ALPINE Infrastructure and Algorithms

Approved for public release

EXASCALE COMPUTING PROJECT

Los Alamos National Laboratory: James Ahrens (PI), Terece Turton (PC), Ayan Biswas, Soumya Dutta, Subhashis Hazarika. Lawrence Livermore National Laboratory: Eric Brugger, Timo Bremer, Cyrus Harrison, Matthew Larsen, Sergei Shudler. Lawrence Berkeley National Laboratory: Gunther Weber, Oliver Rübel. Oak Ridge National Laboratory: Dave Pugmire, Jieyang Chen. Argonne National Laboratory: Joe Insley, Silvio Rizzi, **Victor Mateevitsi**, Saumil Patel. Sandia National Laboratories: Janine Bennett, Marco Arienti. Kitware: Berk Geveci, Utkarsh Ayachit, Dan Lipsa. University of Oregon: Hank Childs (Deputy PI), Nicole Marsaglia, Yuya Kawakami. University of Utah: Valerio Pascucci, Xuan Huang, Steve Petruzza. University of Leeds: Hamish Carr, Petar Hristov.

2021 ECP Community BOF Days, March 30 2021

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY | Office of Science

ALPINE: In Situ Visualization, Analysis, and Infrastructure for ECP Science

- What is ALPINE?
  - Cross-institution effort to leverage and develop new technologies
  - In situ visualization and data analysis algorithms and infrastructure integrated directly to ECP Science Applications

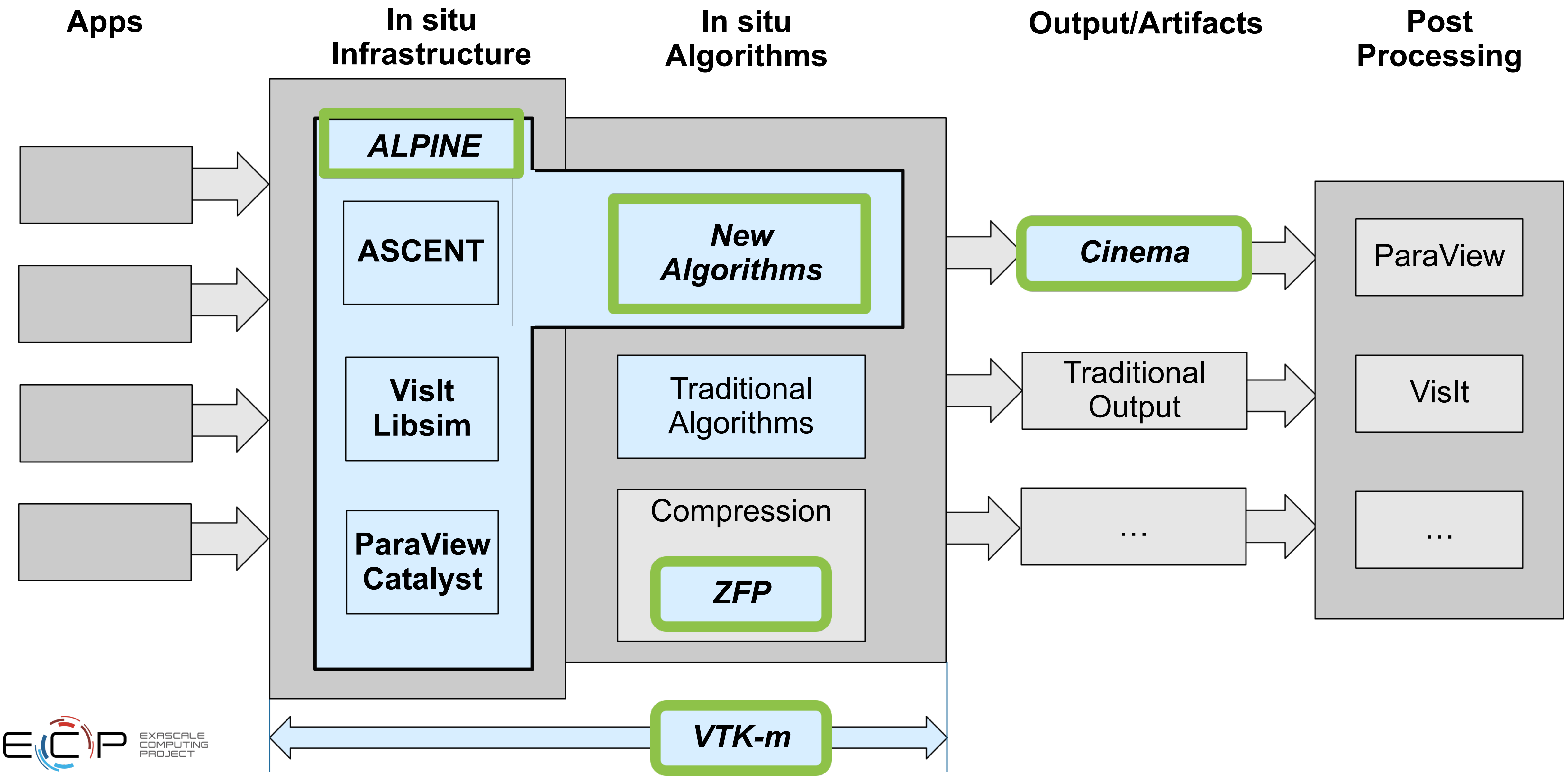
In Situ Data Visualization and Analysis

- What is In Situ Data Visualization and Analysis?
  - Data is processed in situ: as it is generated
  - Visualization and analysis code is coupled with the simulation code
  - Goal: Avoid file system I/O

- Pros
  - Reading/writing (post-hoc) files is slower than running the simulation. Greatly improves vis/analysis speed
  - Can access all the data
  - Take advantage of the computational power of the entire supercomputer

- Cons
  - A-priori knowledge of what to look for
  - Increased complexity due to code instrumentation
  - Memory and network constrains

ECP Software Technology Data and Visualization projects provide an integrated workflow



ALPINE Infrastructure: in situ & post hoc

IN SITU

- Catalyst – ParaView's in situ library

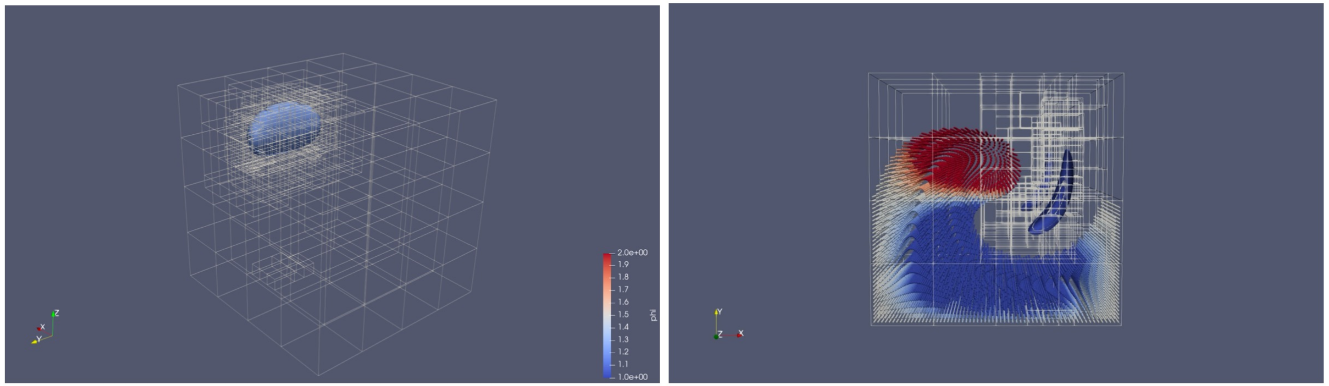
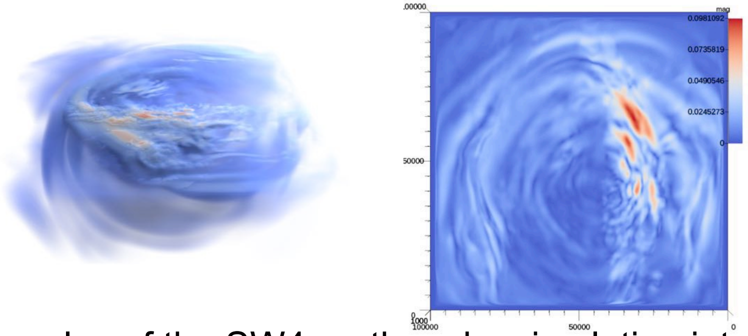


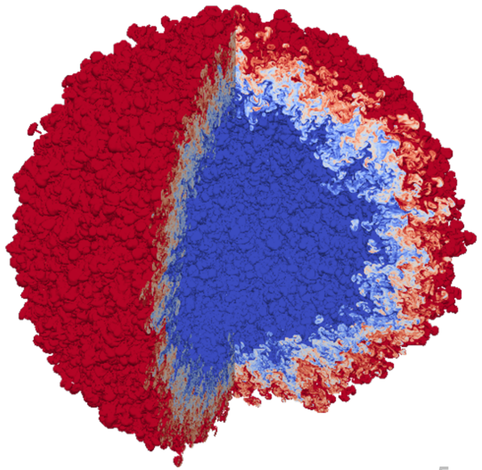
Figure 1: (a) The AmrCore example using amrex::AmrMesh. (b) The AmrLevel example utilizing both amrex::Amr and amrex::ParticleContainer simultaneously.

- Ascent -- ALPINE's new flyweight in situ API



Examples of the SW4 earthquake simulation integrated into ALPINE's Ascent infrastructure. Right: the displacement magnitude in a shock wave. Left: using a VTK-m renderer to visualize the SW4 simulation.

Visualizing turbulent fluid mixing using 16,384 GPUs on LLNL's Sierra: Visualization of an idealized Inertial Confinement Fusion (ICF) simulation of Rayleigh-Taylor instability with two fluids mixing in a spherical geometry.

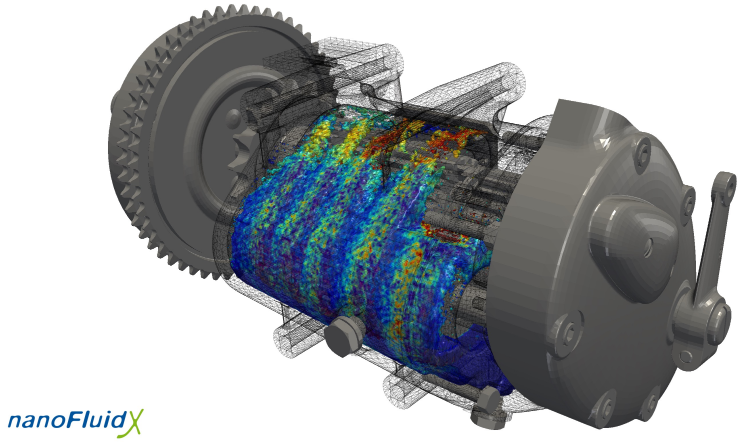


EXASCALE COMPUTING PROJECT

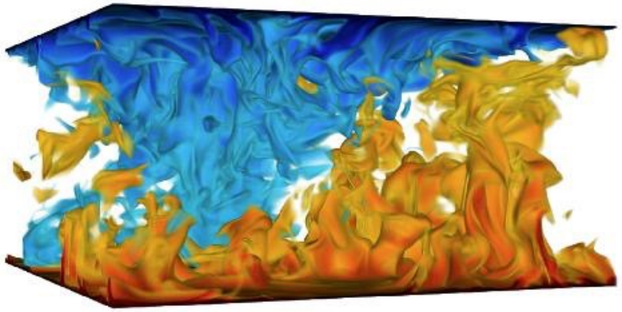
ALPINE Infrastructure: in situ & post hoc

POST HOC

- ParaView



- VisIt



ASCENT: https://github.com/Alpine-DAV/ascent
Paraview: https://www.paraview.org
VisIt: https://wci.llnl.gov/simulation/computer-codes/visit

ALPINE in situ data analysis and visualization algorithms

- Data reduction
- Feature detection in situ
- Identifying important events in situ

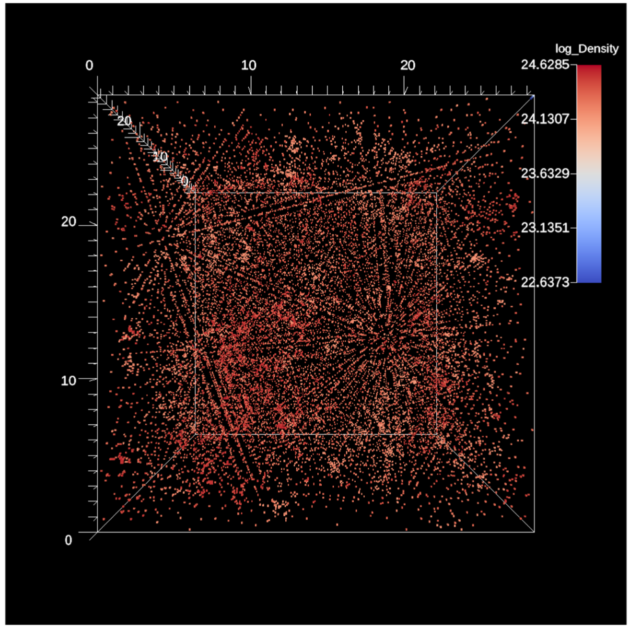
ALPINE algorithms:
- Sampling
- Statistical feature detection
- Topology: contour tree
- Optimal viewpoint metrics
- Task-based feature extraction
- Moments-based pattern detection
- Lagrangian field flow

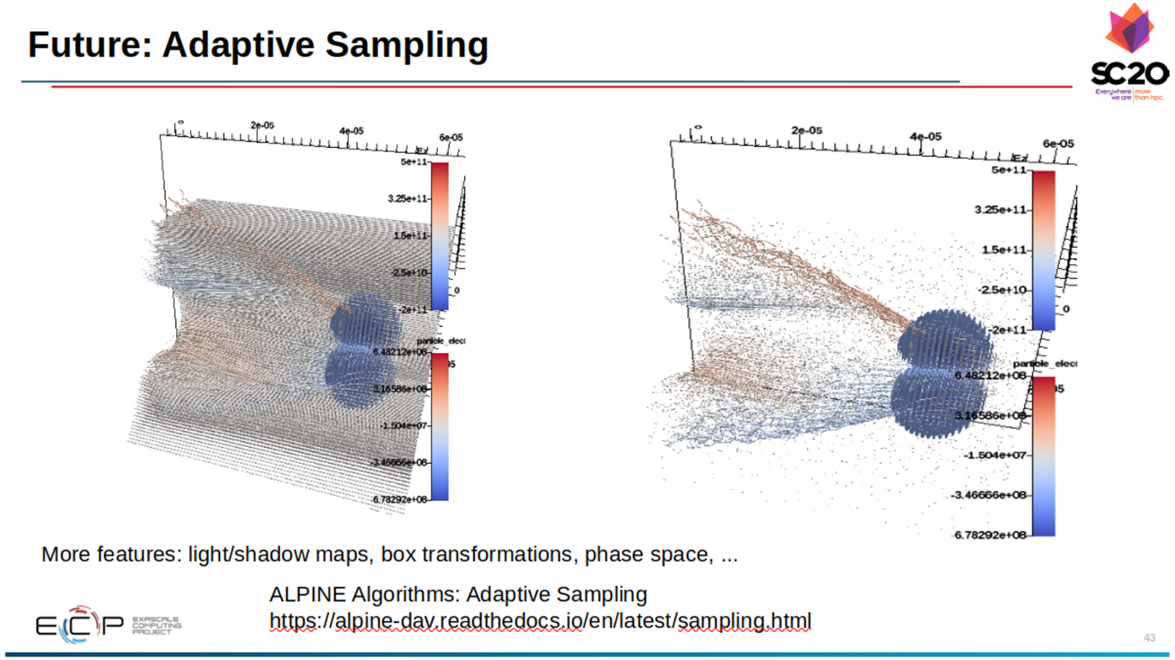
Integrating with and leveraging other ECP software technologies
- Leveraging VTK-m for cross-platform portability
- Spack/E4S for interoperability, fast builds, reproducibility
- Adding I/O capability for HDF5, ADIOS, Cinema

Sampling Algorithm Integrated into Ascent & ECP Applications

Goal: In situ data reduction

Data-driven sampling enables probabilistic identification of interesting regions in the data automatically, prioritizing important regions. Applied in situ to Nyx, important halo regions are preserved.





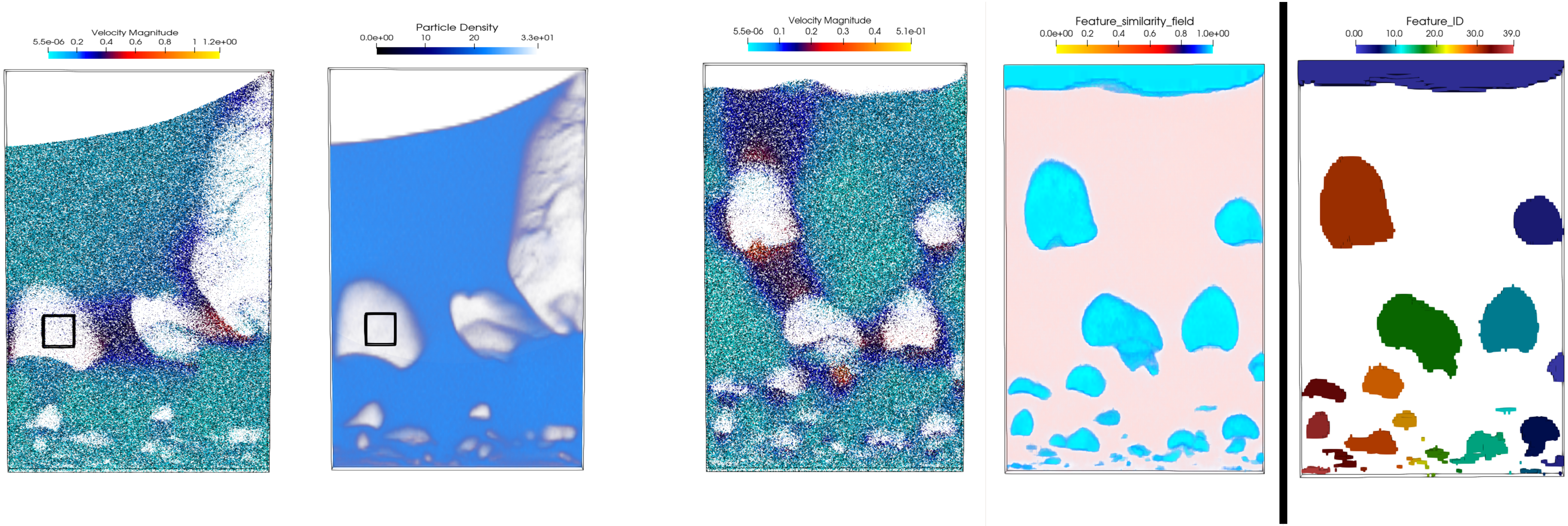
WarpX applies in situ sampling to transverse momentum to preferentially select particles behaving unexpectedly.

Image curtesy of A. Huebl (WarpX)

EXASCALE COMPUTING PROJECT

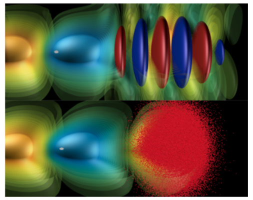
Statistical Feature Detection integrated into Catalyst & MFiX Exa

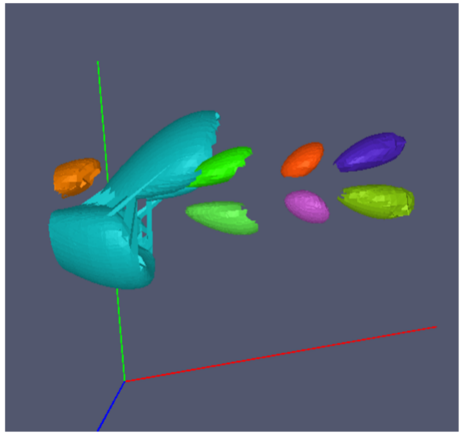
In situ **statistical feature detection** detects features in particle data sets using statistical data modeling and probabilistic similarity measures.

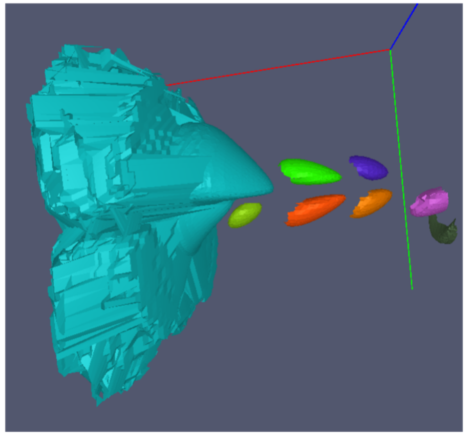


Goal: in situ data reduction & feature detection with post hoc interactive analysis via Cinema

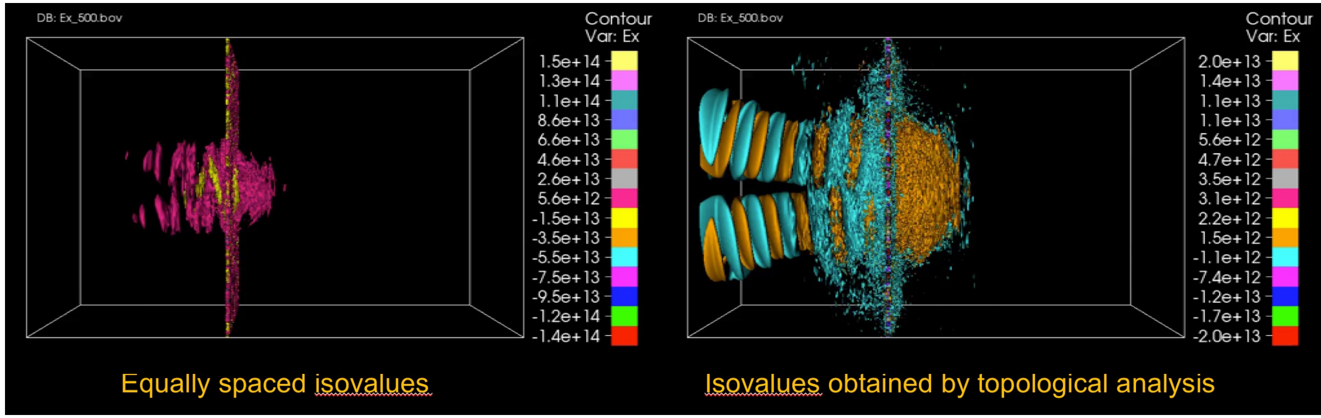
Contour tree integration with WarpX







Topological analysis is used to identify most relevant contours & create isosurface visualizations in situ; saving resulting images for post hoc analysis. Images are saved to a Cinema DB in a format that supports arbitrary combination of contours during post hoc visualization. Right: Most relevant contours in WarpX simulation selected using two importance measures: persistence, volume.

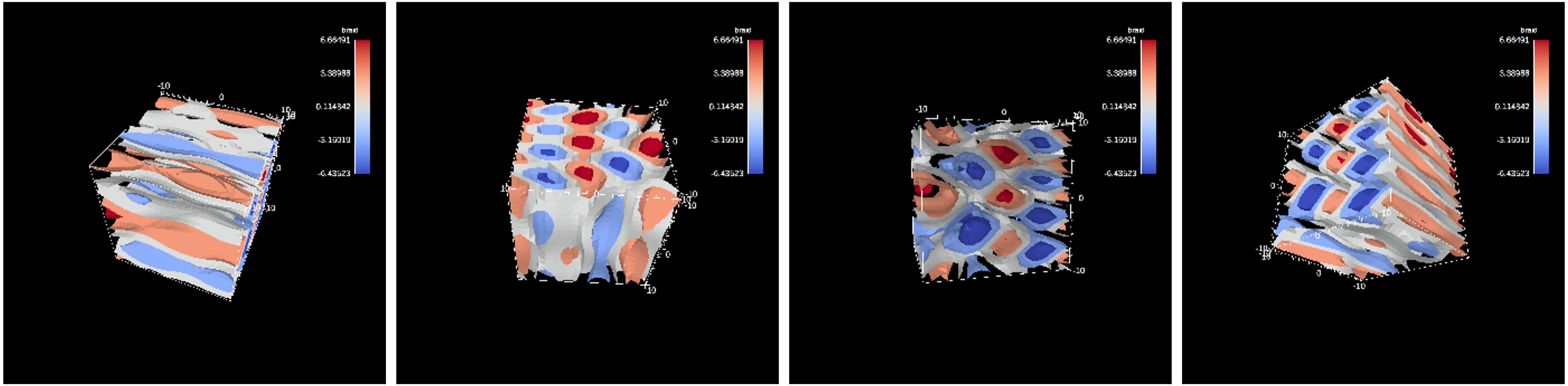


Topological analysis can be used to detect the most significant isosurfaces in complex simulations. At left, equally spaced isovalues in an ion accelerator simulation. Above, our method chooses isovalues using topological analysis to more fully represent complex behavior in the data.

EXASCALE COMPUTING PROJECT

Algorithms

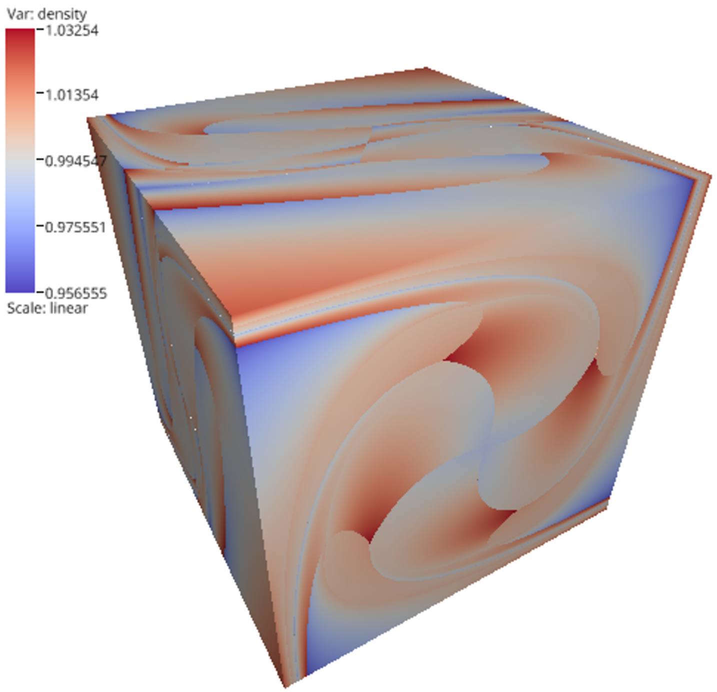
Optimal Viewpoint Metrics based on data properties can automate in situ visualization decisions to only capture interesting views



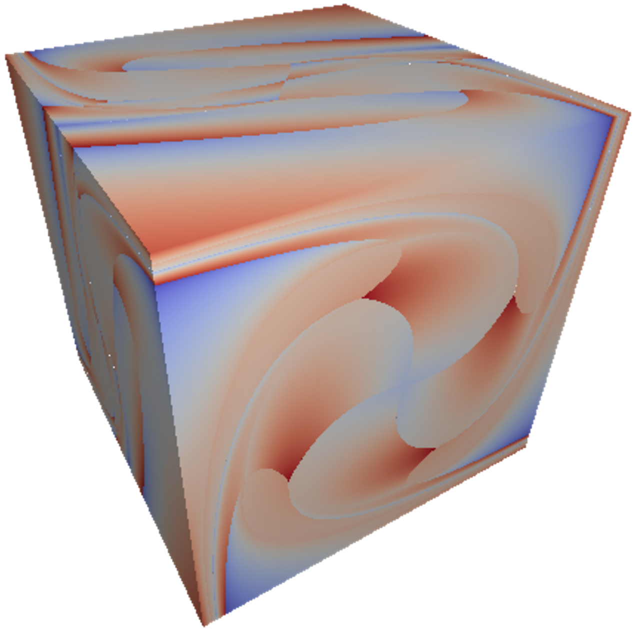
Metrics left to right: data entropy, depth entropy, max depth, projected area

Algorithms

Task-based hierarchical feature extraction algorithm based on segmented merger tree. The algorithm is implemented using a multi-runtime abstraction layer, BabelFlow, which can be used to execute arbitrary analysis and visualization dataflows using different task-based runtime systems.



Rendered/composited via Devil Ray & VTK-h



Rendered/composited via Devil Ray & BabelFlow's RadIxK algorithm

Thank you!