Exascale IO and in situ Data Processing on Leadership Computers

Scott Klasky (ORNL) (klasky@ornl.gov), Matthew Wolf (ORNL)

OLCF meeting (May 22, 2019)

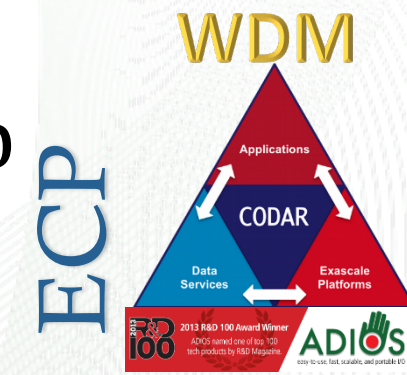
Scientific Data Management

Norbert Podhorszki –TL
Mark Ainsworth
Jong Choi
William Godoy
Tahsin Kurc
Qing Liu
Jeremy Logan
Kshitij Mehta
Eric Suchyta
Ruonan Wang
Lipeng Wan

Scientific Data Analytics

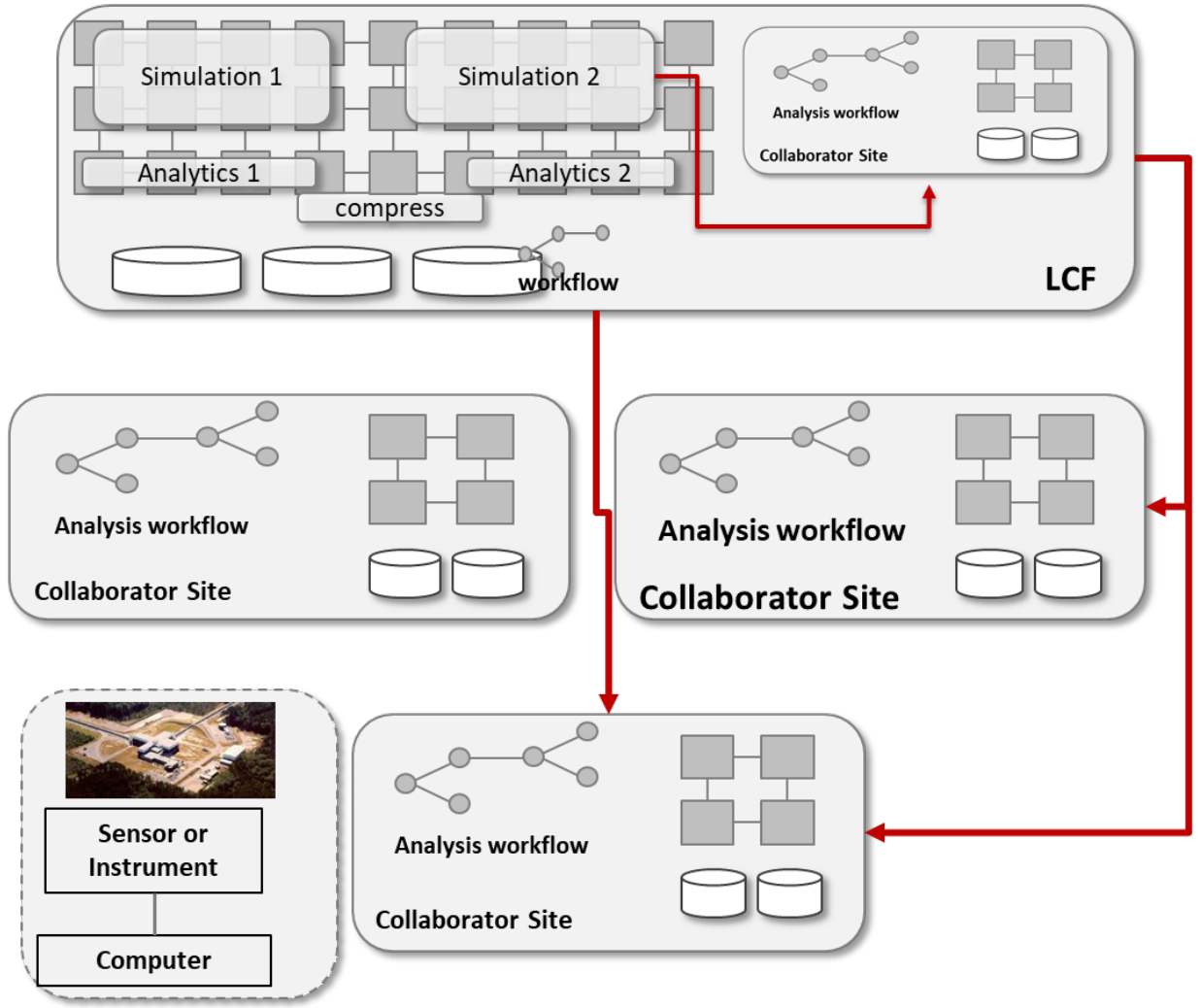
Dave Pugmire
Mark Kim
James Kress
George Ostrouchov
Jieyang Chen
Nick Thompson

https://github.com/CODARcode/MGARD
https://github.com/ornladios/ADIOS2
https://gitlab.kitware.com/vtk/vtk-m
https://pbdr.org/packages.html
https://Adios.ornl.gov



Vision: Enabling High Performance pub/sub I/O

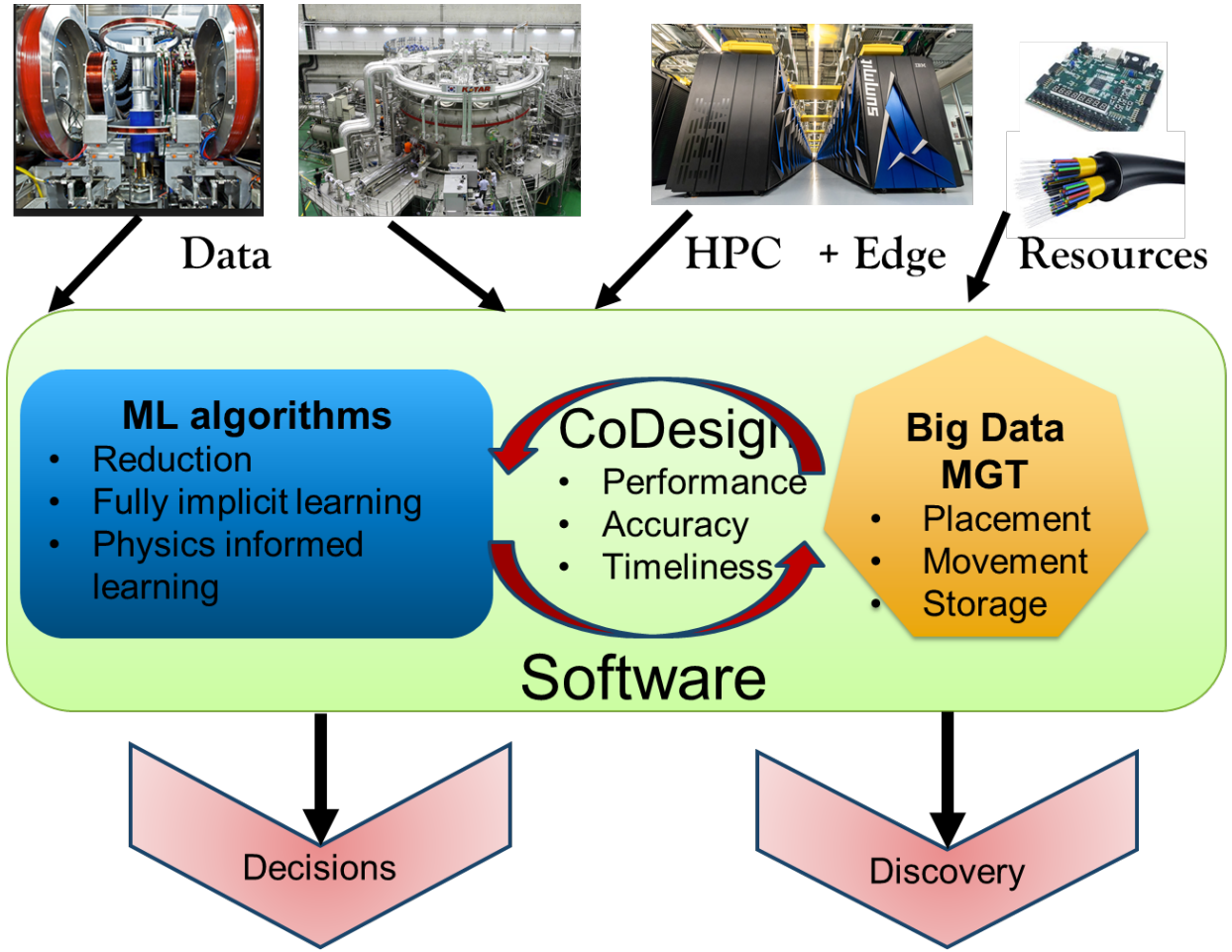
- Create a high performance I/O abstraction to allow for on-line memory/file data subscription service
- Create an abstraction for self-describing I/O for files, streams, etc.
- Work at all scales (laptops, desktops, clusters, exascale-platforms)
- Scale out to the number of nodes, sites, timesteps, variables, etc.



S. Klasky, M. Wolf, M. Ainsworth, C. Atkins, J. Choi, G. Eisenhauer, B. Geveci, W. Godoy, et al., A View from ORNL: Scientific Data Research Opportunities in the Big Data Age in 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), IEEE, pp. 1357–1368.

Co-design the next set of tools for federated computing

- The convergence of large DOE instruments with HPC centers dictates that we need to allow coupling/streaming
  - Codesign of what occurs at the edge and at HPC centers is imperative
  - Integration of ML/AI with HPC is essential to process as much data



Foster, Ian, Mark Ainsworth, Bryce Allen, Julie Bessac, Franck Cappello, Jong Youl Choi, Emil Constantinescu et al. "Computing just what you need: Online data analysis and reduction at extreme scales." In *European Conference on Parallel Processing*, pp. 3-19. Springer, Cham, 2017.

Software we need to enable our application use-cases

- Workflow system
  - To coupled" applications, submit the jobs on any system, and to monitor the job
- A Data management system
  - To get/put data to/from one or more consumers to on or more producers
  - To track the provenance and performance of the workflow
  - To write/read data to the storage layer along with pub/sub to multiple consumers
- Data compression/reduction
  - Which can take user Quantiles of Interest and preserve those within an error bound
- Analysis and Visualization services which can be incorporated into the workflow
  - Fast, Scalable, and memory efficient
- A Dashboard to allow scientist to collaborate during a running experiment/simulation
  - Web based, user configurable

Software we need to enable our application use-cases

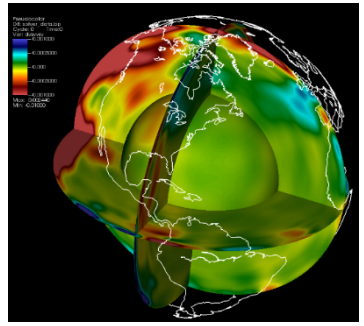
- Workflow system
  - EFFIS, or any other (e.g. EnTK, Kepler, Pegasus, Swift, ..)
- A Data management system
  - ADIOS, …
- Data compression/reduction
  - MGARD, SZ, ZFP, blosc, …
- Analysis and Visualization services which can be incorporated into the workflow
  - VTK-m, pbdR, Visit, Paraview, anaconda …
- A Dashboard to allow scientist to collaborate during a running experiment/simulation
  - Kdash, ..

The new demands from applications

Simulations – Storage: Seismic Tomography Workflow (PBs of data/run)

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

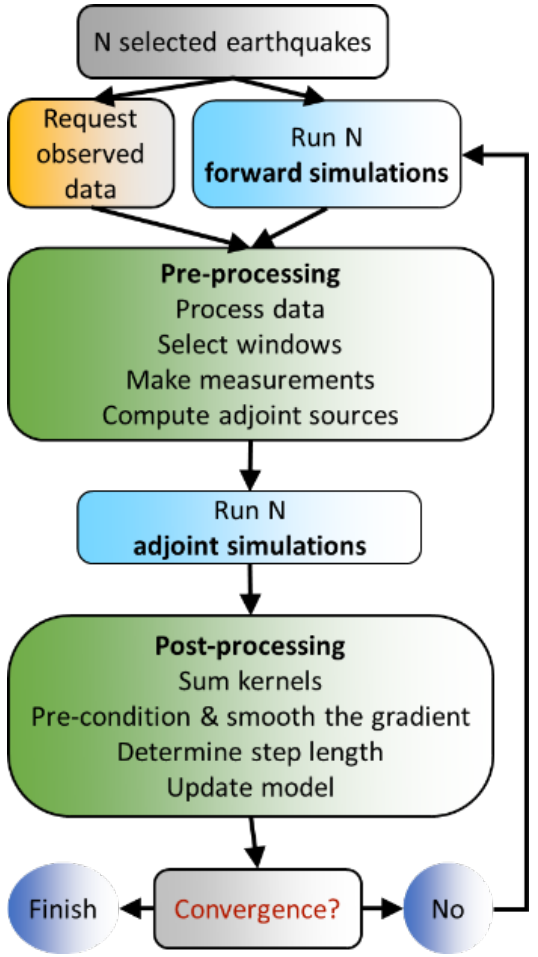


Significance and Impact

- First global seismic model where no approximations were used to simulate how seismic waves travel through the Earth.
- Over 1 PB of data was generated in a 6 hour simulation (on Titan@OLCF)

Research Details

- To improve data movement and flexibility, the Adaptable Seismic Data Format (ASDF) was developed that leverages the Adaptable I/O System (ADIOS) parallel library
- ASDF allows for recording, reproducing, and analyzing data on large-scale supercomputers
- **1.5 PB** of data is produced in a single workflow step, which is fully processed later in another step



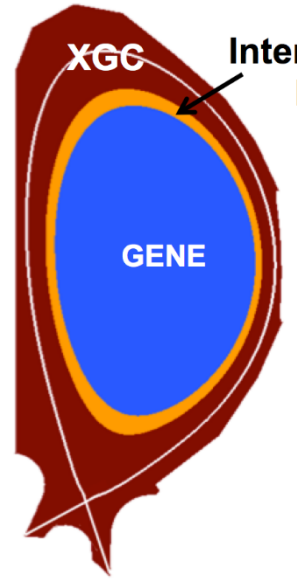
E. Bozdag; D. Peter; M. Lefebvre; D. Komatitsch; J. Tromp; J. Hill; N. Podhorszki; D. Pugmire., **Global adjoint tomography: first-generation model.** *Geophysical Journal International* 2016 207 (3): 1739-1766

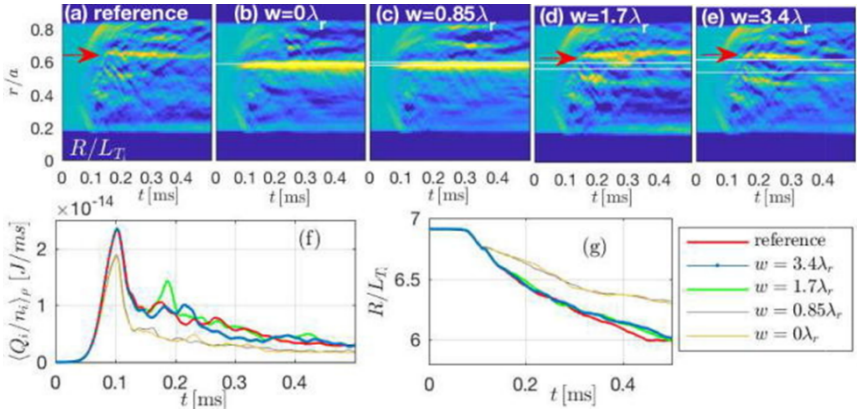
Simulations - Code coupling: ECP WDM - High-Fidelity Whole Device Modeling of Magnetically Confined Fusion Plasmas

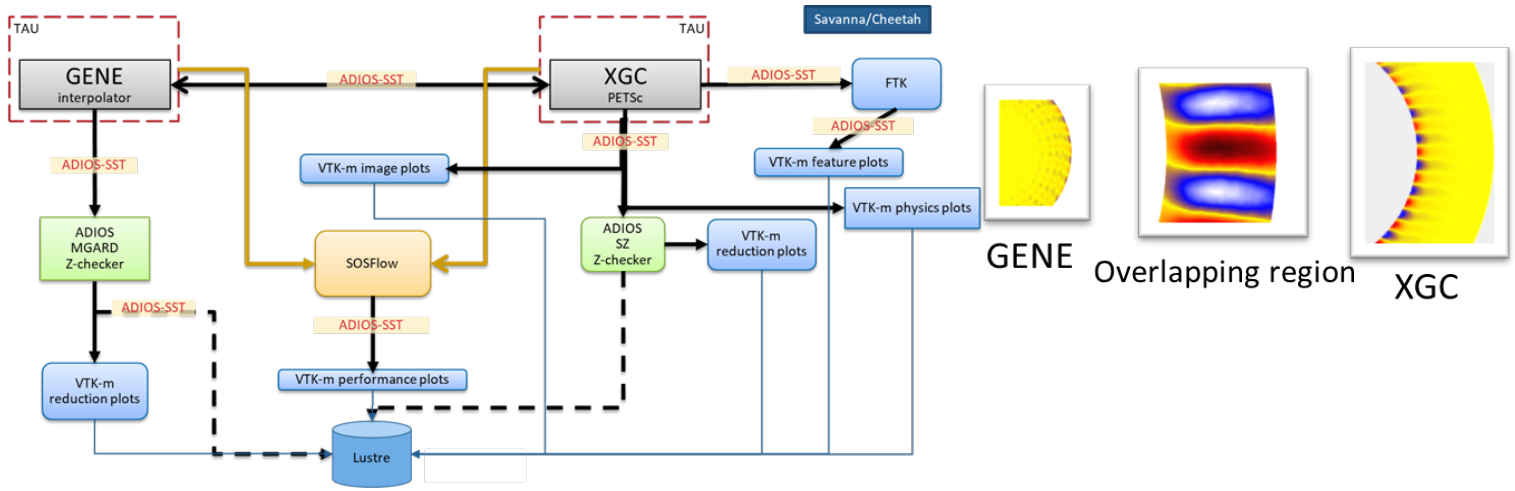
PI: A. Bhattacharjee, PPPL,
C. S. Chang, PPPL

- 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

Core-edge coupling

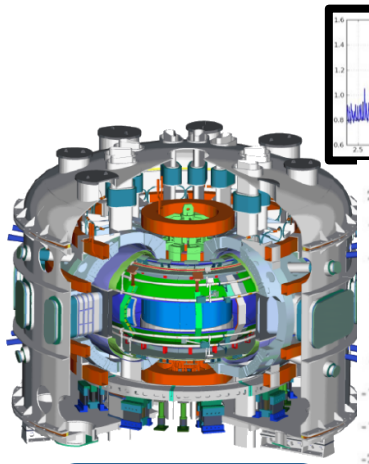


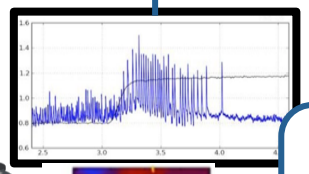


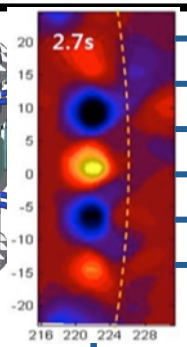


J. Dominski; S. Ku; C.-S. Chang; J. Choi; E. Suchyta; S. Parker; S. Klasky; A. Bhattacharjee; *Physics of Plasmas* **2018,** 25, DOI: 10.1063/1.5044707

Experimental data streaming from KSTAR to OLCF/NERSC







KSTAR Data Stream

Machine learning based data refactoring, data reduction, and prioritization

EFIT magnetic reconstructions (0D magnetic)

Correlation analysis: $|\widetilde{B}_\theta|$ and β_θ

Parameter based, reduced model comparison (1D ECEi)

ECEi Synthetic Diagnostic

Parameter based, reduced model comparison (2D ECEi)

ECEi FFT/Wavelet Decomposition

Machine learning based mode classifier

Campaign Storage

Data access patterns analysis and prediction

MCMC based integration data fusion (2D ECEi, soft-xrays, etc.)

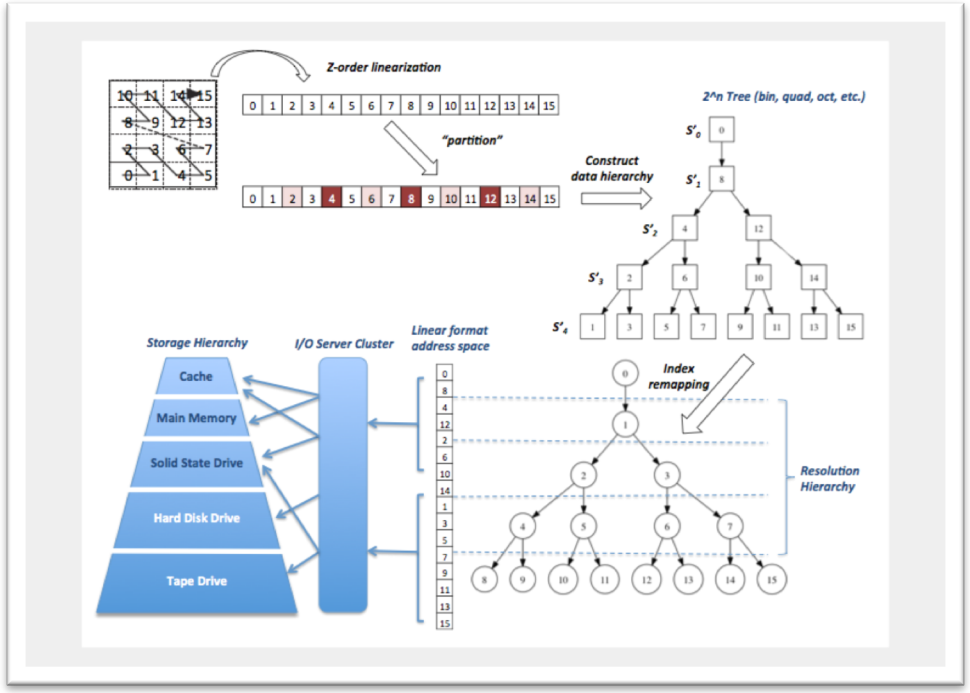
Fusion WDM ECP App (XGC, GENE, M3D-C1, RF, EP)

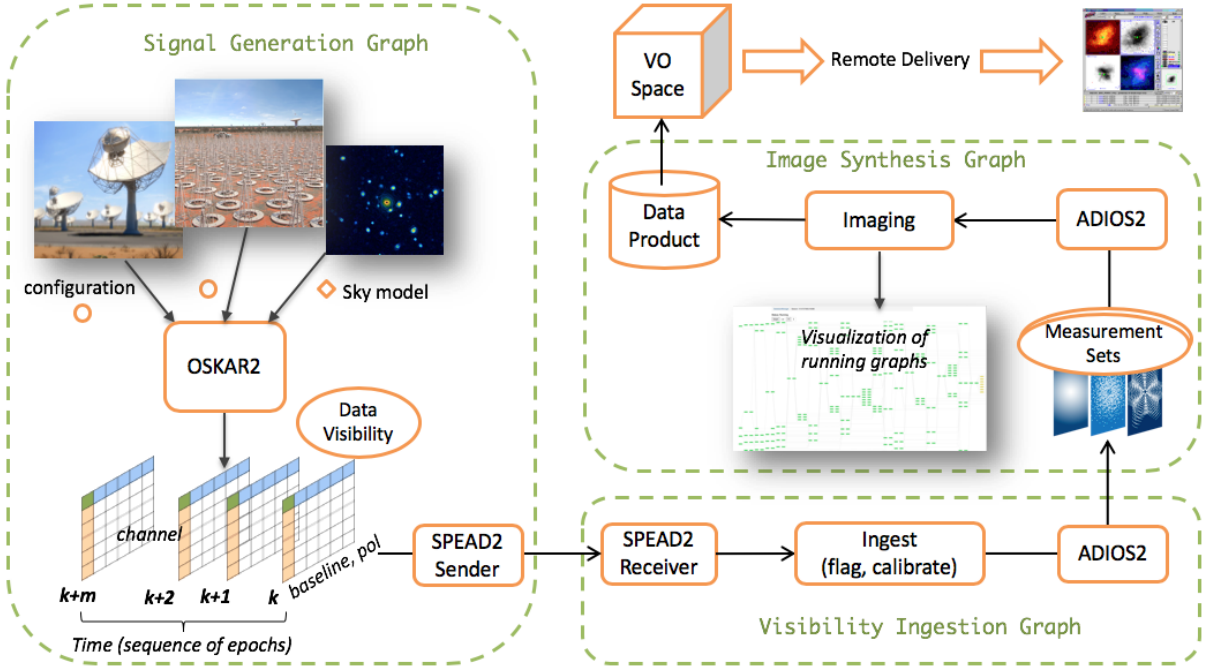
Choi, Jong Y., et al. "Icee: Wide-area in transit data processing framework for near real-time scientific applications." *4th SC Workshop on Petascale (Big) Data Analytics: Challenges and Opportunities in conjunction with SC13*. Vol. 11. 2013.

Observational data: SKA

- The full SKA will use a million antennas to enable astronomers to monitor the sky in unprecedented detail and survey the entire sky much faster than any system currently in existence
- One of SKA's greatest challenges is in the ability to move, process, and store data, without losing information
- The data output of the SKA is limited by the achievable I/O bandwidth

Telescope	Raw Data Rate	Archive Growth
MWA	1.4 TB/hour	5 PB/year
LSST	1.5 TB/hour	6 PB/year
ASKAP	9 TB/hour	5.5 PB/year
SKA1-LOW	1,400 TB/hour	150 PB/year

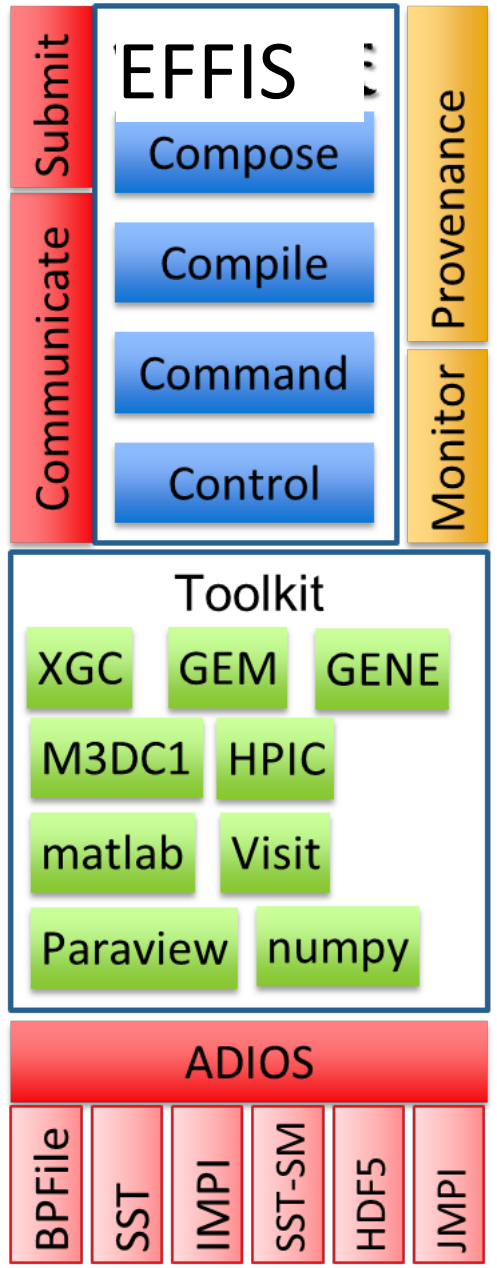




Wang, Ruonan, Christopher Harris, and Andreas Wicenec. "AdiosStMan: parallelizing casacore table data system using adaptive IO system." *Astronomy and computing* 16 (2016): 146-154.

EFFIS 2.0: an End-to-end Framework For coupling Integrated Simulations

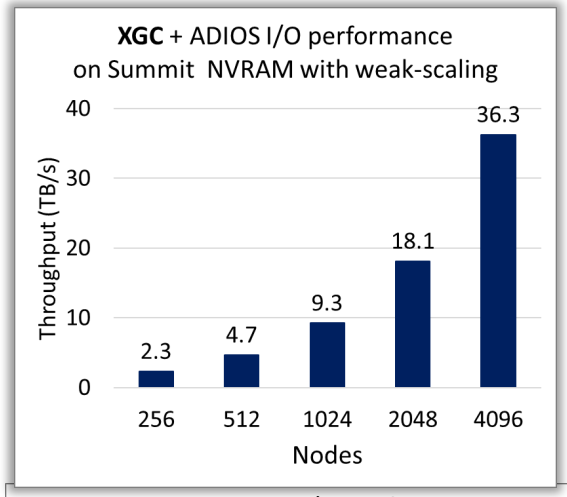
- Why EFFIS?
  - To simplify the complexity of composing, running, and monitoring applications on HPC systems
- What is EFFIS
  - A collection of services to compose, launch, monitor, and control coupled applications
- How does EFFIS work
  - Contains new services to “easily” compose coupled HPC applications on HPC Resources using a python-like interface
    - Cori, Theta, Titan, Summit
  - Allows “easy” integration to visualization tools (Visit, Python notebooks, etc.)
  - ADIOS for data movement

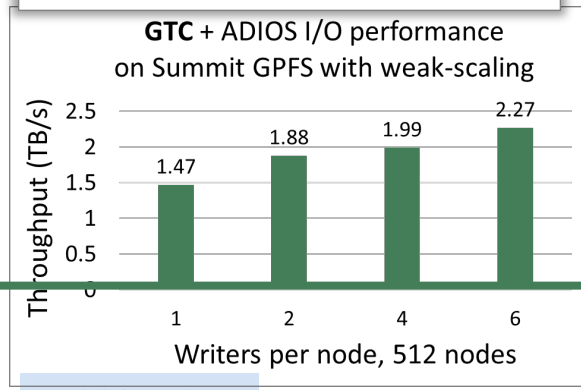


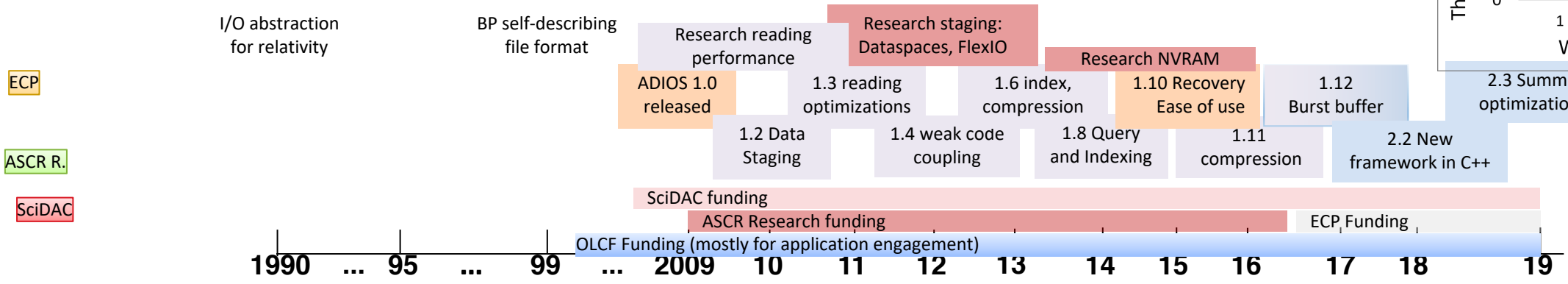
Cummings, J., A. Pankin, N. Podhosrzki, G. Park, S. Ku, R. Barreto, S. Klasky et al. "Plasma edge kinetic-MHD modeling in tokamaks using Kepler workflow for code coupling, data management and visualization." *Communications in Computational Physics* 4, no. 3 (2008): 675-702.

ADIOS I/O Framework for Data Intensive Science

- Problem
  - I/O is severely bottlenecked on HPC systems and experiments because of hardware limitations
  - New I/O patterns from code coupling, AI/ML, simulations require new solutions
- Solution
  - ADIOS is a DOE framework developed for sustainable I/O on Leadership Class Facilities (LCF)
  - ADIOS is an publication/subscribe I/O framework for storage and in situ processing of data
- Impact
  - Over 10X I/O improvement from previous I/O methods on dozens of LCF apps
  - Used by more than 30 LCF application areas, totaling over 1B hours on the LCFs,
  - Outside DOE: Used in Industrial Engineering, Oil Exploration, Computational Fluid Dynamics
  - ADIOS won R&D 100 Award in 2013







ADIOS Approach: "How"

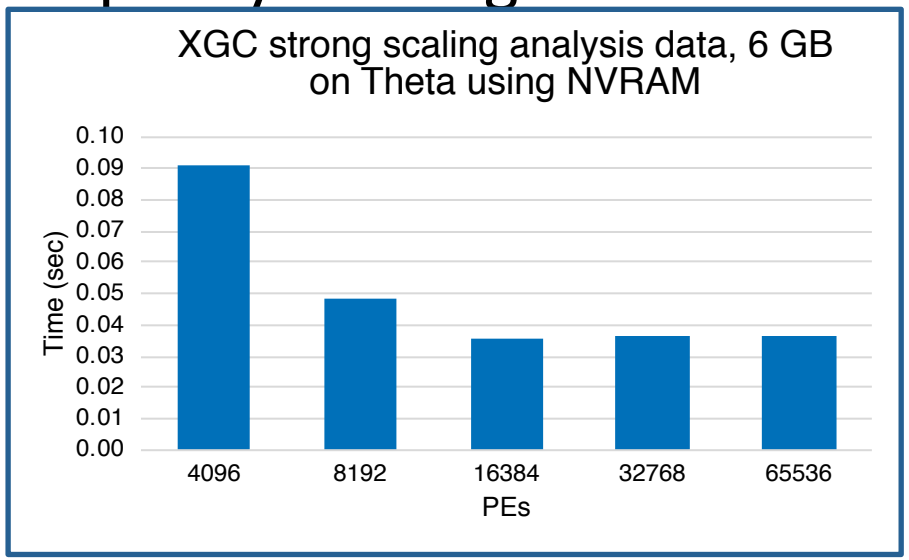
- I/O calls are of declarative nature in ADIOS
  - which process writes what: add a local array into a global space (virtually)
  - adios_close() indicates that the user is done declaring all pieces that go into the particular dataset in that timestep
- I/O strategy is separated from the user code
  - 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

Q. Liu, J. Logan, Y. Tian, H. Abbasi, N. Podhorszki, J. Y. Choi, S. Klasky, R. Tchoua, J. Lofstead, R. Oldfield, et al. Hello ADIOS: the challenges and lessons of developing leadership class I/O frameworks. *Concurrency and Computation: Practice and Experience* 2014, *26*, 1453–1473.

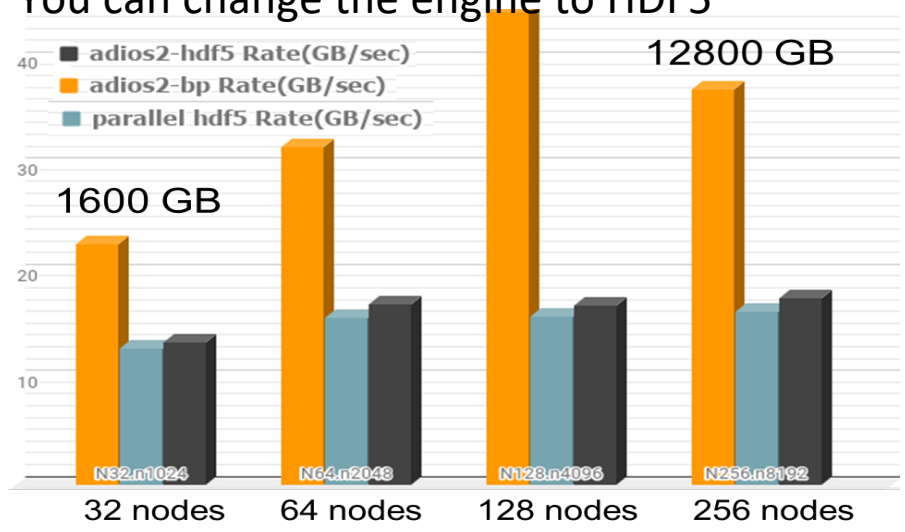
Changing from C/R data to visualization/analysis data is "easy"

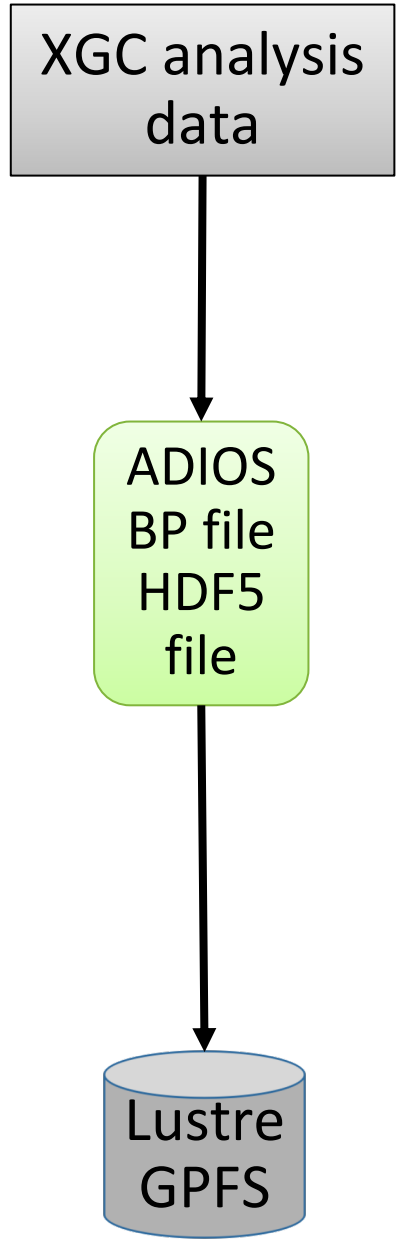
- One change in the code or input file, to specify the engine

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

writer.BeginStep()

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

writer.EndStep()

writer.Close()
```



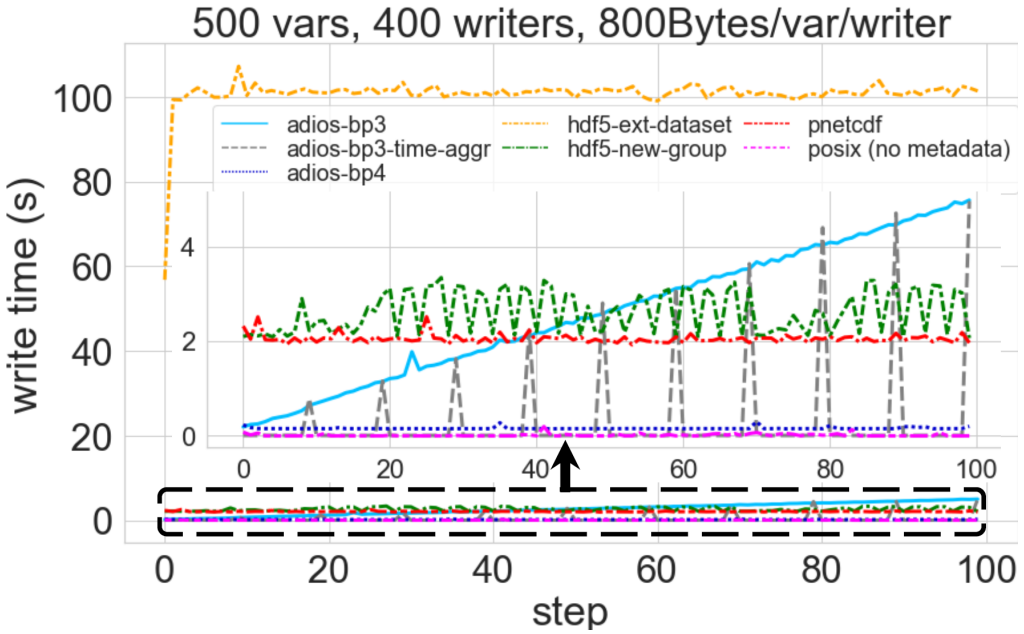
You can change the engine to HDF5





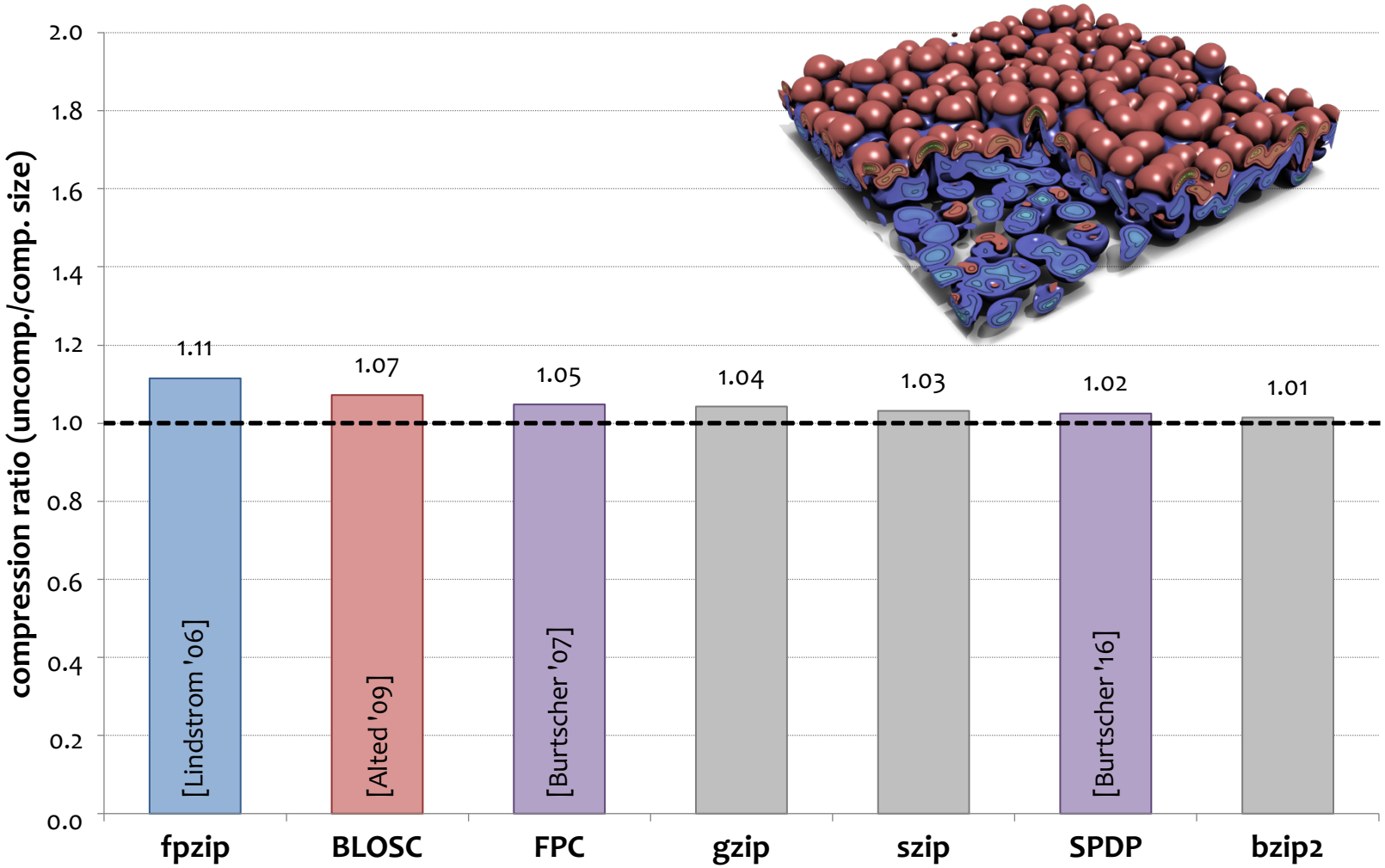
Metadata challenges in big scientific data management

- Creating “containers” to store all of the data products, code, workflow, performance, “raw data” allows scientist to better understand their campaign
- As scientist continue to save data products from their simulations/experiments, the amount of variables grow
- The cost of managing the metadata from all the data produced in the scientific process can grow to be “painful” for scientist
- There is a large overhead of generating and managing the metadata grows with number of data objects (variables, attributes), number of processes, number of simulation steps



L. Wan, K. Mehta, S. Klasky, M. Wolf, H. Wang, W. Wang, J. Li, Z. Lin, Data Management Challenges of Exascale Scientific A case study with the Gyrokinetic Toroidal Code and ADIOS, ICCM 2019.

Numerical data is challenging to compress losslessly (P. Lindstrom)



Lindstrom, Peter. "Fixed-rate compressed floating-point arrays." *IEEE transactions on visualization and computer graphics* 20.12 (2014): 2674-2683.

MGARD: MultiGrid Adaptive Reduction of Data

Decomposes data into contributions from a hierarchy of meshes,

$$u = Q_0 u + (Q_1 - Q_0)u + \ldots + (Q_L - Q_{L-1})u$$

Adaptive reduction of data based on discarding least important contributions

$$u = \sum_{\ell=0}^{L} \sum_{n \in N_\ell} \alpha_n^l \varphi_n^\ell \qquad \tilde{\alpha}_n^\ell = \begin{cases} \alpha_n^\ell, & \text{for } |\alpha_n^\ell| \geqslant \tau_c \\ 0, & \text{otherwise} \end{cases}$$

$$\tilde{u} = \sum_{\ell=0}^{L} \sum_{n \in N_\ell} \tilde{\alpha}_n^\ell \varphi_n^l$$

Mathematically proven error bounds

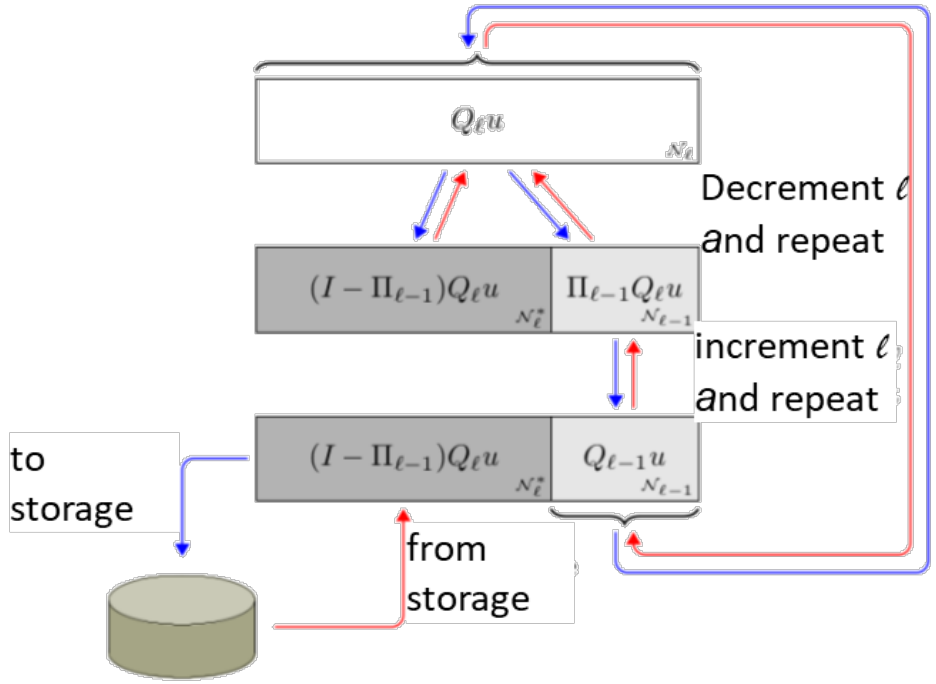
$$C_d \eta_s(v) \leq \|v\|_s \leq \eta_s(v)$$

where

$$\eta_s(v)^2 = \sum_{\ell=0}^{L} 2^{2s\ell} \|\Delta_\ell v\|^2$$

Applicable to structured (tensor product) grids with arbitrary spacing, integrated into ADIOS

Able to preserve quantities of interest (spectrum, averages…)



MGARD decomposition and recomposition workflow

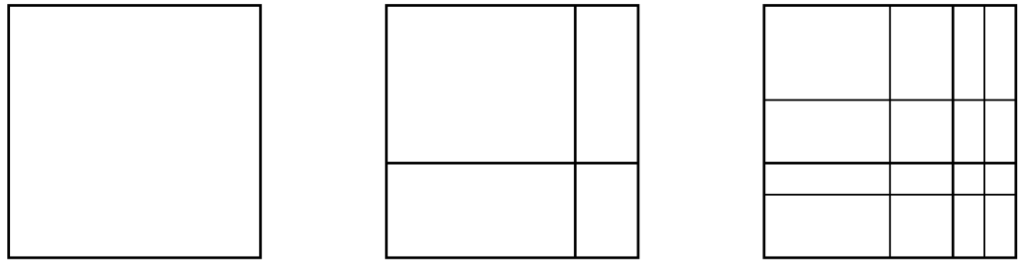


Illustration of a tensor product mesh hierarchy

M. Ainsworth, S. Klasky, B. Whitney. Compression using lossless decimation: analysis and application. SIAM Journal on Scientific Computing 2017, 39, B732–B757.

M. Ainsworth, O. Tugluk, B. Whitney, S. Klasky, "Multilevel Techniques for Compression and Reduction of Scientific Data -- The Multivariate Case", *SIAM Journal on Scientific Computing,* Submitted for publication 2018.

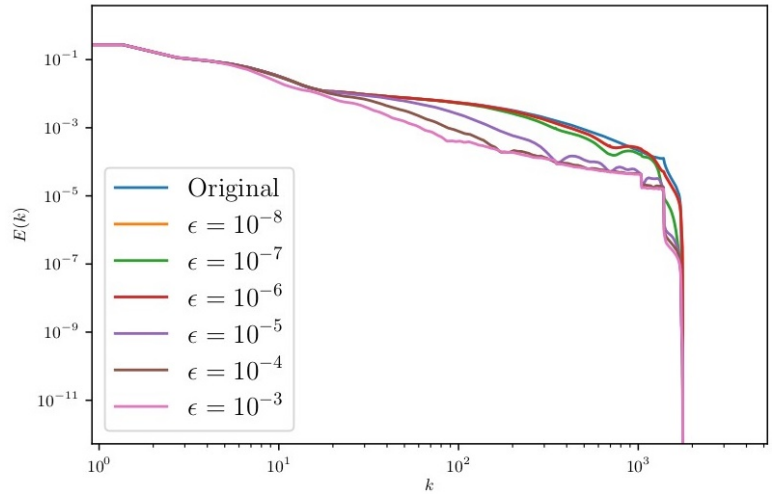
M. Ainsworth, O. Tugluk, B. Whitney, S. Klasky. Multilevel techniques for compression and reduction of scientific data-Quantitative control of accuracy in derived quantities. SIAM Journal on Scientific Computing 2019, TBD, TBD.

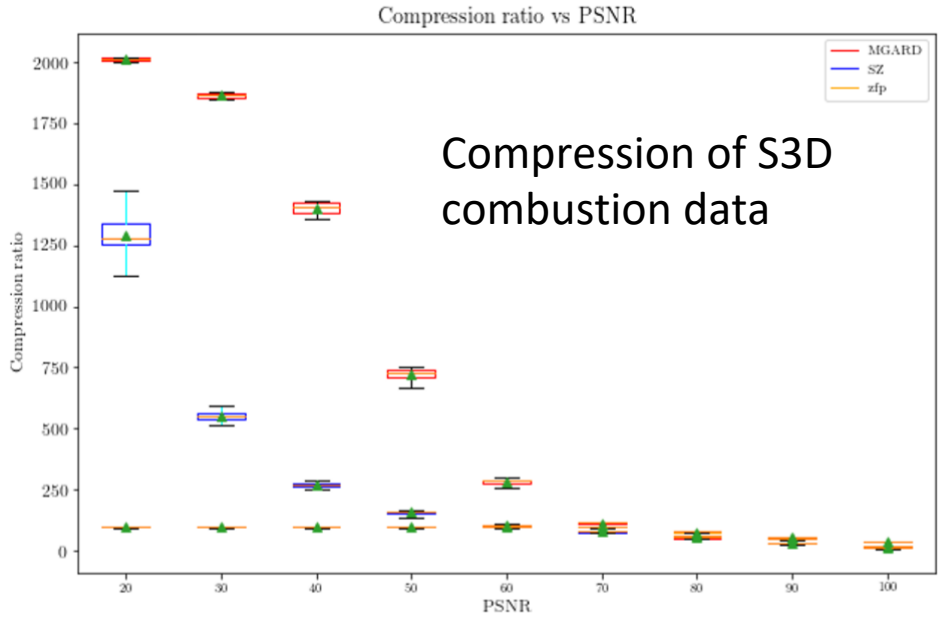
MGARD: MultiGrid Adaptive Reduction of Data

- MGARD can preserve quantities of interest specified by users $|Q(u) - Q(\tilde{u})| = |Q(u - \tilde{u})| \leq r_{L\infty}(Q)\|u - \tilde{u}\|_{L\infty}$
- This can be done by specifying a smoothness parameter (e.g. s=-1/2 for averages/blobs, -1 for streamlines),… $|Q(u) - Q(\tilde{u})| \leq r_s\|u - \tilde{u}\|_s$

- E.g. a user can supply a routine to compute a functional Maximum resolved wave number in energy spectra can be supplied as $|\vec{k}| \leq (\tau/\epsilon)^{-1/s}$

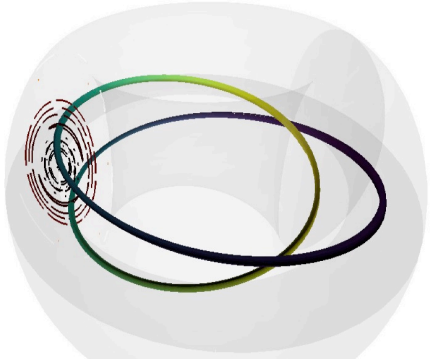
Effects of MGARD compression on turbulent energy spectra.





Puncture plot

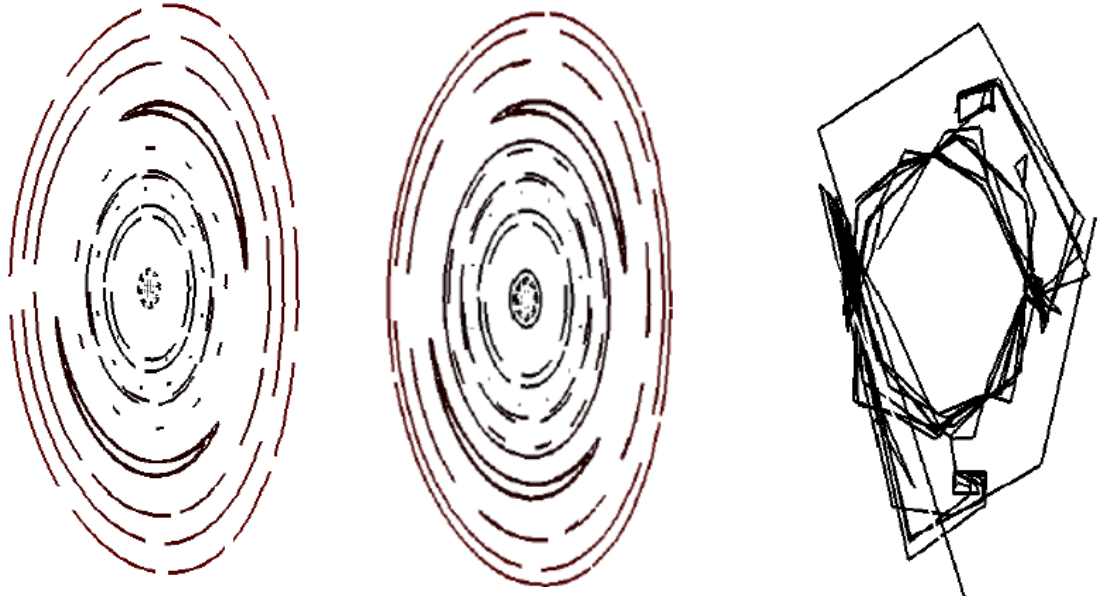
- Can we reduce data from a fusion code of the magnetic field vector (65x257x161) and still compute accurate puncture plots?
- Trajectories depend on
- $Q_k(v) = \int_{t_k}^{t_{k+1}} v(\left(\tilde{x}(t)\right)dt$
- Continuous linear functional
- $|Q(u-\tilde{u})| \leq r_s(Q)\|u-\tilde{u}\|_s$
- For s=(d-1)/2 in d-dimensions
- Here d=3, so take s=1
- Hence, we apply reduction where control loss $\|u-\tilde{u}\|_s$



Trajectories depend on quantity

$$\mathcal{Q}_k(v) = \int_{t_k}^{t_{k+1}} v(\tilde{x}(t))\,\mathrm{d}t$$

No Compression MGARD 100X (SZ, ZFP 100X)



MARK AINSWORTH, OZAN TUGLUK, BEN WHITNEY, AND SCOTT KLASKY, MULTILEVEL TECHNIQUES FOR COMPRESSION AND REDUCTION OF SCIENTIFIC DATA—QUANTITATIVE CONTROL OF ACCURACY IN DERIVED QUANTITIE, accepted to SIAM J. Scientific

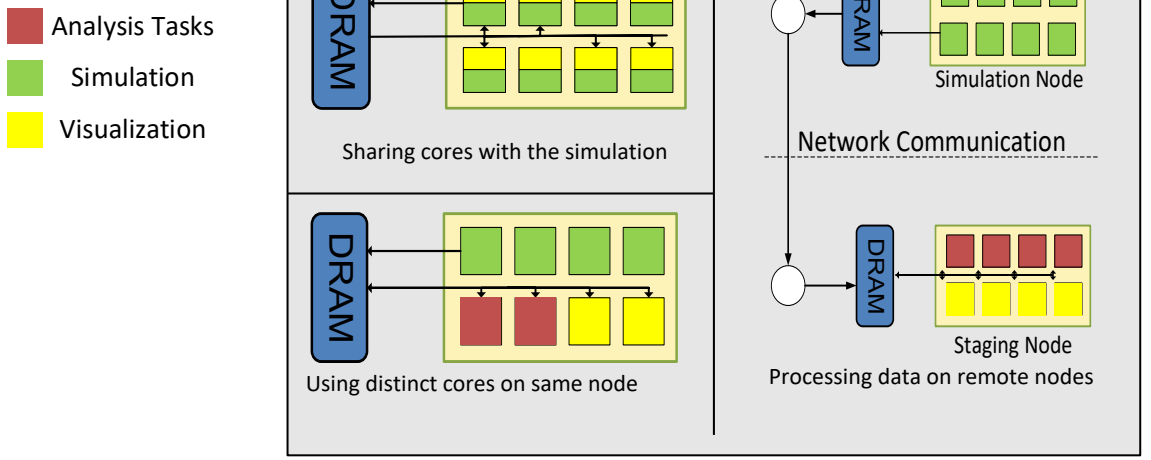
Why In situ

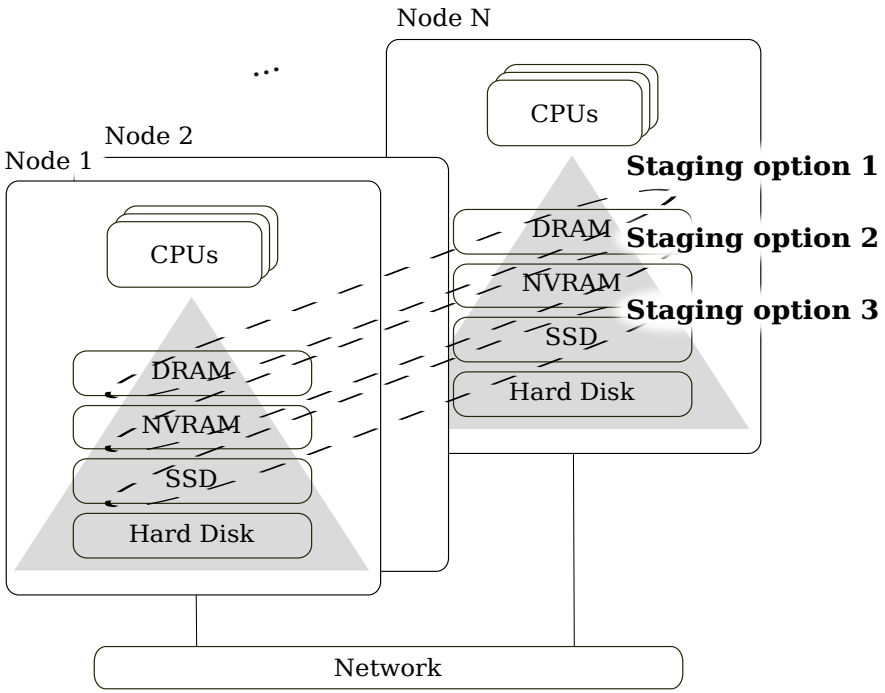
- Timeliness (Resources)
- Reduction… save information, not all the data

F. Zheng, H. Abbasi, C. Docan, J. Lofstead, Q. Liu, S. Klasky, M. Parashar, N. Podhorszki,
K. Schwan, M. Wolf, *PreDatA–preparatory data analytics on peta-scale machines* in *Parallel & Distributed Processing (IPDPS), 2010 IEEE International Symposium on*, IEEE, pp. 1–12.

Rich Design Space for Staging

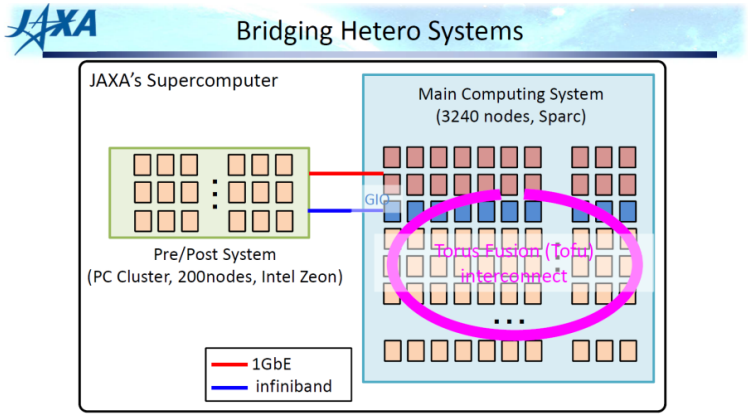
- Location of the compute resources
  - Same cores as the simulation (in situ)
  - Some (dedicated) cores on the same nodes
  - Some dedicated nodes on the same machine
  - Dedicated nodes on an external resource
- Data access, placement, and persistence
  - Direct access to simulation data structures
  - Shared memory access via hand-off / copy
  - Shared memory access via non-volatile near node storage (NVRAM, BB)
  - Data transfer to dedicated nodes or external resources (decoupled in space)
- Synchronization and scheduling
  - Execute synchronously with simulation every n^{th} simulation time step
  - Execute asynchronously (decoupled in time)
  - Dynamic execution





C. Docan, M. Parashar, S. Klasky. Dataspaces: an interaction and coordination framework for coupled simulation workflows. *Cluster Computing* **2012**, *15*, 163–181.

Data Staging in ADIOS



- Sustainable Staging Transport (SST)
  - In situ infrastructure for staging in a streaming-like fashi using RDMA, SOCKETS with "active" connect/disconnect
- InSituMPI
  - One way staging for MPMD applications, for strong coupling
- DataMan
  - WAN transfers using sockets and ZeroMQ for EO data
- SST-SM
  - Staging infrastructure providing a shared memory abstraction for coupling on shared nodes
- InSitu-sync
  - Synchronous in situ, direct pass through of data structures to analytics

Use a combination of ADIOS Staging methods

F. Zheng, H. Abbasi, J. Cao, J. Dayal, K. Schwan, M. Wolf, S. Klasky, N. Podhorszki, *In-situ I/O processing: a case for location flexibility* in *Proceedings of the sixth workshop on Parallel Data Storage*, ACM, pp. 37–42.

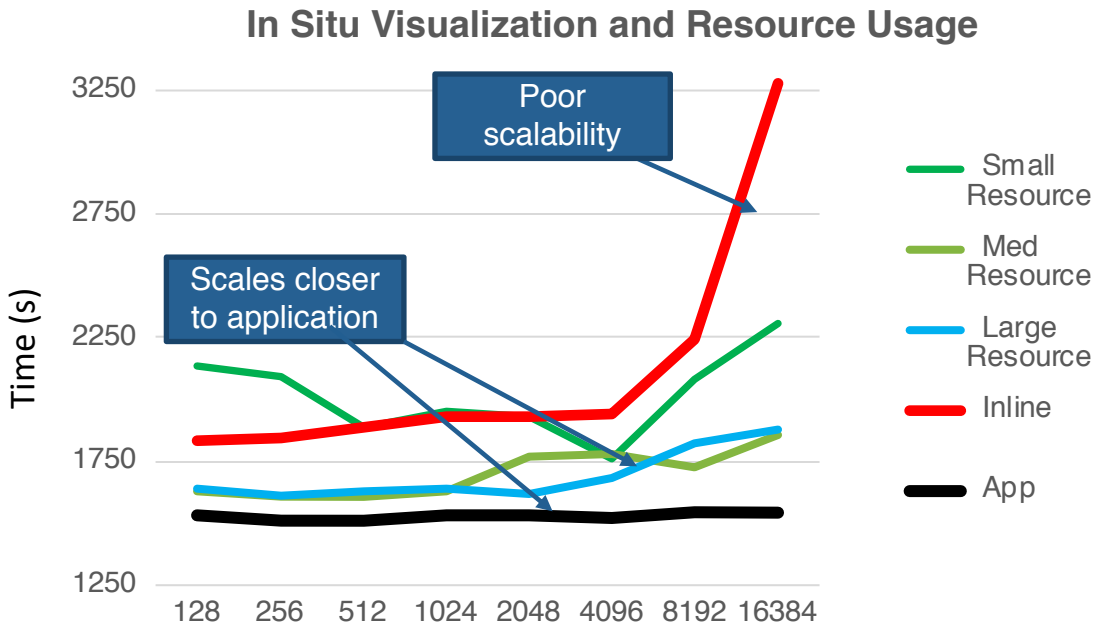
In-line vs. In-transit Visualization Techniques at Scale

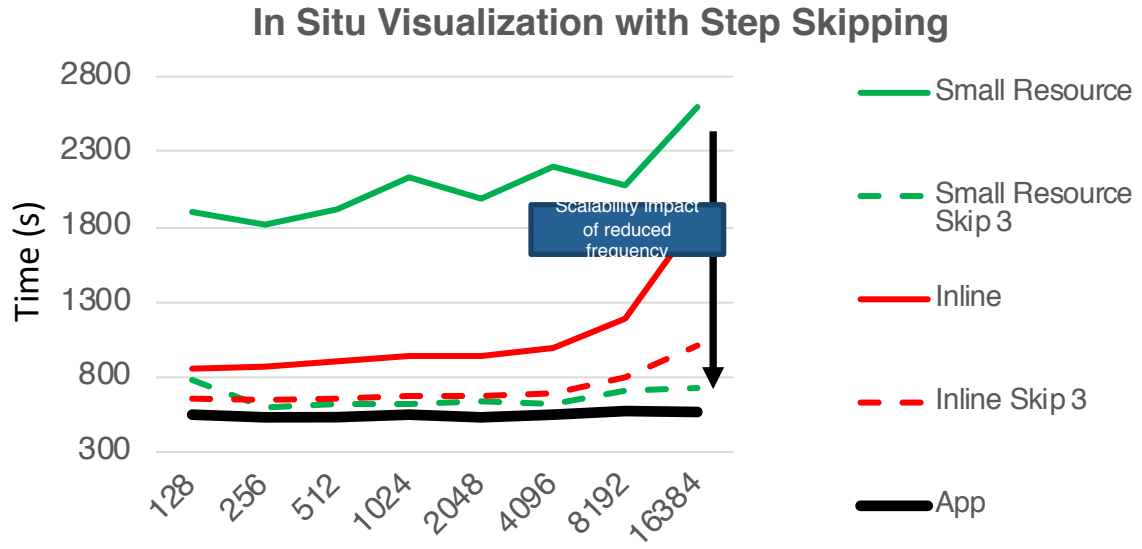
- Scientific workflows are complex
  - Numerous analysis and visualization tasks need to be coordinated
  - Inefficiency = less scientific insight, longer simulation times, larger cost
- Choosing the in situ strategy can be complex
  - Overheads of in situ visualizations are not well understood
- Study some examples from a Cloverleaf3D MiniApp
  - What configuration will impact the simulation time the least?
  - What configuration will be most cost effective, at scale?
- Study two visualization algorithms: Isocontour, and rendering

J. Kress, M. Larsen, J. Choi, M. Kim, M. Wolf, N. Podhorszki, S. Klasky, H. Childs, D. Pugmire, Comparing the Efficiency of In Situ Visualization Paradigms at Scale, ISC 19, Frankfurt Germany, June 2019 (In submission)

Placement for Visualization Services

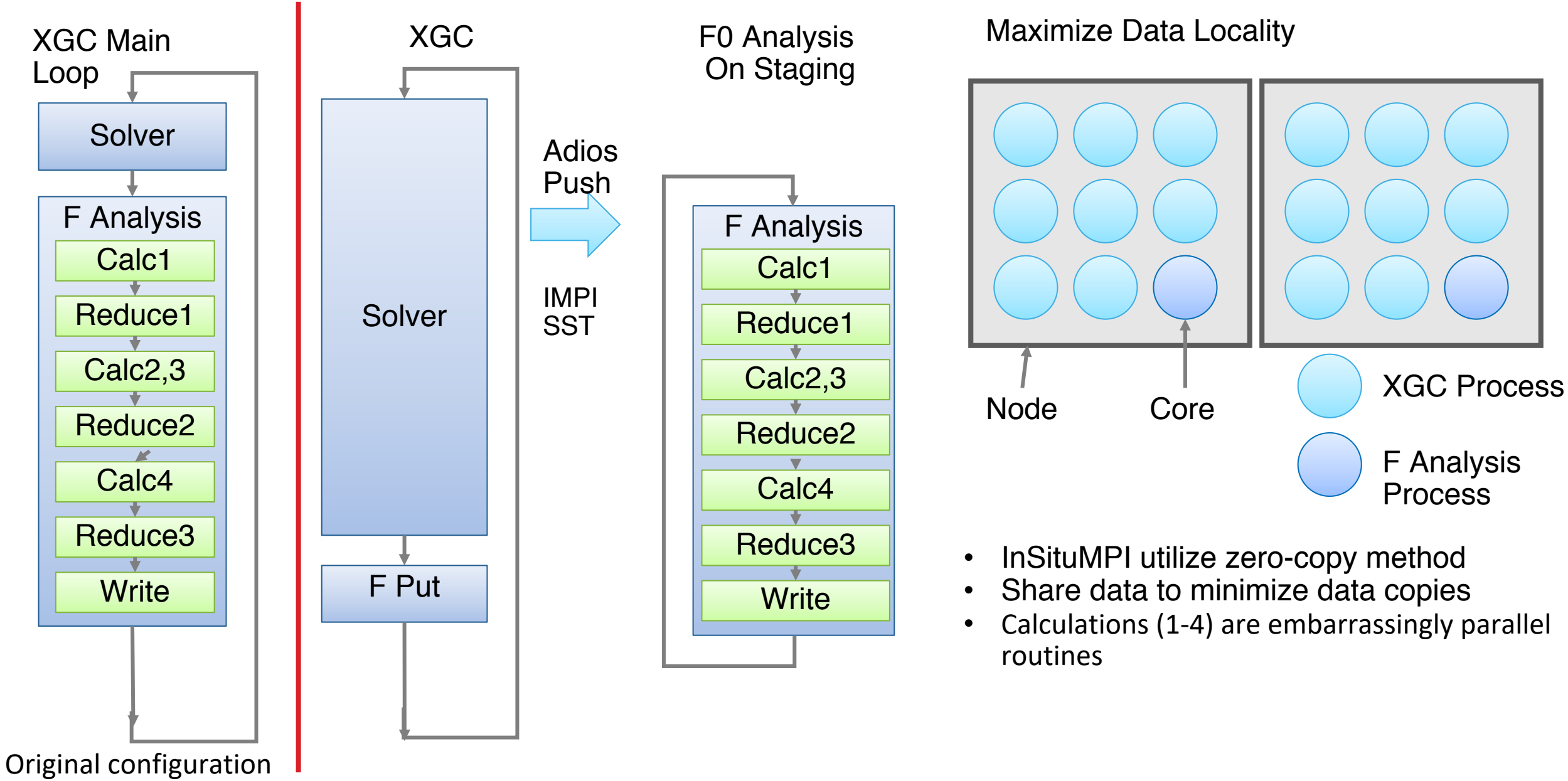
- Placement of services has a dramatic impact on performance
- Algorithms with communication (e.g. parallel rendering) exhibit poor performance at scale
- Move data to staging resource improves scalability of visualization services
- Visualizing every “Nth” step provides improves scalability of smaller resources



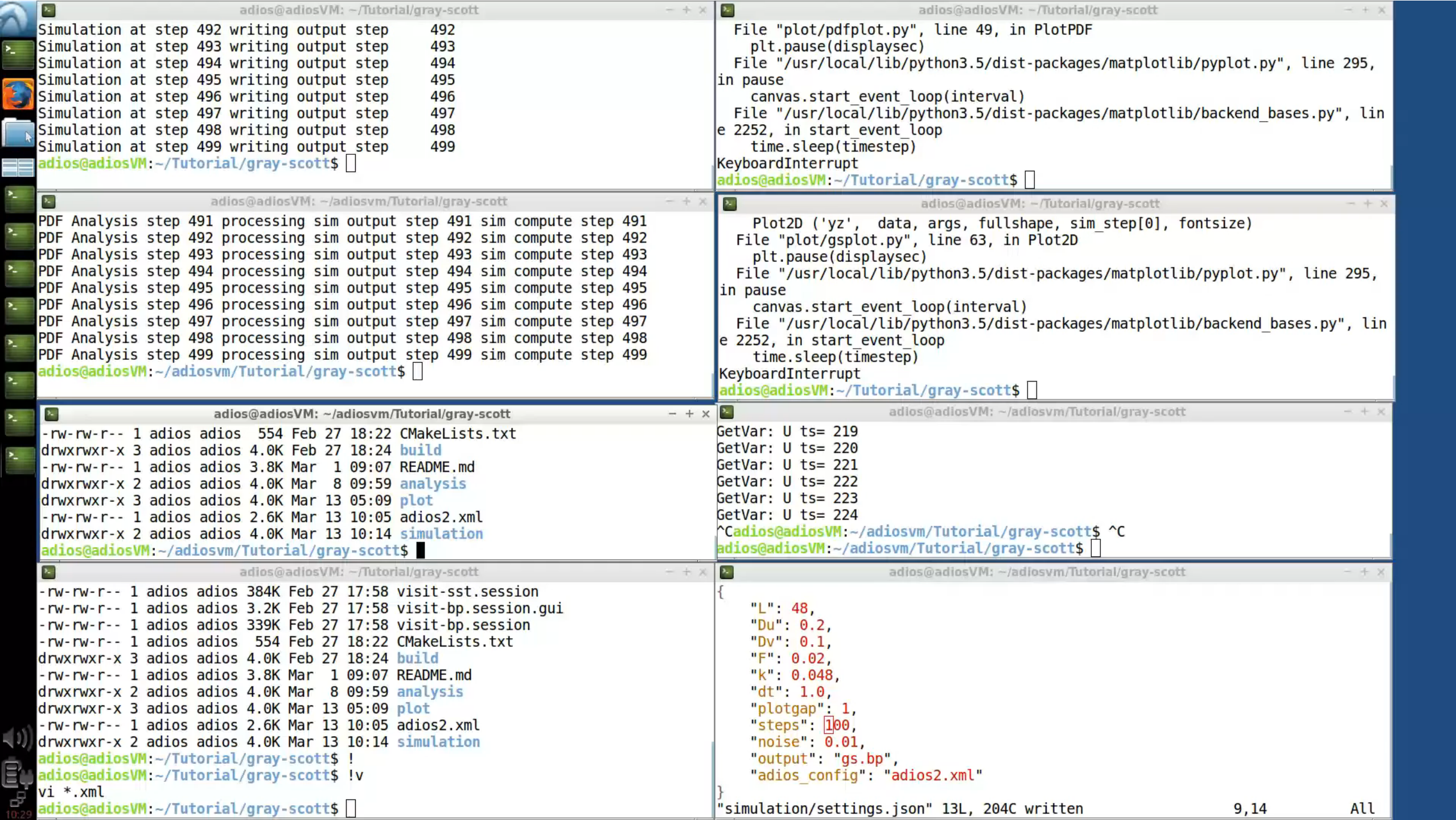


See: Comparing the Efficiency of In Situ Visualization Paradigms at Scale, to appear, ISC 2019, Frankfurt Germany. J Kress, M Larsen, J Choi, M Kim, M Wolf, N Podhorszki, S Klasky, H Childs, D Pugmire.

XGC Analysis – moving from synchronous to staged in situ



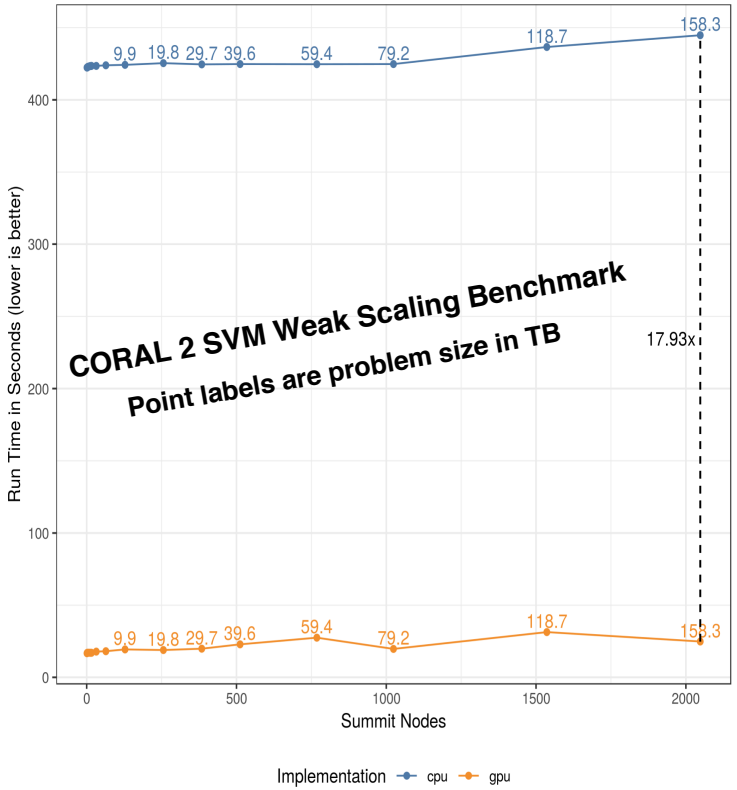
- InSituMPI utilize zero-copy method
- Share data to minimize data copies
- Calculations (1-4) are embarrassingly parallel routines

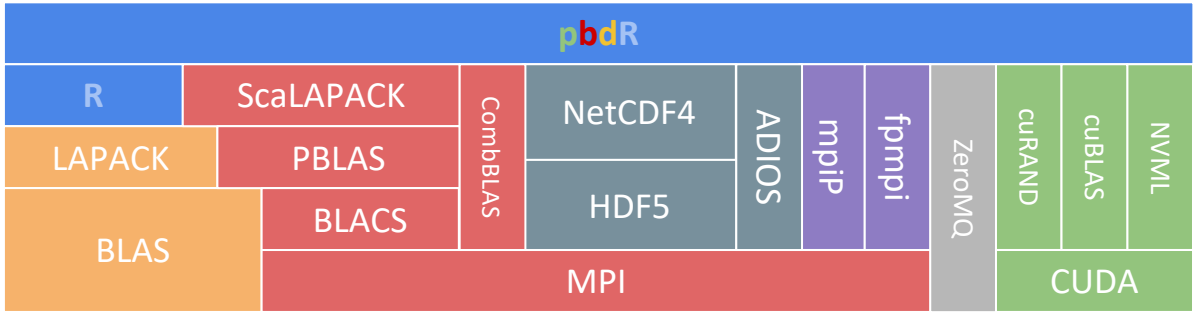

adios@adiosVM: ~/Tutorial/gray-scott
Simulation at step 492 writing output step 492
Simulation at step 493 writing output step 493
Simulation at step 494 writing output step 494
Simulation at step 495 writing output step 495
Simulation at step 496 writing output step 496
Simulation at step 497 writing output step 497
Simulation at step 498 writing output step 498
Simulation at step 499 writing output step 499
adios@adiosVM:~/Tutorial/gray-scott$
adios@adiosVM: ~/Tutorial/gray-scott
File "plot/pdfplot.py", line 49, in PlotPDF
plt.pause(displaysec)
File "/usr/local/lib/python3.5/dist-packages/matplotlib/pyplot.py", line 295, in pause
canvas.start_event_loop(interval)
File "/usr/local/lib/python3.5/dist-packages/matplotlib/backend_bases.py", line 2252, in start_event_loop
time.sleep(timestep)
KeyboardInterrupt
adios@adiosVM:~/Tutorial/gray-scott$
adios@adiosVM: ~/adiosvm/Tutorial/gray-scott
PDF Analysis step 491 processing sim output step 491 sim compute step 491
PDF Analysis step 492 processing sim output step 492 sim compute step 492
PDF Analysis step 493 processing sim output step 493 sim compute step 493
PDF Analysis step 494 processing sim output step 494 sim compute step 494
PDF Analysis step 495 processing sim output step 495 sim compute step 495
PDF Analysis step 496 processing sim output step 496 sim compute step 496
PDF Analysis step 497 processing sim output step 497 sim compute step 497
PDF Analysis step 498 processing sim output step 498 sim compute step 498
PDF Analysis step 499 processing sim output step 499 sim compute step 499
adios@adiosVM:~/adiosvm/Tutorial/gray-scott$
adios@adiosVM: ~/Tutorial/gray-scott
Plot2D ('yz', data, args, fullshape, sim_step[0], fontsize)
File "plot/gsplot.py", line 63, in Plot2D
plt.pause(displaysec)
File "/usr/local/lib/python3.5/dist-packages/matplotlib/pyplot.py", line 295, in pause
canvas.start_event_loop(interval)
File "/usr/local/lib/python3.5/dist-packages/matplotlib/backend_bases.py", line 2252, in start_event_loop
time.sleep(timestep)
KeyboardInterrupt
adios@adiosVM:~/Tutorial/gray-scott$
adios@adiosVM: ~/adiosvm/Tutorial/gray-scott
-rw-rw-r-- 1 adios adios 554 Feb 27 18:22 CMakeLists.txt
drwxrwxr-x 3 adios adios 4.0K Feb 27 18:24 build
-rw-rw-r-- 1 adios adios 3.8K Mar 1 09:07 README.md
drwxrwxr-x 2 adios adios 4.0K Mar 8 09:59 analysis
drwxrwxr-x 3 adios adios 4.0K Mar 13 05:09 plot
-rw-rw-r-- 1 adios adios 2.6K Mar 13 10:05 adios2.xml
drwxrwxr-x 2 adios adios 4.0K Mar 13 10:14 simulation
adios@adiosVM:~/adiosvm/Tutorial/gray-scott$
adios@adiosVM: ~/adiosvm/Tutorial/gray-scott
GetVar: U ts= 219
GetVar: U ts= 220
GetVar: U ts= 221
GetVar: U ts= 222
GetVar: U ts= 223
GetVar: U ts= 224
^Cadios@adiosVM:~/adiosvm/Tutorial/gray-scott$ ^C
adios@adiosVM:~/adiosvm/Tutorial/gray-scott$
adios@adiosVM: ~/Tutorial/gray-scott
-rw-rw-r-- 1 adios adios 384K Feb 27 17:58 visit-sst.session
-rw-rw-r-- 1 adios adios 3.2K Feb 27 17:58 visit-bp.session.gui
-rw-rw-r-- 1 adios adios 339K Feb 27 17:58 visit-bp.session
-rw-rw-r-- 1 adios adios 554 Feb 27 18:22 CMakeLists.txt
drwxrwxr-x 3 adios adios 4.0K Feb 27 18:24 build
-rw-rw-r-- 1 adios adios 3.8K Mar 1 09:07 README.md
drwxrwxr-x 2 adios adios 4.0K Mar 8 09:59 analysis
drwxrwxr-x 3 adios adios 4.0K Mar 13 05:09 plot
-rw-rw-r-- 1 adios adios 2.6K Mar 13 10:05 adios2.xml
drwxrwxr-x 2 adios adios 4.0K Mar 13 10:14 simulation
adios@adiosVM:~/Tutorial/gray-scott$!
adios@adiosVM:~/Tutorial/gray-scott$!v
vi *.xml
adios@adiosVM:~/Tutorial/gray-scott$
adios@adiosVM: ~/adiosvm/Tutorial/gray-scott
{
"L": 48,
"Du": 0.2,
"Dv": 0.1,
"F": 0.02,
"k": 0.048,
"dt": 1.0,
"plotgap": 1,
"steps": 100,
"noise": 0.01,
"output": "gs.bp",
"adios_config": "adios2.xml"
}
"simulation/settings.json" 13L, 204C written 9,14 All

New Developments in Scalable Analytics with R

pbdR is R plus…

- More numerical methods
- Large parallel computing
- GPU infrastructure
- I/O tools, advanced profilers, remote computing





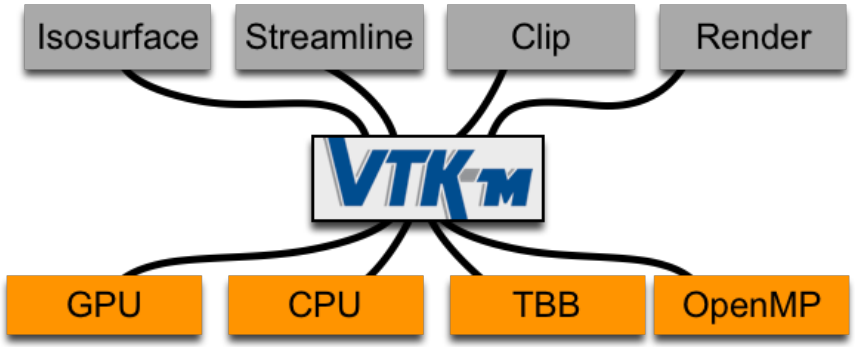
New GPU Capabilities

- New ML methods optimized for Summit GPUs
- Addresses the CORAL 2 benchmarks
- Fast random generators from 6 common distributions
- Bindings for NVIDIA Management Library (NVML)
- Infrastructure for low-level "roll your own" CUDA primitives

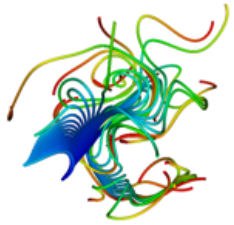
R + summit = pbdR Programming with Big Data in R

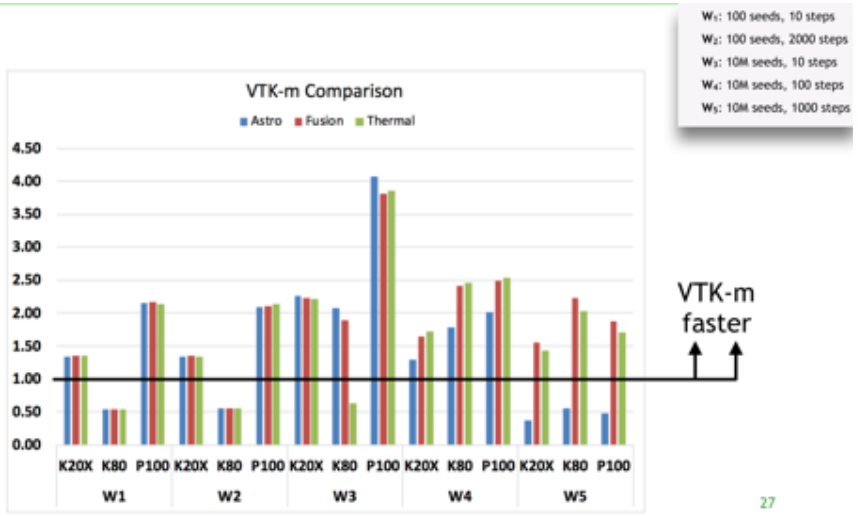
Visualization Services with VTK-m

- Analysis and visualization (1D,2D,3D) are needed to feed a scientific dashboard
- Services are lightweight, configurable components within a scientific workflow
- VTKm is a visualization toolkit aimed at the heterogenous architectures of supercomputers
  - Architecture supports “write once run anywhere” required by visualization services
- Support for large number of key algorithms:
  - Isocontour, slice, histogram, streamline, rendering, …



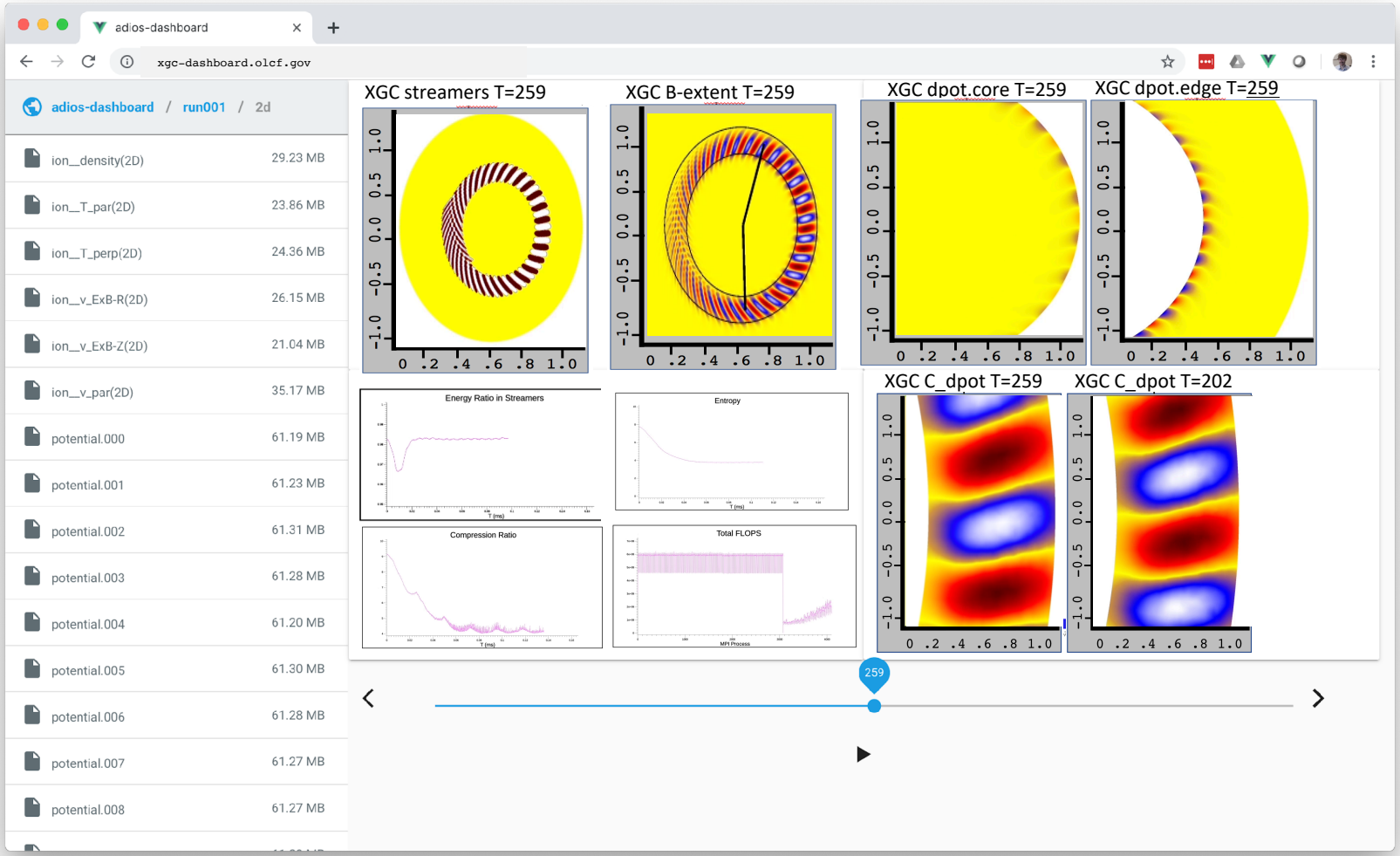
Portable Performance of Particle Advection: The VTKm implementation was shown to be comparable in performance to custom CPU and GPU implementations

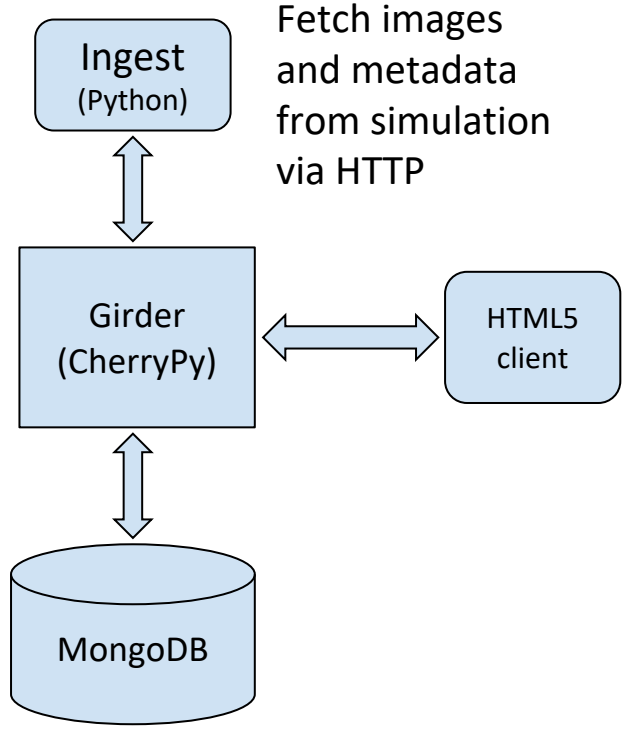




Performance-Portable Particle Advection with VTK-m, D Pugmire, A Yenpure, M Kim, J Kress, R Maynard, H Childs, B Hentschel. EGPGV, 2018

kdash





Fetch images and metadata from simulation via HTTP

Summary: Co-design the next set of tools for federated computing

- The convergence of large DOE instruments with HPC centers dictates that we need to allow coupling/streaming
  - Codesign of what occurs at the edge and at HPC centers is imperative
  - Integration of ML/AI with HPC is essential to process more data
- https://github.com/CODARcode/MGARD
- https://github.com/ornladios/ADIOS2
- https://gitlab.kitware.com/vtk/vtk-m
- https://pbdr.org/packages.html
- https://Adios.ornl.gov