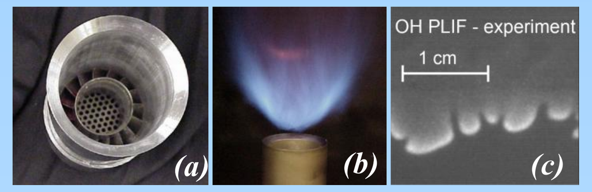
Exploring Combustion Science with SDAV's Technologies

Hasan Abbasi, Janine Bennett, Harsh Bhatia Peer-Timo Bremer, Attila Gyulassy, Scott Klasky, Kwan-Liu Ma, Manish Parashar, Valeio Pascucci, Norbert Podhorszki, Hongfeng Yu

Combustion accounts for the majority of the world's energy needs, and scientists are developing increasingly large and complex simulations to gain a better insight into clean and efficient fuels and burning devices. Visualization and analysis algorithms are integral to answering science questions about combustion; however, these algorithms must be executed concurrently with the simulations without negatively impacting their performance. We present recent results where in-situ and in-transit paradigms are used to achieve efficient topological analysis and high resolution visualizations that are well coupled with combustion simulation via high-throughput data movements that minimize any performance overhead.

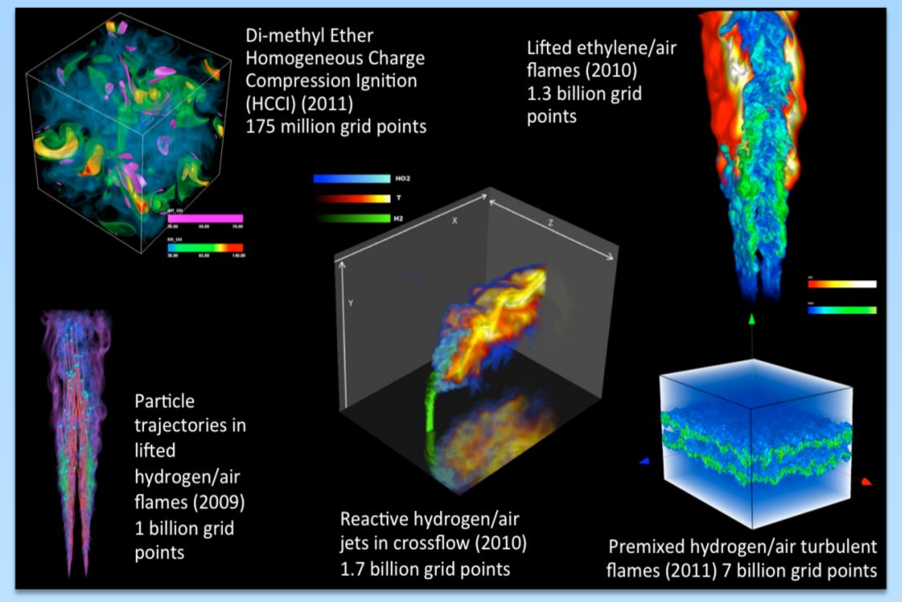
Modeling Turbulent Combustion Scenarios

- DNS codes (S3D by Jackie Chen at SNL, BoxLib by John Bell at LBNL) used to study "turbulence-chemistry" topics, such as interaction of turbulence, chemical reactions, and heat release/dilation
- First principle high-fidelity computational fluid dynamics to resolve spatial and temporal scales of turbulence
- Model various scenarios relevant to fuel efficiency
- Feature identification important to answering science questions such as locating ignition/extinction kernels



Low-swirl injectors have the potential to stabilize lean pre-mixed flames, a scenario simulated using BoxLib

Various jet scenarios, such as flame stabilization in reactive jets in crossflow, are simulated using S3D

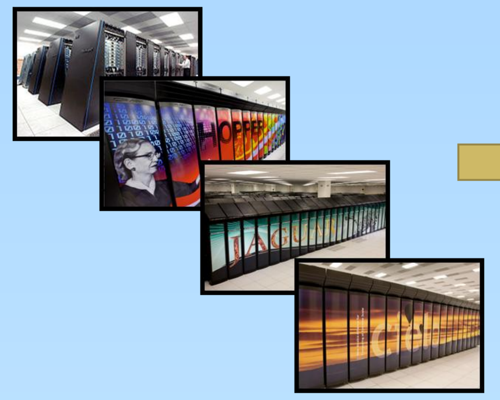


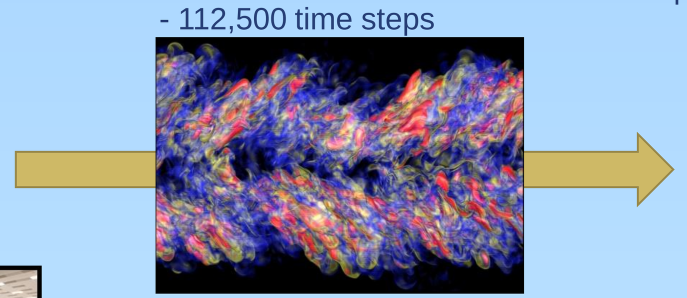
Use Case: Parallel Computation and Output to Persistent Storage

- Leadership class supercomputers are needed to resolve temporal and spatial properties of computation

- Lifted Ethylene Jet Example:
  - Run on Cray XT5 at ORNL
  - 7.5 million CPU-hrs
  - 30,000 processors
  - 112,500 time steps

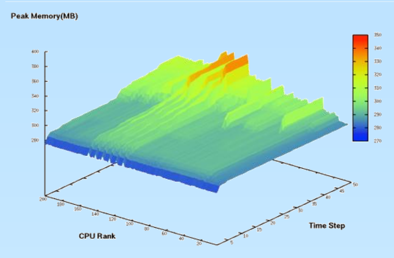
- Data stored at NERSC
- Traditionally, visualization and data analysis are performed as post-process



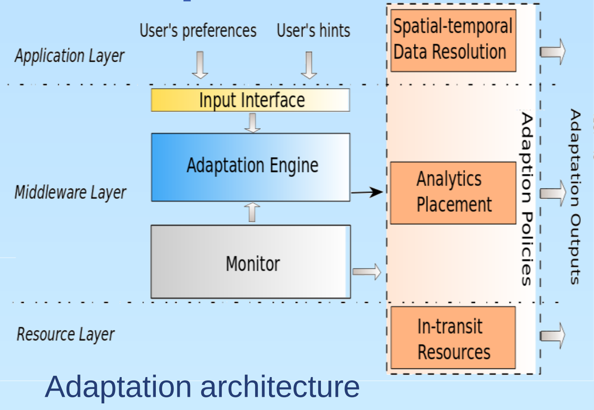


- 1.3 billion grid points
- 22 chemical species, vector & particle data
- 240TB raw field data + 50TB particle data

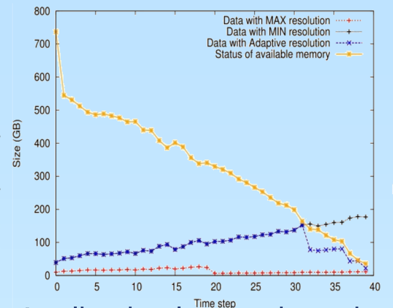
Cross-layer Adaptations for Data Management in Large Scale Coupled Scientific Workflows

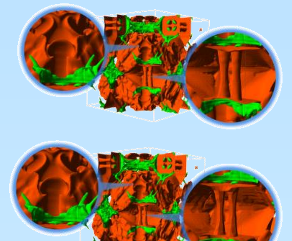


Distribution of peak memory consumption for an AMR-based Polytropic Gas simulation.



Adaptation architecture





Application layer adaptation of the spatial resolution of data using user-defined down-sampling based on runtime memory availability.

Motivation

- The AMR Advection-Diffusion simulation implements an adaptive conservative transport (advection-diffusion) solver.
- Memory and compute intensive
- Dynamic data volume and distribution
- Coupled simulation-analytics workflow based on dynamic formulations such as AMR present new challenges for in-situ/in-transit data management at extreme scale.
- Large and dynamically changing volumes of data
- Imbalanced data distributions

Solution

- Dynamic cross-layer adaptations that can respond at runtime to the dynamic data management and processing requirements
- Application layer: adaptive spatial-temporal data resolution
- Middleware layer: dynamic in-situ/in-transit placement
- Resource layer: dynamic allocation of in-transit resources
- Coordinated approaches: combine mechanisms towards a specific objective (e.g. minimized time-to-solution)

Scalable In-Memory Data Indexing and Querying

Motivation

- Researchers need to track combustion processes on the Flame front defined by range queries
- Online query needed to capture intermittent and transient information
- Overhead of loading data into traditional databases is high.
- The increasing gap between data generation and I/O constraints make it difficult to support online query-driven data analysis over large data volume.

Solution

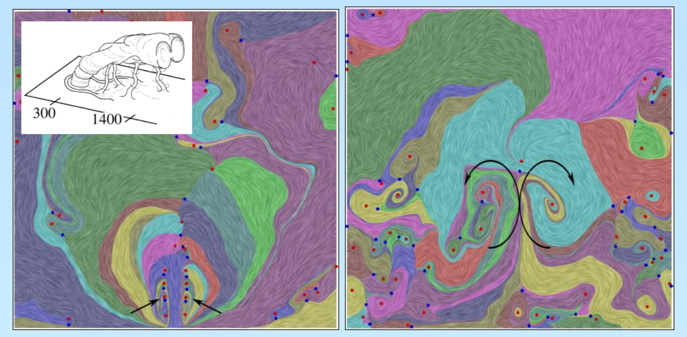
- Use memory to store both raw data and index to accelerate runtime query processing.
- Build indexes and perform queries on data-staging substrate using DataSpaces to reduce the impact on simulation.
- Achieve high concurrency using parallel bitmap indexes through non-overlapping data partition to make efficient use of large numbers of distributed many-core processors.

Results

- 35 times speed-up for query processing compare with file-based approach
- Demonstrated the scalability of our framework coupled with the S3D to perform runtime value-based querying.

Harmonic Invariant Flow Analysis

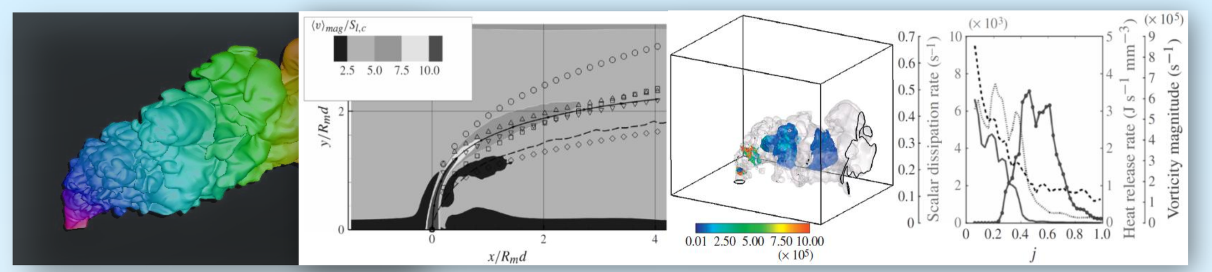
- Traditional vector field topology is not Galilean invariant
- Separate vector field into intrinsic and external flows, using Helmholtz-Hodge decomposition
- Apply Eulerian techniques to extract features
- New embarrassingly parallel algorithm



Topological analysis of the harmonic invariant flow shows the presence of jet-shear vortices.

Jet-Based Coordinates Systems

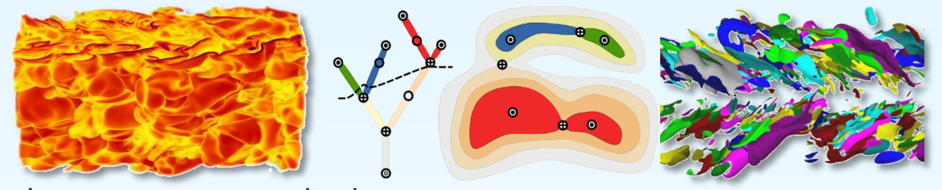
- Need for a stable coordinate system to parameterize jet in cross-flow
- Extract isosurface of mixture fraction
- Find largest components, re-mesh interior and solve Laplacian
- Center of mass of isosurfaces gives center line



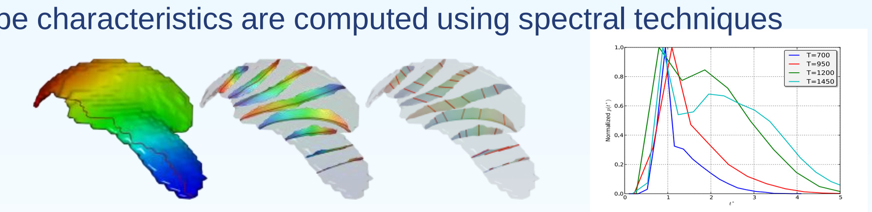
Comparison of center line approaches | Conditioned heat release rate | Vorticity magnitude vs. parameter

Shape characterization of Scalar Dissipation Rate

- Turbulent mixing is characterized locally by scalar dissipation rate χ
- Length and thickness of locally high χ correlated with length scales of turbulence

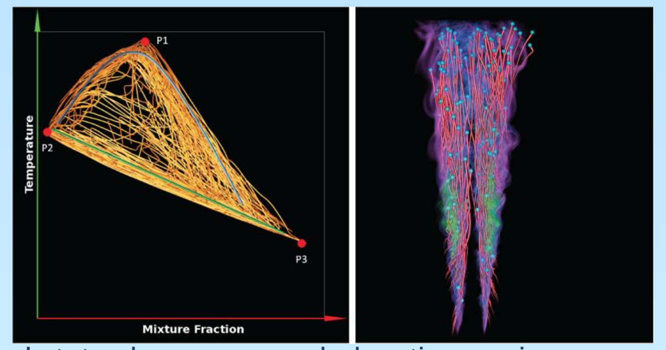


- Local structures extracted using merge trees
- Shape characteristics are computed using spectral techniques



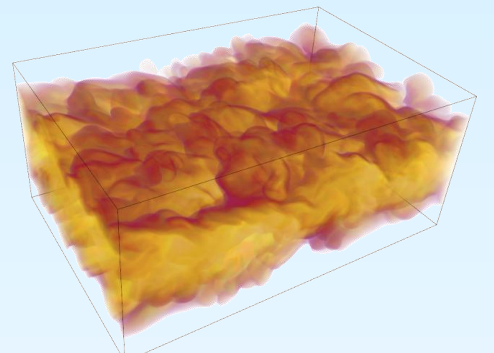
Dual Space Analysis of Particle Data

- Advanced combustion simulations are instrumented with particles to capture the dynamic behavior of turbulent flames.
- A new analysis method incorporating domain knowledge for semi-supervised learning enables studying particle thermo-chemical state changes recorded as time series curves in the phase space and the corresponding particle trajectories in the physical space of simulation.
- The classified curves correlating OH and the mixture fraction. Particles with distinct patterns of evolution curves traverse the physical space differently.

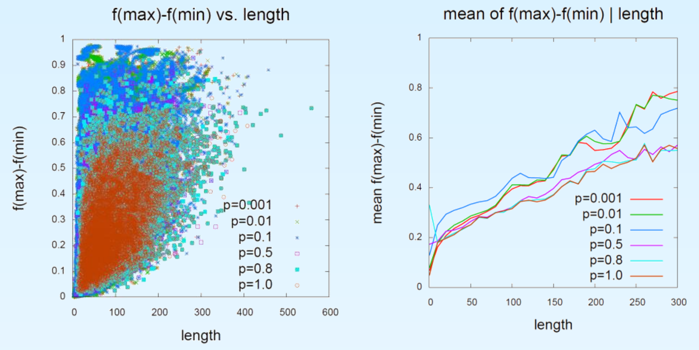


Computing Turbulence Length Scales

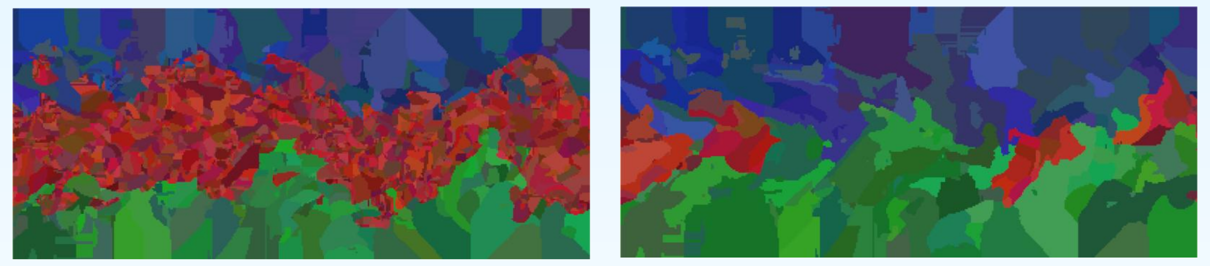
- Dissipation elements have been used to characterize length scales of turbulence
- Defined as set of points whose integral lines terminate at same min/max pair
- New approach computes DE as cells of the Morse-Smale complex
- New ability to consider the affects of perturbation and instability in the field
- Show that the measure itself is unstable in certain configurations



A jet of fuel between two planes of oxidizer evolve temporally



The characteristic length scales relating to mixture fraction is shown to be unstable

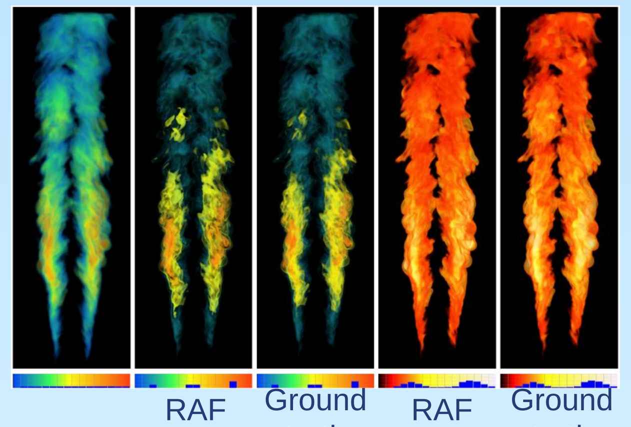


Identification of the top and bottom laminar zones and the middle turbulence zone is heavily influenced by perturbation. Numerical noise (left) is removed (right), a small perturbation having drastic effects on the segmentation.

Interactive Systems for Visualization, Parameter Exploration, and Feature Tracking

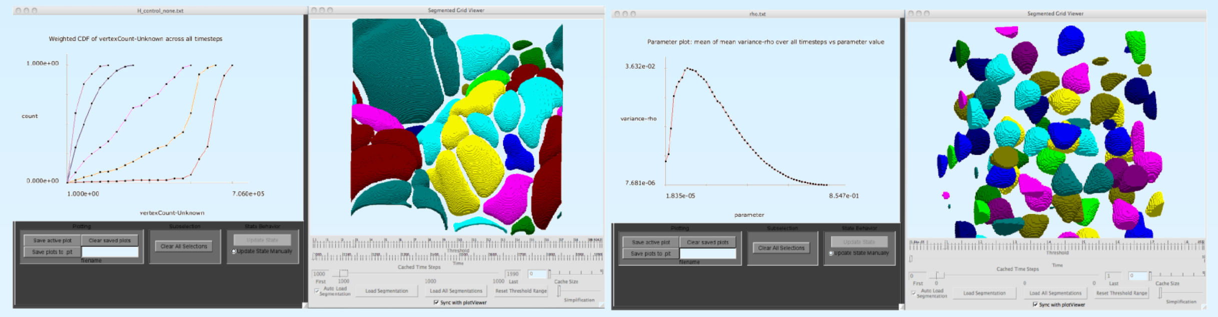
Exploratory Visualization

- Traditional in-situ visualization uses a fixed transfer function
- Computing Ray Attenuation Functions in-situ allows later exploratory visualization
- Interactive modification of transfer functions without re-loading data
- Compression reduces I/O cost of RAF



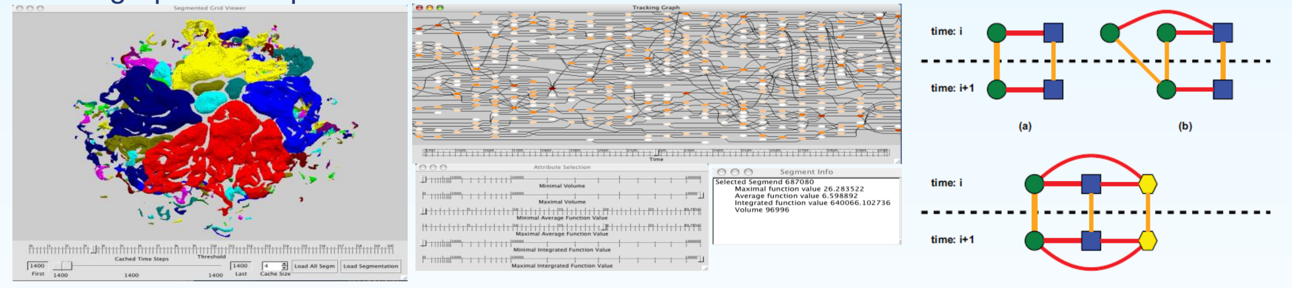
Parameter Exploration for Feature-Based Statistics

- Pre-compute feature families using topological or other segmentation techniques
- Features combine hierarchically
- Each feature has attributes and segmentations associated with them
- Combine hierarchical features interactively and aggregate statistics
- Plots of species distribution, time-series, and parameter studies



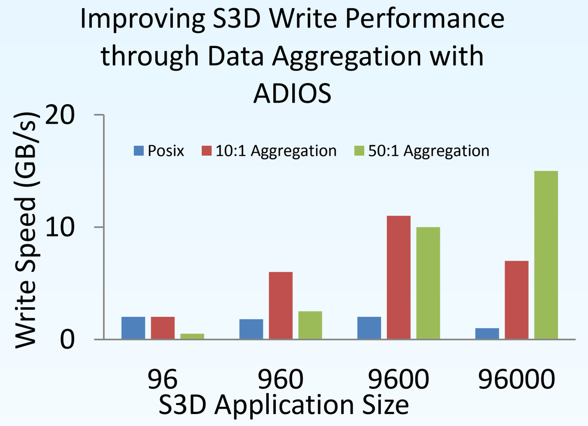
Tracking Graphs

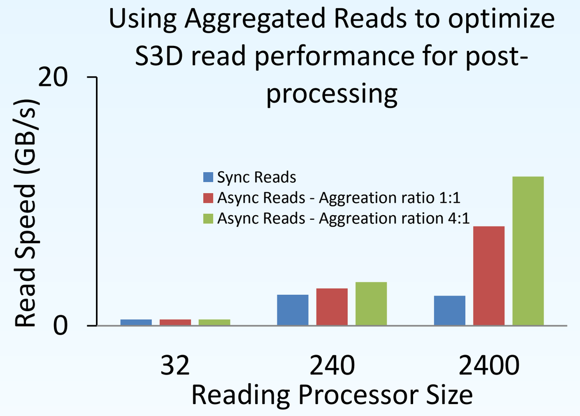
- Track the evolution (creation, merging, splitting, death) of features
- Pre-computed feature hierarchies
- Interactive selection of features, linked with segmentation viewer
- Interactive modification of feature definition parameters
- New extension to multiple fields using attribute relational graphs
- New ability to match known multi-attribute events (dependent split/merge) using subgraph-isomorphism in ARG



Improving S3D Large-scale I/O with ADIOS

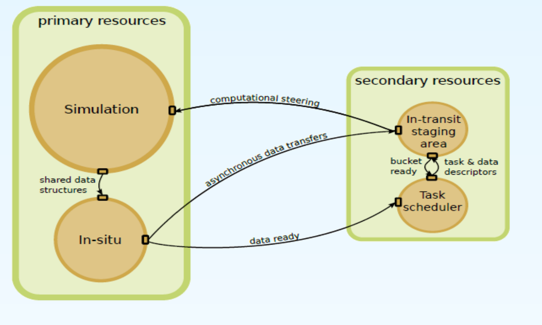
- ADIOS aggregation method provides very large performance improvements for write operations compared to standard POSIX approach
- Flexibility in selecting aggregation ratio can further improve performance
- Even better scalability can be achieved by threading metadata operations
- Similar method for read performance also dramatically improves read throughput for post-processing data



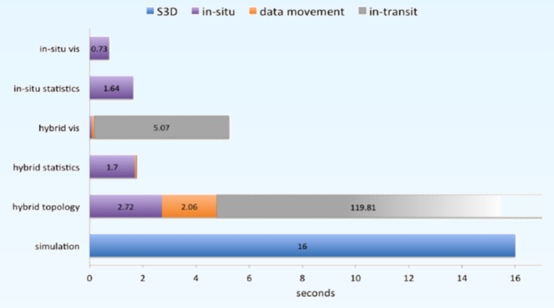


Improving temporal resolution of visualization and analysis through hybrid staging

- Combustion codes output 1/400 time steps to manage output size and I/O time
- Short-lived, fast moving, or small features are well resolved in the simulation, but can be difficult to detect and track in post-process
- Data staging allows increased analysis/visualization frequency without I/O cost
- Up to 40x speedup in analysis frequency



Using in-situ and in-transit resources allows pipelining of analysis computations



Component times for each algorithm determine resource allocation for desired frequency