Future Directions for In Situ Visualization and Analysis

Area Leads: Jim Ahrens, Wes Bethel, Scott Klasky, Kwan-Liu Ma, Valerio Pascucci, Rob Ross

Project Overview
- Aims at set of challenges in enabling scientific knowledge discovery within context of in situ processing at extreme-scale concurrency.
- Motivated by widening gap between FLOPS and I/O capacity.
- Focuses on new algorithms for analysis and visualization suitable for in situ use aimed at scientific discovery in several application areas of interest to DOE.
- Involves efforts on several leading in situ infrastructures, tackling research questions germane to new algorithms at scale across diversity of existing in situ implementations.

Problem Statement
- Lost science, due to widening gap between our ability to compute and save data.
- As a relatively new idea, want to promote quality in in situ infrastructure suitable for use at extreme-scale.
- In many cases, data analysis and visualization (DAV) algorithms need to be completely redesigned for use in situ.

Approach
- Algorithmic R&D for in situ SAV algorithms.
- Scalability and optimization for use at extreme scale and on emerging architectures.
- Direct interactions with science stakeholders and broad community engagement.

Technique: Hybrid Staging with ADIOS
- Use compute and deep-memory hierarchies to optimize overall workflow for power vs. performance tradeoffs.
- Abstract complex/deep memory hierarchy access.
- Placement of analysis and visualization tasks in a complex system.
- Impact of network data movement compared to memory movement.

Technique: Improving Flow Analysis with In Situ
- Flow analysis is typically done using the Eulerian frame of reference, we developed a technique from the perspective of the Lagrangian frame of reference.
- In situ extraction of particle trajectories, with post hoc exploration of a flow field via interplay between trajectories.
- Impact: This method is faster, more accurate, and takes less storage.
- Speed: Interaction up to 80X faster.

Technique: In Situ Topological Analysis
- Precise mathematical language to define a wide range of: Level-set and/or gradient based features, local thresholds, confidences etc.
- Efficient, combinatorial algorithms.

Science Story: Global Magnetospheric Simulations
- Problem: Poor spatial resolution is primary obstacle in understanding magnetic reconnection processes.
- Post hoc analysis results in (i) prohibitive I/O costs; (ii) "missing important science" since not all time steps are available (iii) poor time resolution limits understanding and analysis of dynamics.
- Approach: Reduce the I/O costs by identifying the magnetosphere, and applying feature detection algorithms in situ. Compact representations will be generated greatly increasing the temporal resolution of data and reducing disk I/O constraints.
- Results and Impact: Enables high resolution analysis of evolution and dynamics of the magnetosphere and its features such as magnetic reconnection, flux transfer events, and Kelvin Helmholtz generated events.

Science Story: Cosmology (Nyx)
- Problem: Post hoc halo finding in cosmology simulations, i.e., identifying stars, galaxies, etc. (i) Results not available to influence physical simulation model; (ii) "Miss important science" in time steps not available for analysis.
- Approach: In situ halo finding in Nyx (BoxLib) code (BoxLib).
- Topological methods saved reduced data representation to disk while supporting interactive exploration by changing parameters after simulation run. Use identified halos to set simulation boundary conditions or change physics models.
- Results and Impact: Enable use of halo detection results within simulation, necessary for meaningful simulations. Support analysis on greatly reduced data representation. Avoid "missing important events/science."

Science Story: Fusion
- Problem: Identify transient features known as "blobs" that could cause disruption of confinement of fusion plasma.
- Approach: Perform in situ feature extraction and in transit analysis for overall analysis time.
- Results and Impacts: Developed a feature extraction algorithm based on an outlier detection approach that provided significant time savings. Developed an approach at SC'14 conference. Enables scientists to accelerate the detection of spatio-temporal features in streaming data that can be described with conjunctive or disjunctive range predicate style conditions.

Science Story: 3D Structure of Block Co-Polymers
- Problem: Determine 3D structure of block co-polymers. Scattering data will provide insight but analysis tools need to be developed to interpret the experimental data (e.g., angular, spatial). New approach will enable on-the-fly analysis to study the evolution of 3D structure.
- Approach: Develop methods to use 3D data from MD simulations to study scattering. Compare with experimental data, iterate to modify parameters until scattering pattern matches.
- Results and Impact: Enable scientists to study 3D structure as a function of materials and processing parameters.

Science Story: Turbulent Mixing and Combustion
- Problem: The LES/EDD code is used in design studies for gas turbine combustors and rocket engines. The simulations solve large eddy simulations of complete real geometry with chemistry and Lagrangian droplets for liquid to gas phase changes and combustion.
- Approach: Implement in situ including in transit methods, use feature detection methodologies to automatically explore salient flow features such as flame fronts and turbulent vortex cores.
- Results and Impact: Enable scientists and engineers to more readily analyze the effect of system design variations such as mixer shapes and system pressures resulting in increased throughput of the total number of simulations achievable for any given design study.

Science Story: Global Magnetospheric Simulations
- Problem: Poor spatial resolution is primary obstacle in understanding magnetic reconnection processes.
- Post hoc analysis results in (i) prohibitive I/O costs; (ii) "missing important science" since not all time steps are available (iii) poor time resolution limits understanding and analysis of dynamics.
- Approach: Reduce the I/O costs by identifying the magnetosphere, and applying feature detection algorithms in situ. Compact representations will be generated greatly increasing the temporal resolution of data and reducing disk I/O constraints.
- Results and Impact: Enables high resolution analysis of evolution and dynamics of the magnetosphere and its features such as magnetic reconnection, flux transfer events, and Kelvin Helmholtz generated events.

Science Story: Cosmology (Nyx)
- Problem: Post hoc halo finding in cosmology simulations, i.e., identifying stars, galaxies, etc. (i) Results not available to influence physical simulation model; (ii) "Miss important science" in time steps not available for analysis.
- Approach: In situ halo finding in Nyx (BoxLib) code (BoxLib).
- Topological methods saved reduced data representation to disk while supporting interactive exploration by changing parameters after simulation run. Use identified halos to set simulation boundary conditions or change physics models.
- Results and Impact: Enable use of halo detection results within simulation, necessary for meaningful simulations. Support analysis on greatly reduced data representation. Avoid "missing important events/science."

Science Story: Fusion
- Problem: Identify transient features known as "blobs" that could cause disruption of confinement of fusion plasma.
- Approach: Perform in situ feature extraction and in transit analysis for overall analysis time.
- Results and Impacts: Developed a feature extraction algorithm based on an outlier detection approach that provided significant time savings. Developed an approach at SC'14 conference. Enables scientists to accelerate the detection of spatio-temporal features in streaming data that can be described with conjunctive or disjunctive range predicate style conditions.

Science Story: 3D Structure of Block Co-Polymers
- Problem: Determine 3D structure of block co-polymers. Scattering data will provide insight but analysis tools need to be developed to interpret the experimental data (e.g., angular, spatial). New approach will enable on-the-fly analysis to study the evolution of 3D structure.
- Approach: Develop methods to use 3D data from MD simulations to study scattering. Compare with experimental data, iterate to modify parameters until scattering pattern matches.
- Results and Impact: Enable scientists to study 3D structure as a function of materials and processing parameters.

Science Story: Turbulent Mixing and Combustion
- Problem: The LES/EDD code is used in design studies for gas turbine combustors and rocket engines. The simulations solve large eddy simulations of complete real geometry with chemistry and Lagrangian droplets for liquid to gas phase changes and combustion.
- Approach: Implement in situ including in transit methods, use feature detection methodologies to automatically explore salient flow features such as flame fronts and turbulent vortex cores.
- Results and Impact: Enable scientists and engineers to more readily analyze the effect of system design variations such as mixer shapes and system pressures resulting in increased throughput of the total number of simulations achievable for any given design study.