarized shear wave speed perturbations in GLAD-M15 (2017) and GLAD-M25 (2020) in the Indian Ocean. New features have emerged in GLAD-M25, such as the Reunion, Marion, Kerguelen, Maldives, Seychelles, Cocos and Crozet hotspots.

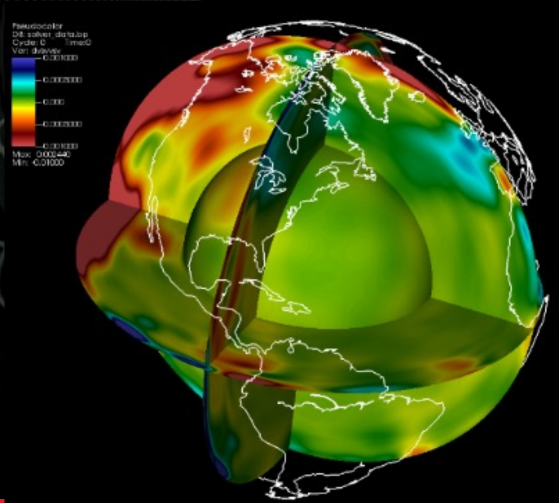
Significance and Impact

- Updated (transversely isotropic) global seismic model GLAD-M25 where no approximations were used to simulate how seismic waves travel through the Earth. The data sizes required for processing are challenging even for leadership computer
- 7.5 PB of data is produced in a single workflow step
    - which is fully processed later in another step.

Improvement by appending steps

- 3200 nodes ensemble run, 19200 GPUs
- 50 tasks at once
- 5.2 TB per task in 133 steps
- 260 TB total per 50 tasks
- 7.5 PB per 1500 tasks (total run)

50 tasks, 133 steps, 3200 nodes	Time
No I/O	94s
BP3, one file per step	235s
BP4 one dataset per job 133x reduction in # of files	156s



Wenjie Lei, Youyi Ruan, Ebru Bozdağ, Daniel Peter, Matthieu Lefebvre, Dimitri Komatitsch, Jeroen Tromp, Judith Hill, Norbert Podhorszki, David Pugmire **Global adjoint tomography—model GLAD-M25**, Geophysical Journal International, Volume 223, Issue 1, October 2020, Pages 1–21, https://doi.org/10.1093/gji/ggaa253

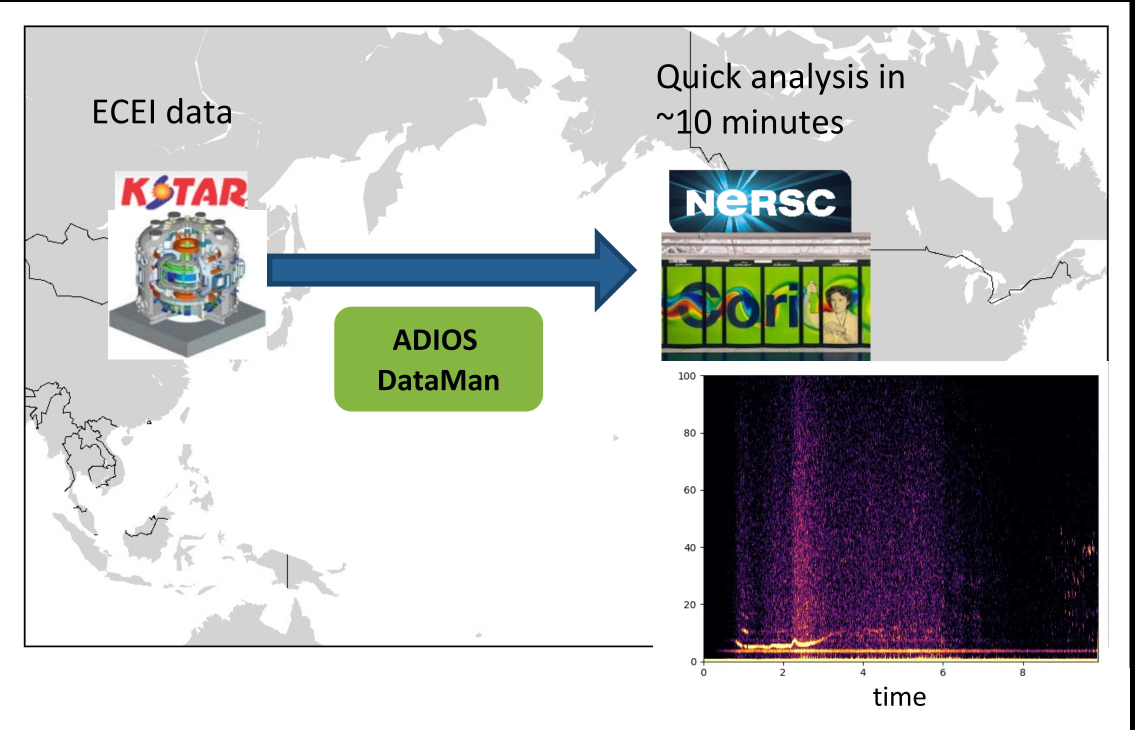
FES Highlight: Established capability for near-real time networked analysis of big KSTAR data at NERSC (PPPL, ORNL, ESnet, NERSC, KSTAR, KISTI)

Objectives

- Research and develop a streaming workflow framework, to enable near-real-time streaming analysis of KSTAR data on a US HPC
- Allow the framework to adopt ML/AI algorithms to enable adaptive near-real-time analysis on large data streams

Impact

- Created a framework to enable US fusion researchers to have broader and faster access to the KSTAR data, enabling
  - Faster analysis of data
  - Faster and autonomous utilization of ML/AI algorithms for incoming data
  - More informed steering of experiment
  - Quicker utilization of US HPC for KSTAR collaboration



Accomplishments

- Created end-to-end Python framework DELTA, streams data using ADIOS DataMan over WAN (at rates > 4 Gbps), asynchronously processes on multiple workers with MPI multi-threading
- Applied to KSTAR streaming data to NERSC Cori. Reduces time for an ECEi analysis from 12 hours on single-process to 10 minutes on 6 Cori nodes.
- Implemented deep convolutional neural networks for working with multi-scale fusion data, e.g. ECEi, for recognizing events of interest.[2]
- On-going: improve "adaptive" nature of data stream: adaptive compression at KSTAR source

Churchill RM, Klasky et al. A Framework for International Collaboration on ITER Using Large-Scale Data Transfer to Enable Near-Real-Time Analysis. Fusion Science and Technology. 2021 Feb 17;77(2):98-108

[2]R.M. Churchill, NeurIPS 2019

E3SM

PI: Mark Taylor, SNL

E3SM Energy Exascale Earth System Model

Search | Select Language | GitHub | ESGF

ABOUT | RESEARCH | MODEL | DATA | PUBLICATIONS | RESOURCES

ABOUT

Vision and Mission
- Long Term Roadmap
- Science Drivers

Organization
- The Leadership Team
- NGD Sub-Projects
  - NGD Atmospheric Physics
  - NGD Land and Energy
  - NGD Nonhydrostatic Atmosphere
  - NGD Software and Algorithms
  - NGD BISICLES
  - NGD Coastal Waves

Events
- E3SM Conferences
  - 2019 E3SM Spring Meeting
- E3SM Tutorials
- All-Hands Presentations

Collaboration
- Collaboration Request
- Ecosystem Projects
- Closely Related Projects

News

Home > About > News > PIO2 + ADIOS = Performance Improvement

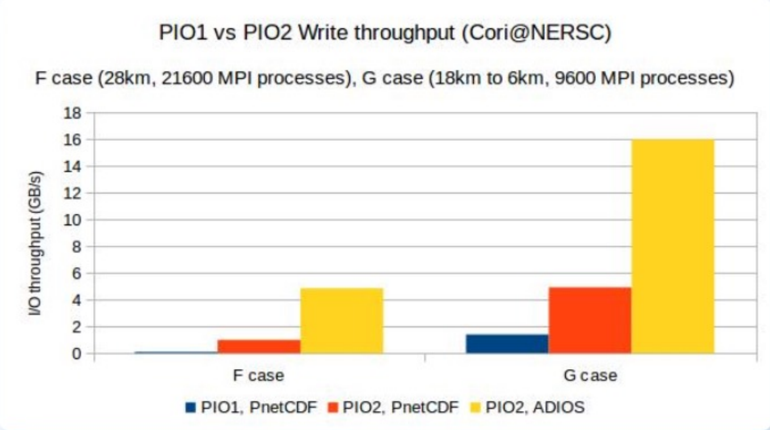
PIO2 + ADIOS = PERFORMANCE IMPROVEMENT

August 8, 2019

Releases

The **P**arallel **I**nput/**O**utput (I/O) library (PIO) is used by all the model components in E3SM for reading input and writing model output. The library supports reading and writing data using low-level I/O libraries like PnetCDF and NetCDF. The user data in E3SM is not typically decomposed across the compute processes in an "I/O friendly" way and this requires some data rearrangement, supported by PIO, before using these low-level I/O libraries to write the model data. PIO2 is the latest version of the PIO library that includes a complete rewrite of the original Fortran PIO (PIO1) library into C/C++. PIO2 also supports advanced caching and data rearrangement algorithms and includes support for more low-level I/O libraries.

In recent months developers added support in PIO to read and write data using the The Adaptable I/O System (ADIOS) library. The ADIOS library provides a flexible way to describe scientific data that may be read, written or processed outside a simulation. The library supports MPI individual I/O, MPI collective I/O, POSIX I/O, asynchronous I/O and a visualization engine to process scientific data. The library also supports a NULL output option to disable all model output. The data is written out in the ADIOS file format (which uses the .bp extension) and can be converted to the NetCDF format using a post processing tool included with PIO. Since the user data is decomposed across multiple compute processes it typically requires some data rearrangement in PIO or the low level I/O libraries to write the data efficiently in contiguous chunks, as required by the NetCDF format. Since ADIOS writes data out in multiple files and does not require data to be written out in contiguous chunks, it saves time by partially rearranging data and reducing contention in the file system.



As shown in the adjacent figure, the performance of PIO2 was measured using two E3SM simulation configurations on Cori: a configuration with high resolution atmosphere and a configuration with high resolution ocean. The high resolution atmosphere case, F case, runs active atmosphere and land models at 1/4 degree (28km) resolution, the sea ice model on a regionally refined grid with resolutions ranging from 18km to 6km and the runoff model at 1/8 degree resolution. The atmosphere component is the only component that writes output data. In this configuration all restart output is disabled and the component only writes history data, which has been historically shown to have poor I/O performance. A one-day run of this configuration generates two output files with a total size of approximately 20 GB. The high resolution ocean case, G case, runs ocean and sea ice models on a regionally refined grid with resolutions ranging from 18km to 6km. All output from the sea ice component is disabled in this configuration. A one-day run of this configuration generates 80 GB of model output from the ocean model.

The I/O write throughput for the F case was < 100 MB/s with PIO1 on Cori. PIO2 with its improved caching and data rearrangement algorithms provides a 10x improvement in the write throughput. Using ADIOS as the I/O library provides about 4x improvement over the PnetCDF library and results in a 40x improvement (ignoring post processing to convert ADIOS files to NetCDF) over PIO1. The I/O write throughput for the G case was about 1.4 GB/s with PIO1. PIO2 provides a 4x improvement in performance compared to PIO1 and using ADIOS with PIO2 provides a further 4x improvement in the write performance, resulting in a 16x improvement in write performance on Cori.

The high resolution G case was also run with PIO2 on Summit, using PnetCDF as the low-level I/O library, and the measured I/O write throughput was around 22 GB/s. Using ADIOS as the I/O library and leveraging the asynchronous I/O feature in ADIOS provided a 5x performance improvement in the write throughput compared to PnetCDF. Increasing the model output (higher output frequency) can further increase the ADIOS I/O throughput to about 7x compared to PnetCDF. This is a work in progress and the developers will continue to measure and tune performance of PIO2 on Summit.

https://e3sm.org/pio2-adios-performance-improvement/

SKA

https://www.icrar.org/summit/

WORLD'S FASTEST SUPERCOMPUTER PROCESSES HUGE DATA RATES IN PREPARATION FOR MEGA-TELESCOPE PROJECT

Wang, Ruonan, et al. "Processing full-scale square kilometre array data on the summit supercomputer." *2020 SC20: International Conference for High Performance Computing, Networking, Storage and Analysis (SC)*. IEEE Computer Society, 2020.

HOME / WORLD'S FASTEST SUPERCOMPUTER PROCESSES HUGE DATA RATES IN PREPARATION FOR MEGA-TELESCOPE PROJECT

SHARE ARTICLE: TWITTER FACEBOOK

TAGS

BIG DATA
OAK RIDGE NATIONAL LABORATORY
SHANGHAI ASTRONOMICAL OBSERVATORY
SHAO
SKA-LOW
SQUARE KILOMETRE ARRAY
SUMMIT
SUPERCOMPUTING

October 22, 2019

Scientists have processed 400 gigabytes of data a second as they tested data pipelines for the Square Kilometre Array (SKA) telescope.

Researchers from ICRAR in Perth, Oak Ridge National Laboratory in the US and Shanghai Astronomical Observatory in China used the world's most powerful supercomputer—Summit—to process simulated observations of the early Universe ahead of the radio telescope being built in Western Australia and South Africa.

Summit — Oak Ridge National Laboratory's 200 petaflop supercomputer. Credit: Oak Ridge National Laboratory.

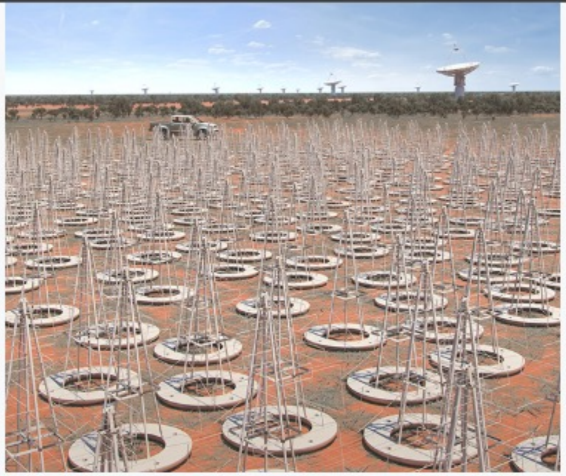
The data rate achieved was the equivalent of more than 1600 hours of standard definition YouTube videos every second.

Professor Andreas Wicenec, the director of Data Intensive Astronomy at the International Centre for Radio Astronomy Research (ICRAR), said it was the first time radio astronomy data has been processed on this scale.

"Until now, we had no idea if we could take an algorithm designed for processing data coming from today's radio telescopes and apply it to something a thousand times bigger," he said.

2020 Gordon Bell Nominee

https://www.icrar.org/summit/



Computer generated image of what the SKA-low antennas will look like in Western Australia. Credit: SKA Project Office.

The billion-dollar SKA is one of the world's largest science projects, with the low frequency part of the telescope set to have more than 130,000 antennas in the project's initial phase, generating around 550 gigabytes of data every second.

Summit is located at the US Department of Energy's Oak Ridge National Laboratory in Tennessee.

It is the world's most powerful scientific supercomputer, with a peak performance of 200,000 trillion calculations per second.

Oak Ridge National Laboratory software engineer and researcher Dr Ruonan Wang, a former ICRAR PhD student, said the huge volume of data used for the SKA test run meant the data had to be generated on the machine itself.

"We used a sophisticated software simulator written by scientists at the University of Oxford, and gave it a cosmological model and the array configuration of the telescope so it could generate data as it would come from the telescope observing the sky," he said.

"Usually this simulator runs on just a single computer, generating only a small fraction of what the SKA would produce.

"So we used another piece of software written by ICRAR, called the Data Activated Flow Graph Engine (DALiuGE), to distribute one of these simulators to each of the 27,648 graphics processing units that make up Summit.

"We also used the Adaptable IO System (ADIOS), developed at the Oak Ridge National Laboratory, to resolve a bottleneck caused by trying to process so much data at the same time."

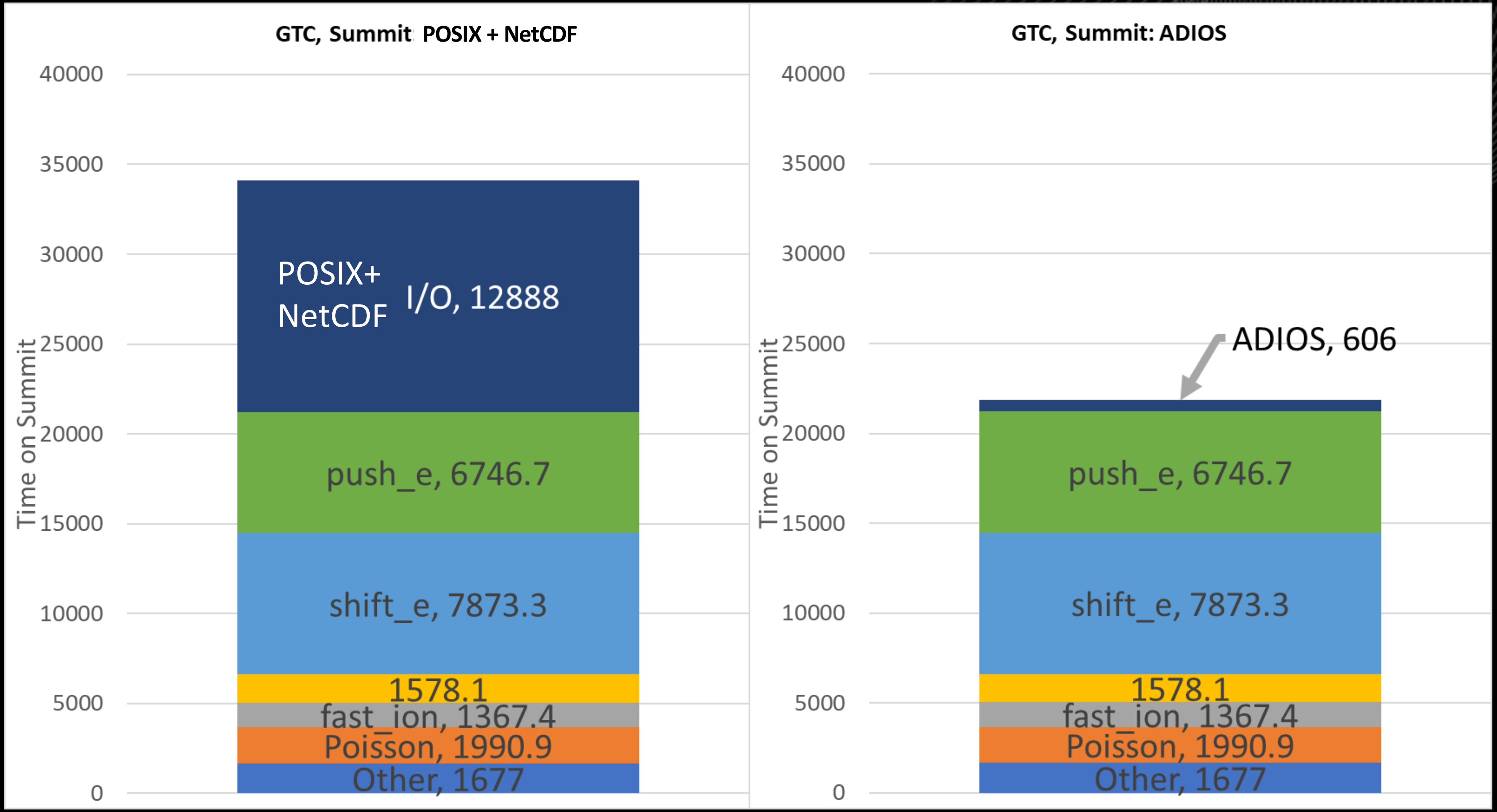
The test run used a cosmological simulation of the early Universe at a time

ANDREAS WICENEC, ICRAR, PI

GTC

PI: Zhihong Lin, UC Irvine

- Change to ADIOS I/O: Total simulation time reduced from 9.5 hours to 6.1 hours on 1024 nodes on Summit



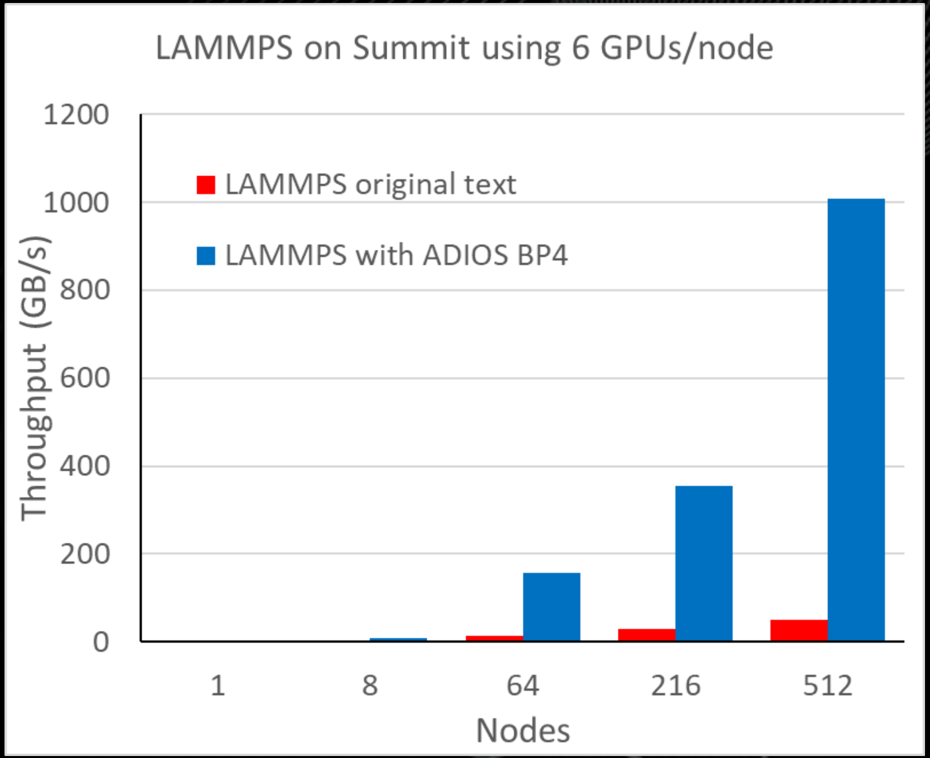
Results: LAMMPS

PI: Steve Plimpton, Sandia

https://github.com/lammps/lammps/tree/master/src/USER-ADIOS

- USER-ADIOS package in LAMMPS for dump commands
  - dump atom/adios
  - dump custom/adios
- Output goes into an I/O stream
  - BP4 file by default
  - Can use staging engines
- Concurrent reading is enabled

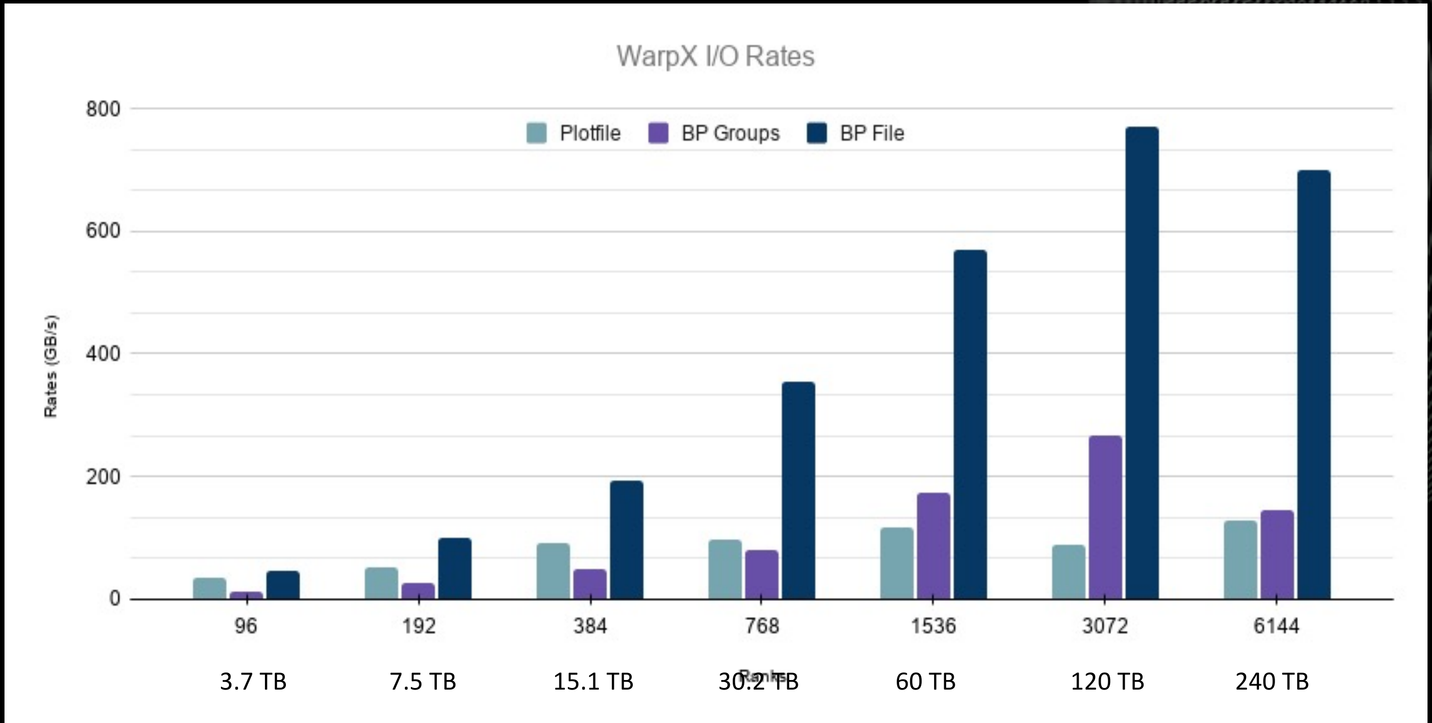
Results from ECP EXAALT Q4/FY19 milestone report (for 2.2.1.04 EXAALT ADSE04-54)
Summit 512 nodes
12B atoms, 5 TB



Results: WarpX

PI: Jan-Luc Vay, LBNL

- BPFile:
  - BP4: Use one file for all outputs.
- BPGroups:
  - BP3: Use one file / timestep
- Plot:
  - The AMReX plot file.
- One way to improve the I/O performance, is to use one ADIOS file for all time steps

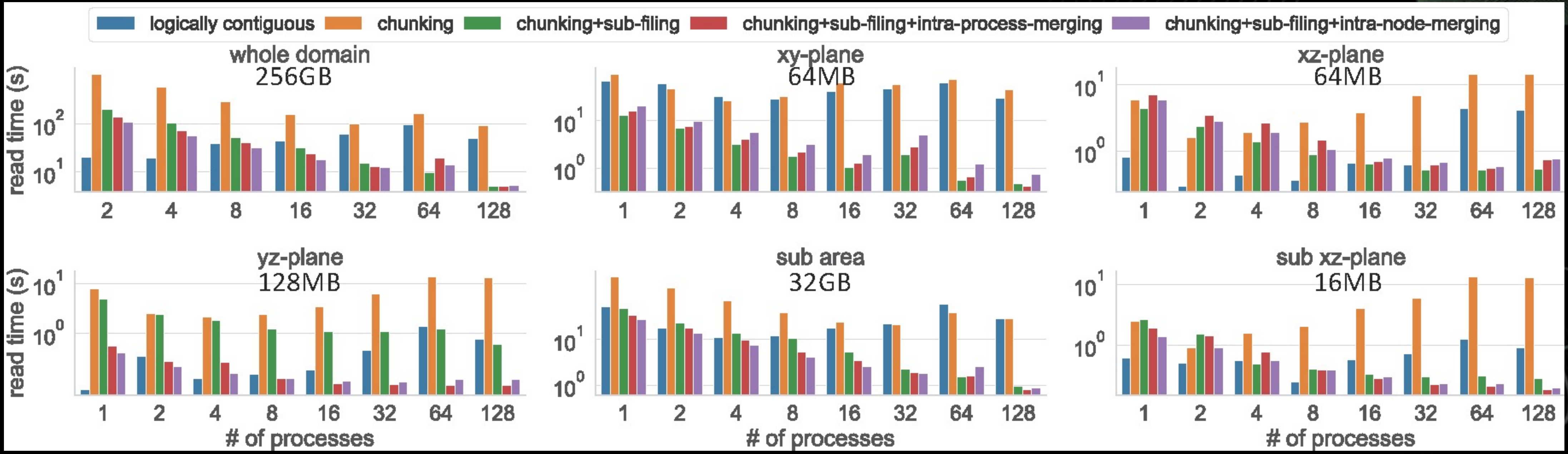


Summit, 6 GPUs, 6 cores per node, up to 1024 nodes

slide courtesy of Junmin Gu, LBNL

Writing performance is great but what about reading?

- Codes such as the WarpX code, which uses AMReX can take advantage of ADIOS-BP4 for “fast” writing
- The challenge is reading
- Development of a clustering algorithm for WarpX/AMReX data for fast writing/reading performance



Lipeng Wan et al.: "Data Layout Strategies for Parallel I/O: The Good, The Bad and The Ugly", submitted to TPDS journal 2021/March