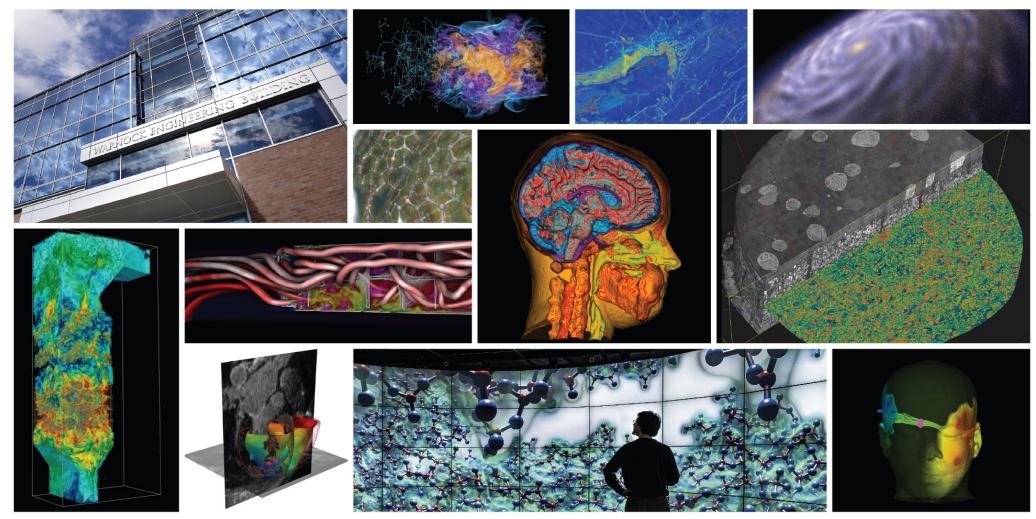
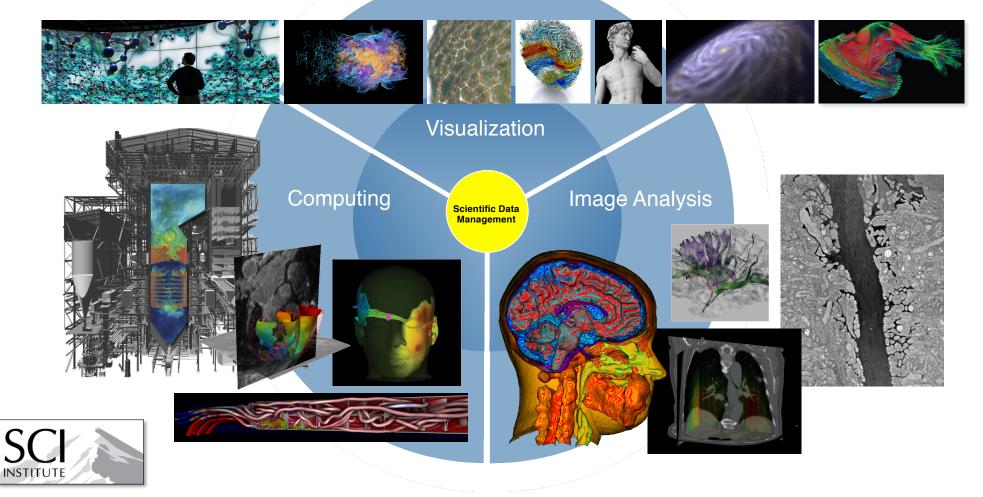
In Situ Visualization Past and Future



Scientific Computing and Imaging Institute



Research Centers at SCI



NIH/NIGMS Center for Integrative Biomedical Computing





Intel Visualization Center





Center for Extreme Data Management, Analysis, and Visualization



National Science Data Fabric



UTAH Center for Computational Earth Sciences



Carbon Capture Multidisciplinary Simulation Center

CDE₃M

Alliance for Computationally-guided Design of Energy Efficient Electronic Materials









Acknowledgments



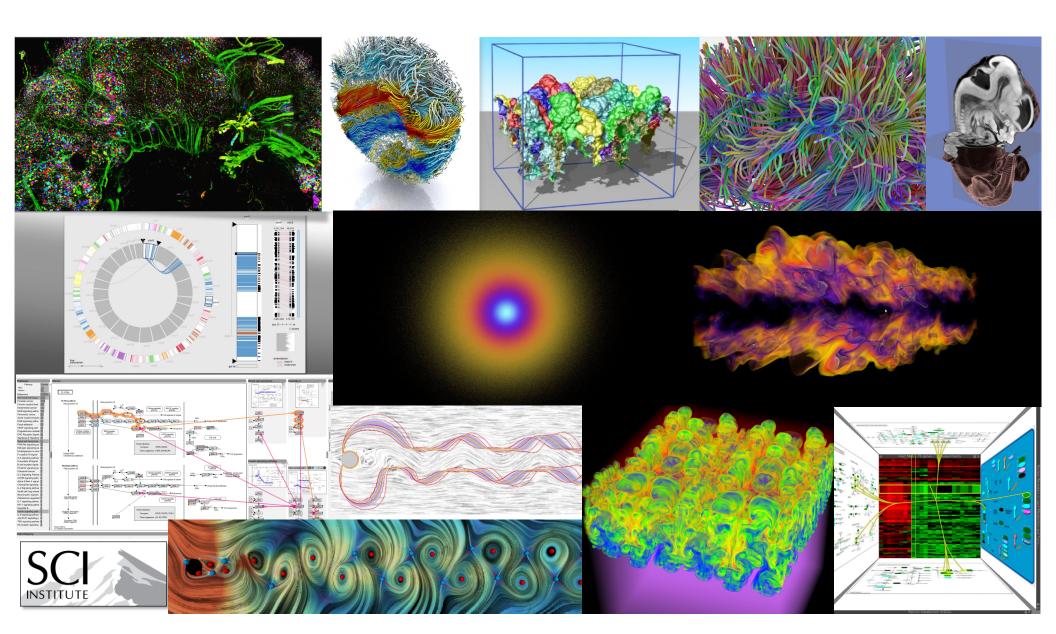






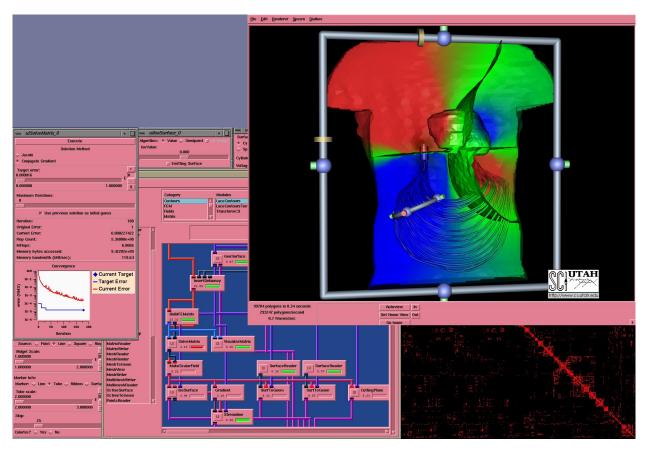






SCIRun - 1994





C.R. Johnson and S.G. Parker. A computational steering model for problems in medicine. *Supercomputing '94*, pp. 540-549, IEEE Press, 1994.



S.G. Parker and C.R. Johnson. SCIRun: A scientific programming environment for computational steering. *Supercomputing '95*, IEEE Press, 1995.





S.G. Parker, D. Beazley, and C.R. Johnson. Computational steering software systems and strategies. *IEEE Computational Science and Engineering*, Vol. 4, Number 4, pp. 50-59, Oct., 1997.

M. Miller, C. Moulding, J. Dongarra, and C.R. Johnson. Grid-enabling problem solving environments: A case study of SCIRun and NetSolve. In *Proceedings of HPC* (pp. 98-103), 2001.

J. Knezevic, R.P. Mundani, E. Rank, A. Khan, and C.R. Johnson. Extending the SCIRun problem solving environment to large-scale applications. *IADIS Conference on Applied Computing*, pp., 171-178, 2012.



A. Narayan, Z. Liu, J. Bergquist, C. Charlebois, S. Rampersad, L. Rupp, D. Brooks, D. White, J. Tate, R.S. MacLeod. UncertainSCI: Uncertainty Quantification for Computational Models in Biomedicine and Bioengineering. *Available at SSRN 4049696, 2022.*

- G. A. Geist, J. A. Kohl, P. M. Papadopoulos, ``CUMULVS: Providing Fault-Tolerance, Visualization and Steering of Parallel Applications," Environment and Tools for Parallel Scientific Computing Workshop, Lyon, France, August 21-23, 1996.
- Interactive Exploration and Modeling of Large Data Sets: A Case Study with Venus Light Scattering Data. J.J. van Wijk, H.J.W. Spoelder, W-J. Knibbe, K.E. Shahroudi, IEEE Visualization '96
- Jeffrey Vetter "Computational Steering Annotated Bibliography". <u>SIGPLAN Notices</u>. **32** (6): 40– 44. doi:10.1145/261353.261359, 1997.



- Nanoscopic visualization of localized corrosion damage by means of atomic force microscope, WIT Transactions on Engineering Sciences 13 (1970). "...nanoscopic in-situ visualization by using Atomic Force Microscopy."
- Computer assisted stereotactic neurosurgery." *Image and vision* computing 11.8, 1993, "Since no direct in situ visualization of the cerebral structures along the trajectory is possible..."
- R. P. Kale, M. E. Fleharty and P. M. Alsing, Parallel molecular dynamics visualization using MPI with MPE graphics, *Proceedings. Second MPI Developer's Conference*, 1996, pp. 104-110. <u>"The main</u> thrust of our research efforts is currently directed towards in situ visualization of the MD simulations."



1997 Sandia Technical Report about ASCI Red:

"Although other distributed strategies for visualizing large data sets are also being considered, several parallel tools are currently being implemented directly on the ASCI Red machine to enable **in situ visualization** of machine capacity data sets thereby avoiding the need to move the data prior to visualization"

> SAND--97-0463C CONF-9706112--1

ASCI Red -Experiences and Lessons Learned with a Massively Parallel TeraFLOP Supercomputer (U)

Mark A. Christon, David A. Crawford, Eugene S. Hertel, James S. Peery, and Allen C. Robinson

> Computational Physics R&D Department Sandia National Laboratories Albuquerque, New Mexico 87185-0819 E-mail: machris@sandia.gov





In Situ Visualization Papers

```
1990 - 2000 - 26
2000 - 2005 - 79
2005 - 2010 - 284
```



- Ma, Kwan-Liu. "VisFiles: the next surge of visualization research." ACM SIGGRAPH Computer Graphics 41.3, 2007.
- K. Ma, "In Situ Visualization at Extreme Scale: Challenges and Opportunities," in *IEEE Computer Graphics and Applications*, vol. 29, no. 6, pp. 14-19, Nov.-Dec. 2009. 139 Citations.
- Cummings, J. and Pankin, A. and Podhosrzki, N. and Park, G. and Ku, S. and Barreto, R. and Klasky, S. and Chang, C. S. and Strauss, H. and Sugiyama, L. and Snyder, P. and Pearlstein, D. and Ludäscher, B. and Bateman, G. and Kritz, A. *Plasma Edge Kinetic-MHD Modeling in Tokamaks Using Kepler Workflow for Code Coupling, Data Management and Visualization.* Communications in Computational Physics, 4 (3). pp. 675-702, 2008
- Yu, H., Wang, C., Grout, R. W., Chen, J. H., & Ma, K. L. In situ visualization for large-scale combustion simulations. *IEEE computer graphics and applications*, *30* (3), 45-57, 2010.



In Situ Visualization Papers

```
1990 - 2000 - 26
2000 - 2005 - 79
2005 - 2010 - 284
2010 - 2015 - 803
2015 - 2020 - 1,580
2020 - 2022 - 783 (~ 2,000 by 2025)
```



Fabian, N., Moreland, K., Thompson, D., Bauer, A.C., Marion, P., Gevecik, B., Rasquin, M. and Jansen, K.E. The Paraview coprocessing library: A scalable, general purpose in situ visualization library. In *2011 IEEE Symposium on Large Data Analysis and Visualization* (pp. 89-96), 2011. 281 Citations

In Situ Motivation - Early Days

Jackie Chen - Kwan-Liu Ma - Combustion simulation data that is multiscale, multi-variate, time varying and three-dimensional. The data was intrinsically intermittent and highly transient (turbulence, unsteady ignition and extinction events) necessitating performing the analysis and visualization in situ since we weren't able to store data at sufficient frequency and the IO rates and storage capacity were limited.

Success Story: Then postdoc, Hongfeng Yu (now a professor at U. Nebraska), was resident at Sandia with Jackie where he immersed himself with combustion scientists to learn what their needs were. From this interaction he developed in situ multi-variate and particle visualization and in-situ parallel distance field computation with respect to dynamical turbulent flame surfaces.



Jackie Chen: Future requirements for in situ visualization, analytics, ML in simulation workflows

- Adaptive data placement for staging-based coupled scientific workflows to address complex and dynamic data exchange patterns exhibited by the workflows.
- Take advantage of application-specific data access patterns to adaptively place data with an awareness of the system network topology to reduce data access costs and enable efficient data sharing.
- Identify and characterize the dynamic data access patterns of data consumer applications at runtime using a combination of user provided hints and knowledge of prior access behaviors.



Jackie Chen: Future requirements for in situ visualization, analytics, ML in simulation workflows (continued)

- Including visualization as part of a more complex simulation and data science workflow at exascale with triggers for steering analytics and reduced order modeling on the fly and visualization the results is a challenge.
- Programming which transient events to look for, extract, compute statistics on, and track forwards and backwards in time is still a challenge, especially for large multi-scale data.



In Situ Motivation - Early Days

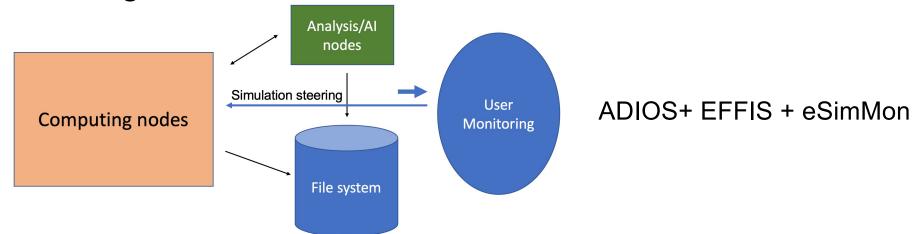
CS Chang: Extreme-scale Fusion Simulation. In 2005 DOE awarded the Prototype Fusion Simulation Project (FSP) Center for Plasma Edge Simulation (CPES), converted to SciDAC 2.

There was a need to monitor, during the simulation:

- when and how the kinetic buildup of tokamak edge pressure induces a sudden MHD instability
- how turbulence is developed and affecting the edge pressure buildup
- how to steer the simulation by parameter injection
- Code coupling between kinetic XGC and an MHD code was a necessity for this work
- Scott Klasky's DM group (then PPPL), led by Arie Shoshani (LBNL, then director of SDM center), used Kepler to create an in situ visualization workflow called EFFIS for code coupling
- Service oriented Architecture was used in **EFFIS** (End-to-end Framework for Fusion Integrated Simulation)
- **eSimMon** Dashboard was created for collaborative in situ visualization of the simulation data, including macroscopic plasma profile evolution, turbulent dynamics and code coupling status
- Nagiza Samatova (NCSU) added mathematical data-analysis capability in eSimMon
- Steve Parker (then U. Utah) added visually pin-pointed simulation analysis capability in eSimMon
 - Use the mouse click on a spot to get the local result data

During SciDAC 3 and 4, Kepler was replaced by ADIOS code-coupling workflow framework.

CS Chang: Future requirements for in situ visualization, analysis and monitoring simulation workflows

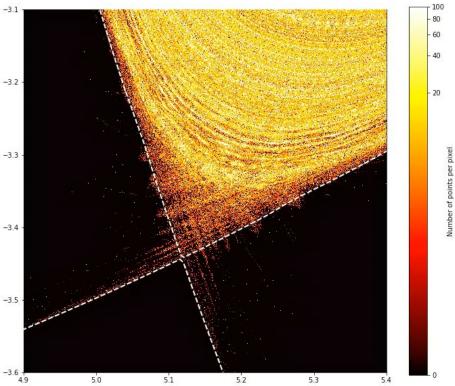


Utilize a hybrid in situ monitoring and analysis method

- In-line analysis for well parallized part.
- On-line asynchronous analysis and visualization for poorly parallized but time-consuming part: offload the in situ analysis/visualization to independent analysis nodes.
- Utilize AI/ML in analysis nodes for feature detection, UQ, validation and simulation steering
 - Simulation steering: AMR where needed, filter out known instabilities, etc.
- Analysis nodes also reduce and compress data for further post processing.
- Utilize accelerators as much as possible.
- Extend this technique for collaborative research on big experiments.

CS Chang: On-line visualization of Poincare puncture movie in XGC from fluctuating homoclinic tangle in ITER edge

- Electromagnetic turbulence could partially destroy the last closed magnetic surface called "separatrix" surface (white dashed line)
- At every simulation timestep, the fluctuating magnetic field data is asynchrously off-loaded to an analysis load for on-line visualization
 - XGC simulation continues without interruption
- In-line visualization would have doubled the XGC simulation time (full-scale Summit is used)
 - The Poincare visualization routine is difficult to be massively parallelized, like XGC.
- ✤ XGC simulation by S. Ku (PPPL)
- ADIOS2.0 data movement by S. Klasky's group (ORNL)
- Poincare puncture-movie by D. Pugmire's group (ORNL)



Top In Situ Visualization Challenges - 2019 Dagstuhl Workshop

- Data quality and reduction, i.e., reducing data in situ and then exploring it post hoc, which is likely the form that will enable exploration of large data sets on future supercomputers.
- Workflow execution, i.e., how to efficiently execute specified workflows, including workflows that are very complex.
- Software complexity, heterogeneity, and user-facing issues, i.e., the challenges that prevent user adoption of in situ techniques because in situ software is complex, computational resources are complex, etc.
- Exascale systems, which will have billion-way concurrency and disks that are slow relative to their ability to generate data.
- Algorithmic challenges, i.e., algorithms will need to integrate into in situ ecosystems and still perform efficiently.
- Workflow specification, i.e., how to specify the composition of different tools and applications to facilitate the in situ discovery process.
- Use cases beyond exploratory analysis, i.e., ensembles for uncertainty quantification and decision optimization, computational steering, incorporation of other data sources, etc.
- Exascale data, i.e., the data produced by simulations on exascale machines will, in many cases, be fundamentally different than that of previous machines.
- **Cost models**, which can be used to predict performance before executing an algorithm and thus be used to optimize performance overall.
- Convergence of HPC and Big Data for visualization and analysis, i.e., how can developments in one field, such as machine learning for Big Data, be used to accelerate techniques in the other?

H. Childs, J.C. Bennett, C. Garth (editors). In Situ Visualization for Computational Science, Springer, 2022





In Situ Visualization for Computational Science

Deringer

My In Situ Visualization Challenges and Opportunities

Reproducibility



Reproducibility

National Academies of Sciences, Engineering, and Medicine, *Reproducibility and Replicability in Science*. The National Academies Press, 2019. [Online]. Available: https://www.nap.edu/catalog/25303/reproducibility-and-replicability-in-science

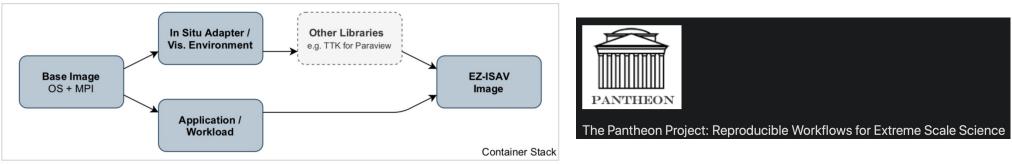
Jean-Daniel Fekete, Juliana Freire. *Exploring Reproducibility in Visualization*. IEEE Computer Graphics and Applications, Institute of Electrical and Electronics Engineers, 2020, 40 (5), pp.108-119.

Reproducibility and Replicability (R&R) tools e.g., ReproZip, Docker, Jupyter and repositories zenodo.org, osf.org—that make it easier to publish transparent, R&R results. *It is worth noting that there are possible limitations, regarding humans, hardware, and software, that can hamper reproducibility. In particular, special hardware is an obstacle to reproducibility. Visualization is particularly rich in special hardware, from HPC to display technologies like VR, AR, wall-sized displays, and physical visualizations.*



Reproducibility

EZ-ISAV - A work-in-progress towards a framework for easy construction of customizable in situ pipelines in container images. Designed for portability and ease of use, these images are intended to serve as proof-of-concept cases for in situ visualization and analysis research...reduce the overhead of developing and evaluating in situ techniques and provide **improved reproducibility** and portability of in situ visualization research.



pantheonscience.org

Figure 1: Overview of the EZ-ISAV container image framework.



Michael Will, Quincy Wofford, John Patchett, David Rogers, Jonas Lukasczyk, and Christoph Garth. 2021. Developing and Evaluating In Situ Visualization Algorithms using Containers. ISAV'21: In Situ Infrastructures for Enabling Extreme-Scale Analysis and Visualization. Association for Computing Machinery, New York, NY, USA, 6–11.

My In Situ Visualization Challenges and Opportunities

- Reproducibility
- High Performance Data Movement



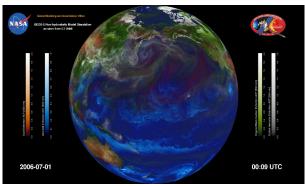
Scalable Deployment: Exploration of 3.5PB of NASA Weather/Climate Data in Real Time

Workflow

Processing _

_

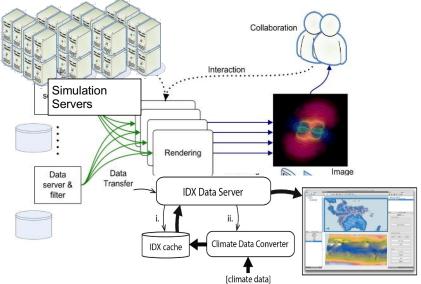
- Data creation
- Data Management ٠
- Analysis - Visualization

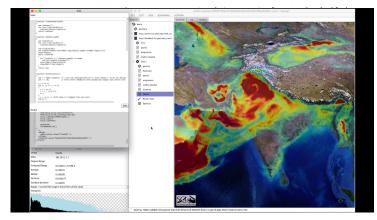


- 7km GEOS-5 "Nature Run" •
- 1 dataset, 3.5 PB
- theoretically: openly accessible
- practically: precomputed pics •

Distributed Resources

- 3.5 PB of data store in NASA
- Primary ViSUS server in LLNL
- Secondary ViSUS server in Utah
- Clients connect remotely
- Work without additional HPC resources _



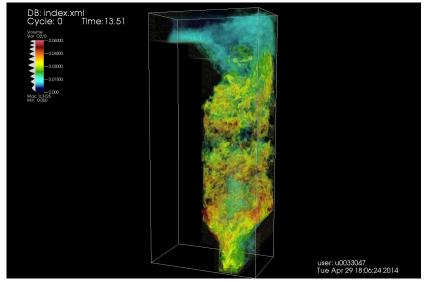


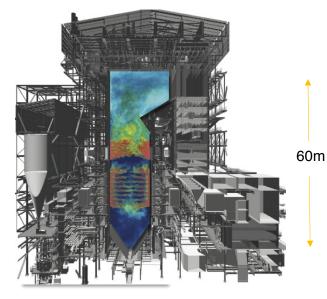
2



DOE PSAAP2 Simulations of GE Clean(er) Coal Boilers

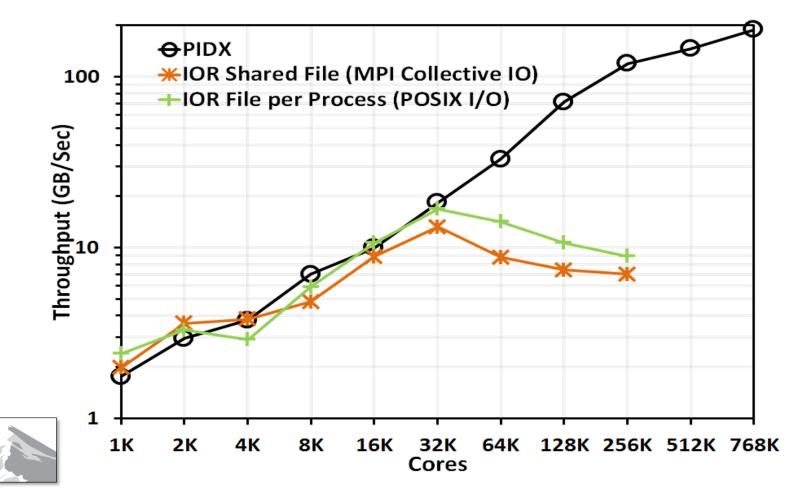
- Large scale turbulent combustion needs mm scale grids 10¹⁴ mesh cells 10¹⁵ variables (1000x more than now)
- Structured, high order finite-volume discretization
- Mass, momentum, energy conservation
- LES closure, tabulated chemistry
- PDF mixing models
- DQMOM (many small linear solves)
- Uncertainty quantification





- Low Mach number approx. (pressure Poisson solve up to 10^12 variables. 1M patches 10 B variables
- Radiation via Discrete Ordinates many hypre solves Mira (cpus) or ray tracing Titan (gpus strong and weak scaling via AMR).
- FAST I/O needed PIDX for scalability

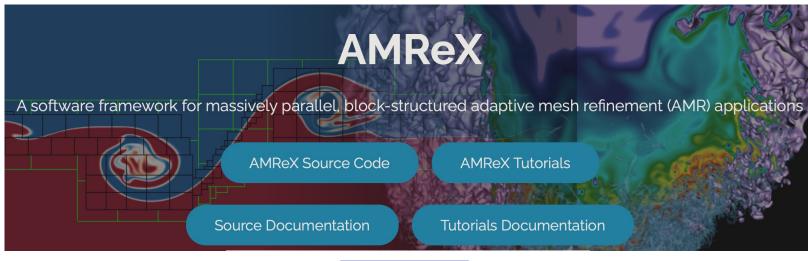
High Performance Data Movement for Real-Time Monitoring of Large Scale Simulations

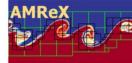


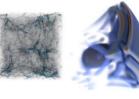
My In Situ Visualization Challenges and Opportunities

- Reproducibility
- High Performance Data Movement
- Adaptive Meshes and High Order Simulation
 - Most large-scale simulation use adaptive meshes and/or high order basis functions, however, most visualization algorithms do not.











Nyx WarpX INT-179 INT-825

AMRWind Pele INT-1350 INT-133





AMR Visualization

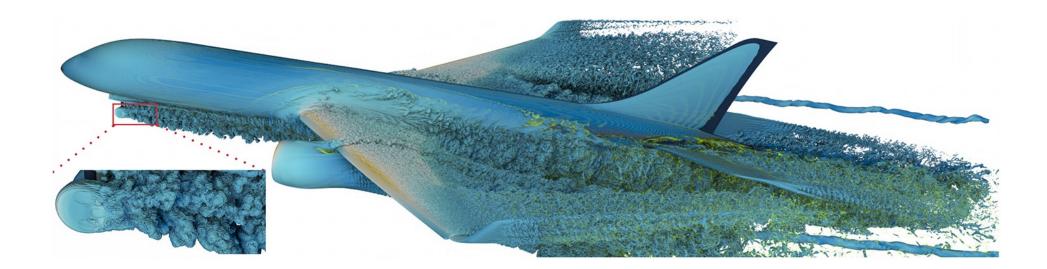


Colliding Black Holes

NASA Exajet Landing Gear



F. Wang, I. Wald, Q. Wu, W. Usher, C. R. Johnson. "**CPU Isosurface Ray Tracing** of Adaptive Mesh Refinement Data," In *IEEE Transactions on Visualization and Computer Graphics*, Vol. 25, No. 1, IEEE, pp. 1142-1151. Jan, 2019.



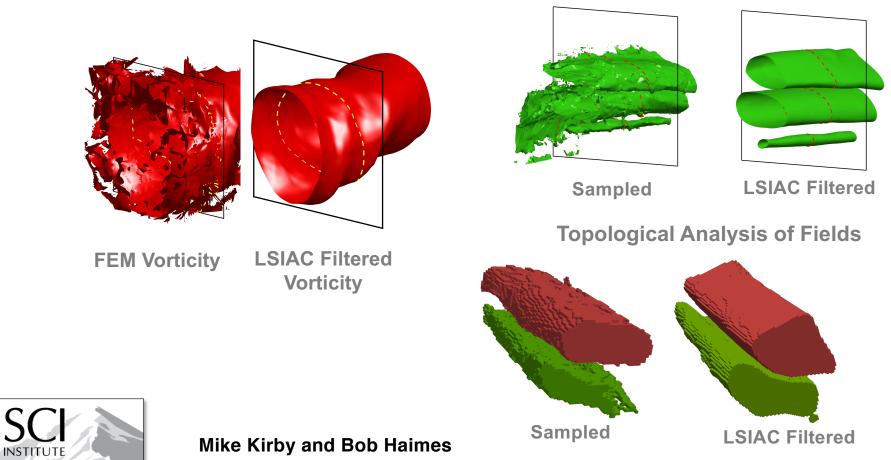
CPU Ray-tracing of Tree-based Adaptive Mesh Refinement Data

Feng Wang, Nathan Marshak, Will Usher, Carsten Burstedde Aaron Knoll, Timo Heister, and Chris R. Johnson



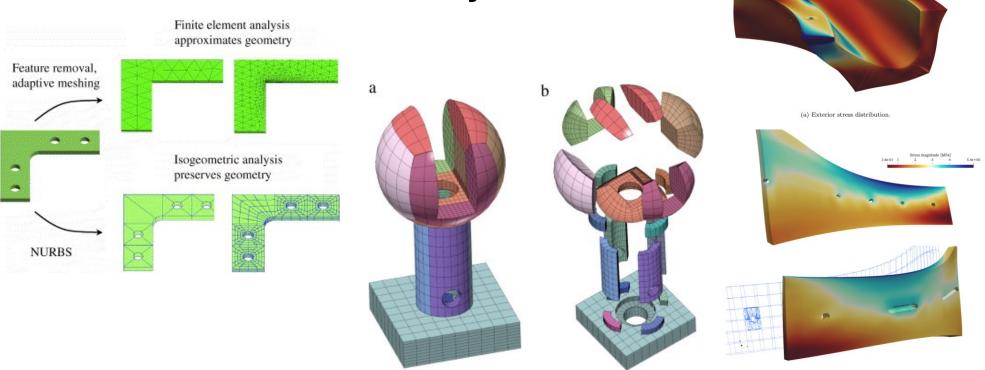
F. Wang, N. Marshak, W. Usher, C. Burstedde, A. Knoll, T. Heister, C. R. Johnson. "CPU Ray Tracing of Tree-Based Adaptive Mesh Refinement Data," In *Eurographics Conference on Visualization (EuroVis)* 2020, Vol. 39, No. 3, 2020.

High-Order FEM Visualization



Counter-Rotating Vortex Vorticity

Iso-Geometric Analysis



(b) Interior stress along some a longitudinal section. Both pictures correspond to the front and back views of a one element thick slice.

Stress magnitude [MPa] 6 8 10 12 3



Cottrell, J.A., Hughes, T.J. and Bazilevs, Y. *Isogeometric Analysis: Toward Integration of CAD and FEA*. John Wiley & Sons, 2009.

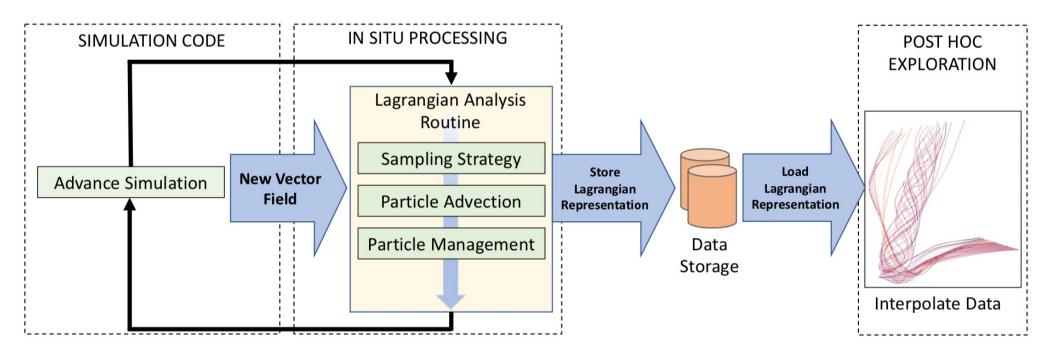
F. Massarwi, G. Elber. A B-spline based framework for volumetric object modeling, Computer-Aided Design, Volume 78, pp. 36-47, 2016.

My In Situ Visualization Challenges and Opportunities

- Reproducibility
- High Performance Data Movement
- Adaptive Meshes and High Order Simulation
 - Most large-scale simulation use adaptive meshes and/or high order basis functions, however, most visualization algorithms do not.
- Domain Expertise and Compact Analysis Techniques
 - Feature Extraction, TDA, Lagrangian Representations



In Situ Lagrangian Analysis

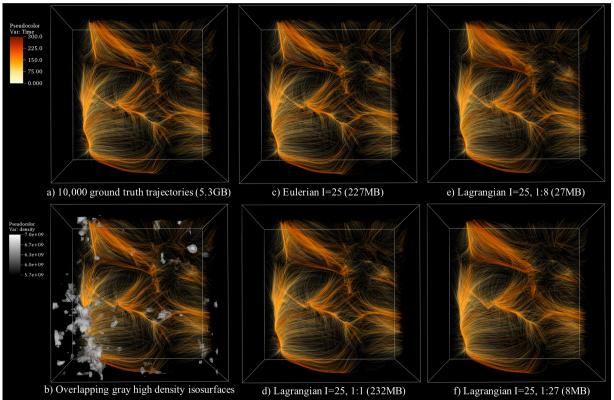


S. Sane, C.R. Johnson, H. Childs. Investigating the Use of In Situ Reduction via Lagrangian Representations for Cosmology and Seismology Applications. *International Conference on Computational Science 2021*. Best Paper Award.



S. Sane, A. Yenpure, R. Bujack, M. Larsen, K. Moreland, C. Garth, C. R. Johnson, and H. Childs. Scalable In Situ Computation of Lagrangian Representations via Local Flow Maps. *Eurographics Symposium on Parallel Graphics and Visualization (EGPGV)* 2021. Best Paper Award.

In Situ Lagrangian Analysis

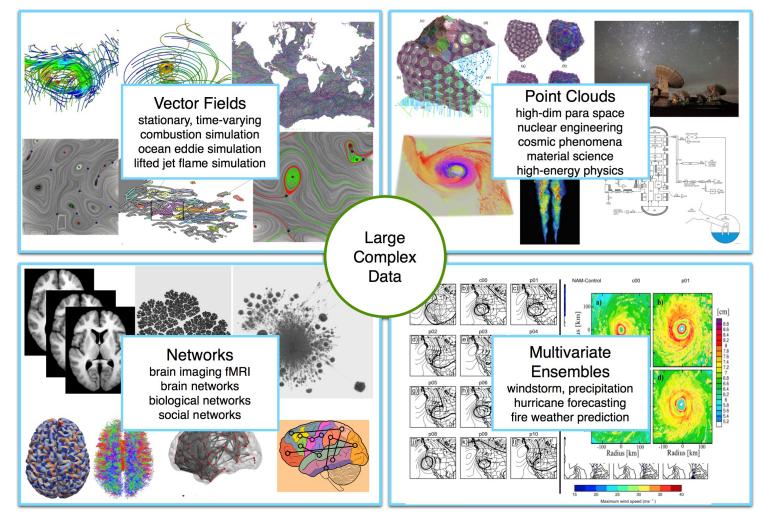


S. Sane, C.R. Johnson, H. Childs. Investigating the Use of In Situ Reduction via Lagrangian Representations for Cosmology and Seismology Applications. *International Conference on Computational Science 2021*. **Best Paper Award**.

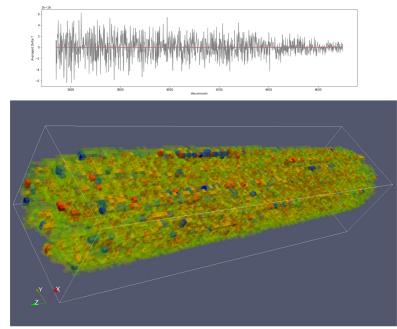


S. Sane, A. Yenpure, R. Bujack, M. Larsen, K. Moreland, C. Garth, C. R. Johnson, and H. Childs. Scalable In Situ Computation of Lagrangian Representations via Local Flow Maps. *Eurographics Symposium on Parallel Graphics and Visualization (EGPGV)* 2021. Best Paper Award.

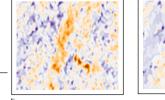
Topological Data Analysis and Visualization

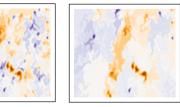


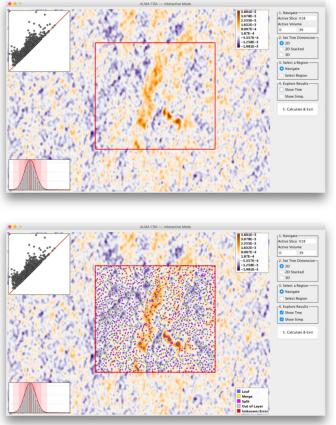
Topological Data Analysis for Astronomical Data Cubes



Analysis of cosmic voids







P. Rosen, A. Seth, E. Mills, A. Ginsburg, J. Kamenetzky, J. Kern, C.R. Johnson and B. Wang. Using Contour Trees in the Analysis and Visualization of Radio Astronomy Data Cubes. In *Topological Methods in Data Analysis and Visualization VI*, pp. 87–108, Springer-Verlag, 2021.



Topological Data Analysis and Visualization

- Most topological analysis, from persistent homology to merge trees to Morse-Smale complexes require global information.
- There are some topological tools used for geometry/topology-based stratification learning that use local homology to infer structure in local neighborhoods:

Brown, A., Wang, B. Sheaf-Theoretic Stratification Learning from Geometric and Topological Perspectives. *Discrete Computational Geometry* **65**, 1166–1198, 2021.



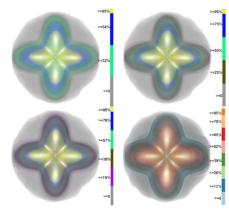
My In Situ Visualization Challenges and Opportunities

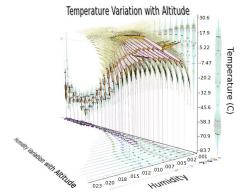
- Reproducibility
- High Performance Data Movement
- Adaptive Meshes and High Order Simulation
 - Most large-scale simulation use adaptive meshes and/or high order basis functions, however, most visualization algorithms do not.
- Domain Expertise and Compact Analysis Techniques
 - Feature Extraction, TDA, Lagrangian Representations
- Uncertainty and Error Propagation
 - Information loss. Data reduction is used throughout in situ visualization pipelines. Plus, there is already uncertainty in the



simulation data and visualization algorithms. How do we assess overall quality?

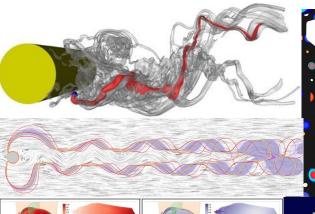
Uncertainty Visualization

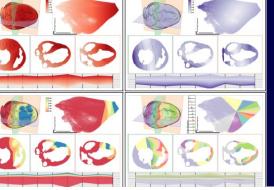


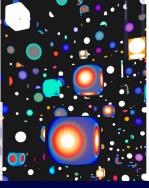




When is the last time you've seen an error bar on an isosurface?







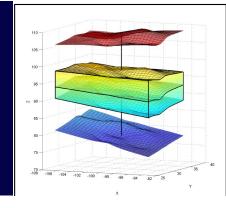
G.P. Bonneau, H.C. Hege, C.R. Johnson, M.M. Oliveira, K. Potter, P. Rheingans, T. Schultz. "Overview and State-ofthe-Art of Uncertainty Visualization," In Scientific Visualization: Uncertainty, Multifield, Biomedical, and Scalable Visualization, Edited by M. Chen and H. Hagen and C.D. Hansen and C.R. Johnson and A. Kauffman, Springer-Verlag, pp. 3-27. 2014.

M.G. Genton, C.R. Johnson, K. Potter, G. Stenchikov, Y. Sun. "Surface boxplots," In *Stat Journal*, Vol. 3, No. 1, pp. 1-11. 2014.

K. Potter, P. Rosen, C.R. Johnson. "From Quantification to Visualization: A Taxonomy of Uncertainty Visualization Approaches," In *Uncertainty Quantification in Scientific Computing*, IFIP Series, Vol. 377, Springer, pp. 226-249. 2012.

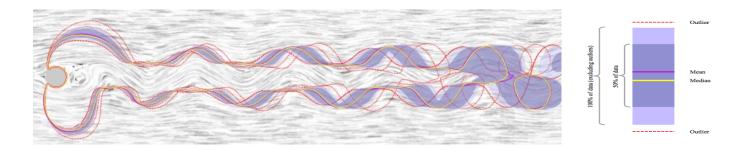
K. Potter, A. Wilson, P.-T. Bremer, D. Williams, C. Doutriaux, V. Pascucci, C.R. Johnson. "Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data," In *Proceedings of the 2009 IEEE International Conference on Data Mining Workshops*, pp. 233-240. 2009.

C.R. Johnson, A.R. Sanderson. "A Next Step: Visualizing Errors and Uncertainty," In *IEEE Computer Graphics and Applications*, Vol. 23, No. 5, pp. 6-10. September/October,



Contour Box Plots

$$S \in \mathrm{sB}(S_1, \ldots S_j) \iff \bigcap_{k=1}^j S_k \subset S \subset \bigcup_{k=1}^j S_k$$



Whitaker, Mirzargar, Kirby, *IEEE Transactions on Visualization and Computer Graphics*, Vol. 19, No. 12, pp. 2713--2722, 2013.



M.G. Genton, C.R. Johnson, K. Potter, G. Stenchikov, Y. Sun. "Surface boxplots," In *Stat Journal*, Vol. 3, No. 1, pp. 1-11. 2014.

Uncertainty Visualization of the Marching Squares and Marching Cubes Topology Cases - VIS 2021

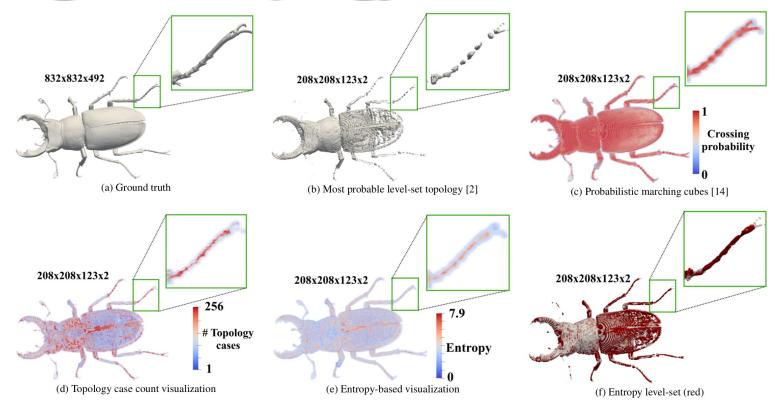




Figure 5: Uncertainty visualizations for the stag beetle [21] hixel dataset at k = 900. The noise in the data results in breaking of the beetle leg in image (b). In probabilistic marching cubes, it is difficult to distinguish between the regions of high and topological uncertainty, which is easier using our visualizations in images (d-f). The relatively high sensitivity of the beetle leg topology to noise is detected in images (d-f) by the red regions. In image (f), the most probable level-set (gray) is overlaid with the entropy volume level-set (red) for entropy isovalue 5.

Back to Computational Steering and In Situ Visualization Together Again?

"How do the above considerations change if in situ interactive exploration (mandating short response times) is considered, e.g. for computational steering applications?"

Report from Dagstuhl Seminar 18271 In Situ Visualization for Computational Science Edited by Janine C. Bennett, Hank Childs, Christoph Garth, and Bernd Hentschel. Available at: https://www.osti.gov/pages/servlets/purl/1492333



Productivity Machines







More Information

www.sci.utah.edu

crj@sci.utah.edu

