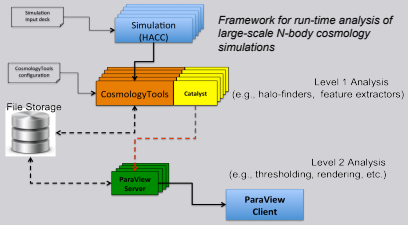
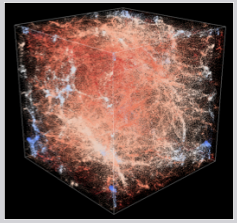
The Visual and Numerical Analysis of Large-Scale Structures in Cosmology and Astrophysics

Objective

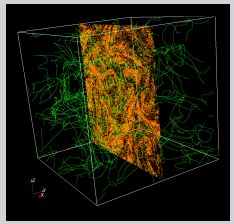
The visual and statistical identification of large-scale structures--halos, voids, magnetic reconnections, and flame surfaces--enables simulations to be correlated with telescope surveys, explain the distribution of galaxies in the universe, and infer properties of dark energy and dark matter. We present a set of tools used in situ, coprocessing, and postprocessing for scientists to locate, track, and evaluate the evolution of such structures. Techniques include run-time data analysis, visualization, and data management.



Framework for run-time analysis of large-scale N-body cosmology simulations



Halos, voids, filaments, and walls are visible in any subdomain of this cosmology simulation. Quantitative analysis, visualization, and data management of such structures are objectives of this work.

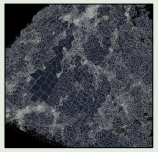


Simulations of magnetic reconnection and turbulent magnetohydrodynamics are used to better understand the structures and evolution of the magnetic field during astrophysical events. Informative visualization of these structures is another objective of this work.

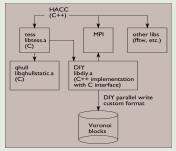
Data Analysis

Voronoi Tessellation

A polyhedral Voronoi mesh is computed in situ from dark matter tracer particles. This continuous density field is used to detect voids and other features.



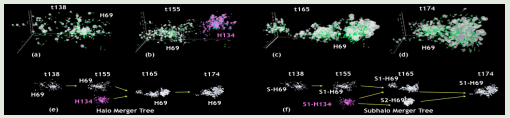
Voronoi tessellation of dark matter tracer particles



In situ analysis software structure

Merger Trees

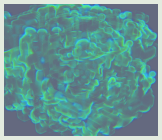
Temporal feature evolution is tracked using merger trees. Computed in situ, they mark births, deaths, and mergers of halos and subhalos.



Merger tree of halos and subhalos includes births, mergers, and continuations

Multistream Structure Identification

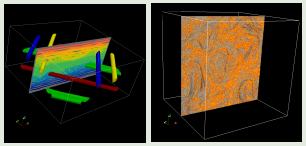
We are developing a parallel feature extractor that classifies regions based on multistream flow events.



Volume rendering of the estimated density field at a given redshift

Magnetic Reconnection Feature Detection

Magnetic reconnection is a driving phenomenon in solar and magnetospheric physics that occurs when two topologically different structures in the magnetic field merge. Poincare mapping is used to locate features. We developed a parallel implementation of the Poincare mapping technique using ParaView.

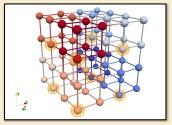


Left: The detection of flux tubes in a sample dataset. Right: the Poincare map applied to an MHD turbulence study shows structures in the magnetic field.

Data Management

Simulation Checkpoint / Restart

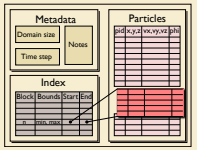
We designed and implemented an efficient parallel I/O checkpoint / restart method directly into the HACC simulation and are testing performance at full scale of BG/P and BG/Q



Topology-aware data movement for I/O. Leveraging the topology of the Blue Gene torus interconnect for smarter aggregation and improved bandwidth.

Simulation Output for Analysis

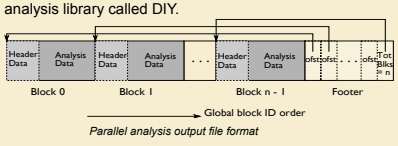
Together with NWU, we are developing a parallel netCDF I/O module to improve data management, annotations, analysis, and performance for particle data generated by HACC. It supports spatial block querying of arbitrary size regions and parallel reading with different numbers of processes.



File schema for analysis output enables spatial queries of particle data in a high-level self-describing format.

In Situ Analysis Output

Derived fields such as Voronoi tessellations are written using the parallel output capability provided by a data analysis library called DIY.

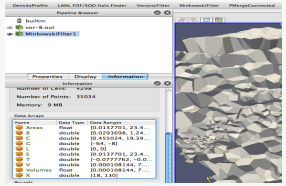


Parallel analysis output file format

Visualization

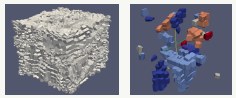
ParaView

A cosmology tools plugin and toolbar has been added to the next release of ParaView. It includes a parallel reader for Voronoi tessellations and Minkowski functionals over connected components of Voronoi cells.

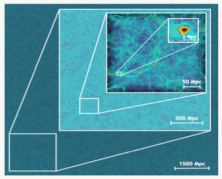


VL3

Vl3 is a parallel visual analysis framework for Blue Gene supercomputers, GPU-based clusters, and scientists' laptops. Its modular design scales to 16K GPU cores and supports interactive visual analysis and data exploration of $(6K)^3$ data volume. We are currently developing scalable out-of-core mechanisms for upcoming $(10K)^3$ data volumes.



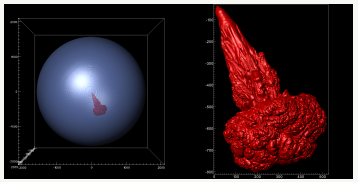
Cosmology tools plugin in ParaView promotes interactive feature exploration.



HACC simulation performed on the ALCF BG/Q supercomputer, and visualized on the ALCF Eureka cluster using vl3.

VisIt

We are using VisIt for visualization of Type Ia supernovae with Chris Malone (UCSC), Andy Nonaka (LBNL) and Stan Woosley (UCSC). The application is a 5-level CASTRO AMR simulation; the finest level is 64 times the resolution of the root level; resolution in root level 8.68 km on a side. With 1 billion cells at the finest level, 2 billion cells overall, and 2 dozen variables per cell, the dataset is 250 GB in size.



Snapshot of the evolution of the buoyant flame after about 0.6 seconds of evolution. The left panel shows the stellar surface in blue and the surface of the flame in red. The right panel is a zoom-in of the flame surface, which exhibits significant small-scale structure.

SDAV
Scalable Data Management, Analysis, and Visualization

Tom Peterka, Venkat Vishwanath, Joe Insley, Salman Habib, Katrin Heitmann (ANL), George Zagaris, Berk Geveci (Kitware), Jingyuan Wang (UTK), Jay Takle (Rutgers), Jon Woodring, Jim Ahrens (LANL), Wei-keng Liao, Saba Sehrish (NWU), Burlen Loring, Gunther Weber, Wes Bethel (LBNL)