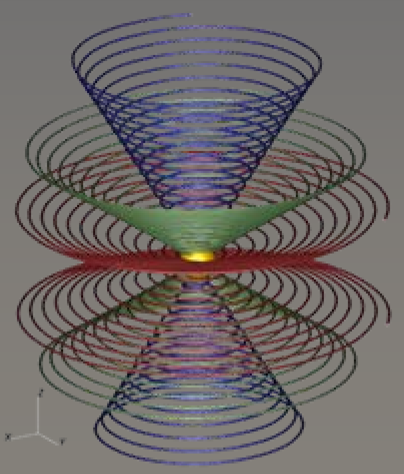
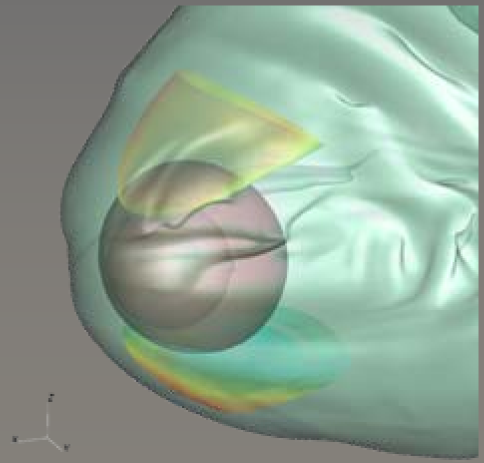
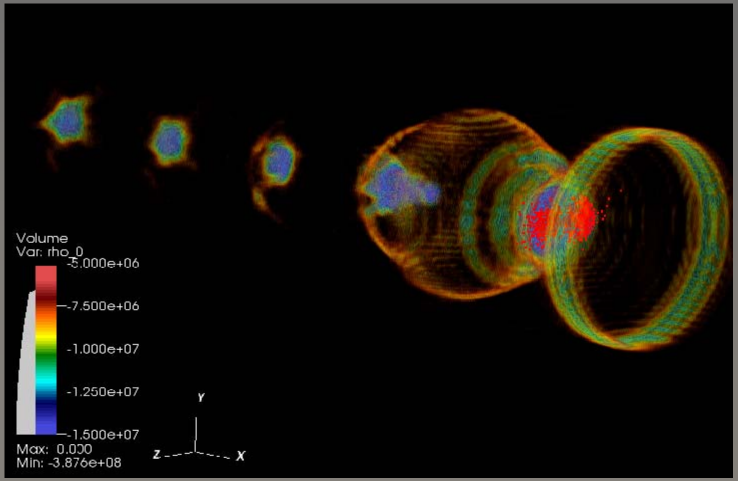
Recent Advances in VisIt: AMR Streamlines and Query-driven Visualization

Gunther H. Weber

Lawrence Berkeley National Laboratory / UC Davis

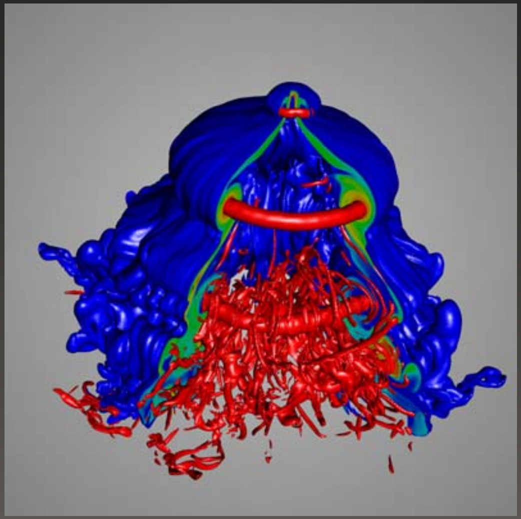




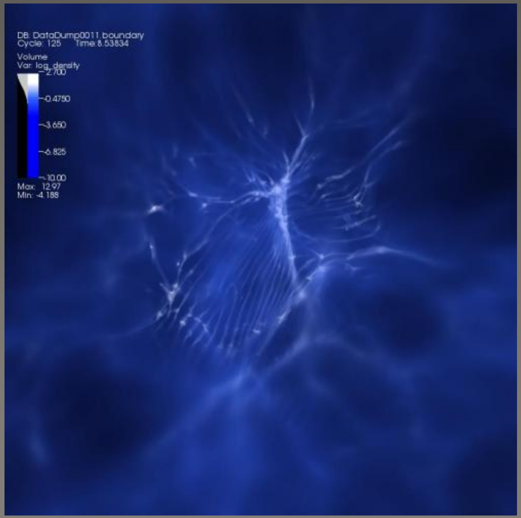


- Richly featured visualization and analysis tool for large data sets
- Built for five use cases:
  - Data exploration
  - Visual debugging
  - Quantitative analysis
  - Presentation graphics
  - Comparative analysis
- Data-parallel client server model, distribution on per patch-basis
- Use of meta-data / contracts to reduce amount of processed data

➔ Ideal basis for specialized AMR visualization tool replacement



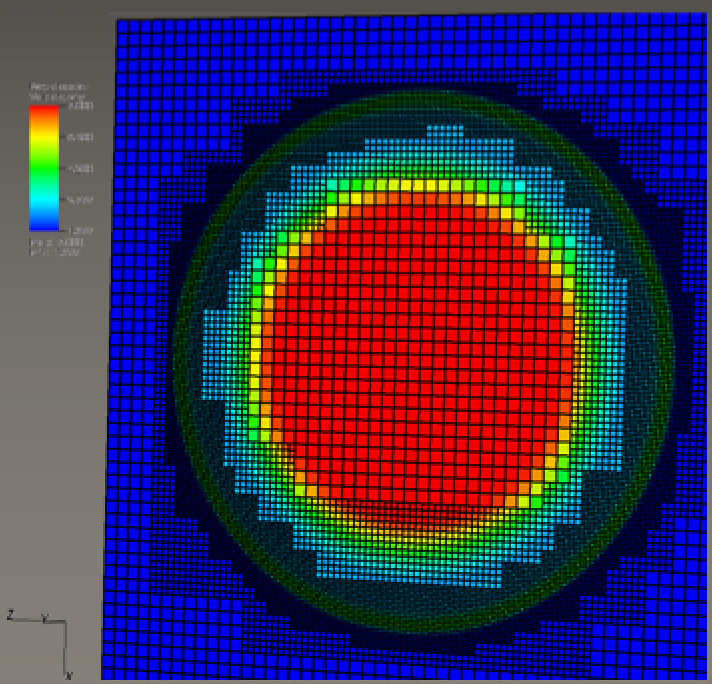
[Argon bubble subjected to shock Jeff Greenbough, LLNL]

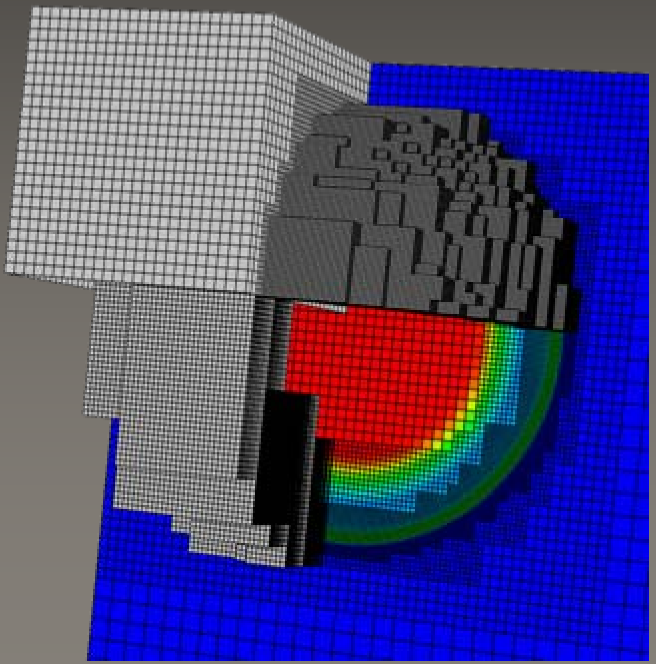


[Logarithm of gas/dust density in Enzo star/galaxy simulation, Tom Abel & Matthew Turk, Kavli Institute]

VisIt and AMR Data

- Supported as “first-class” data type
- Handled via “ghost-cells”: Coarse cells that are refined are marked “ghost” in the lower level
- Work on rectilinear grids and skip ghost cells or “remove” results produced in ghost cells later on





Streamlines for Adaptive Mesh Refinement Data

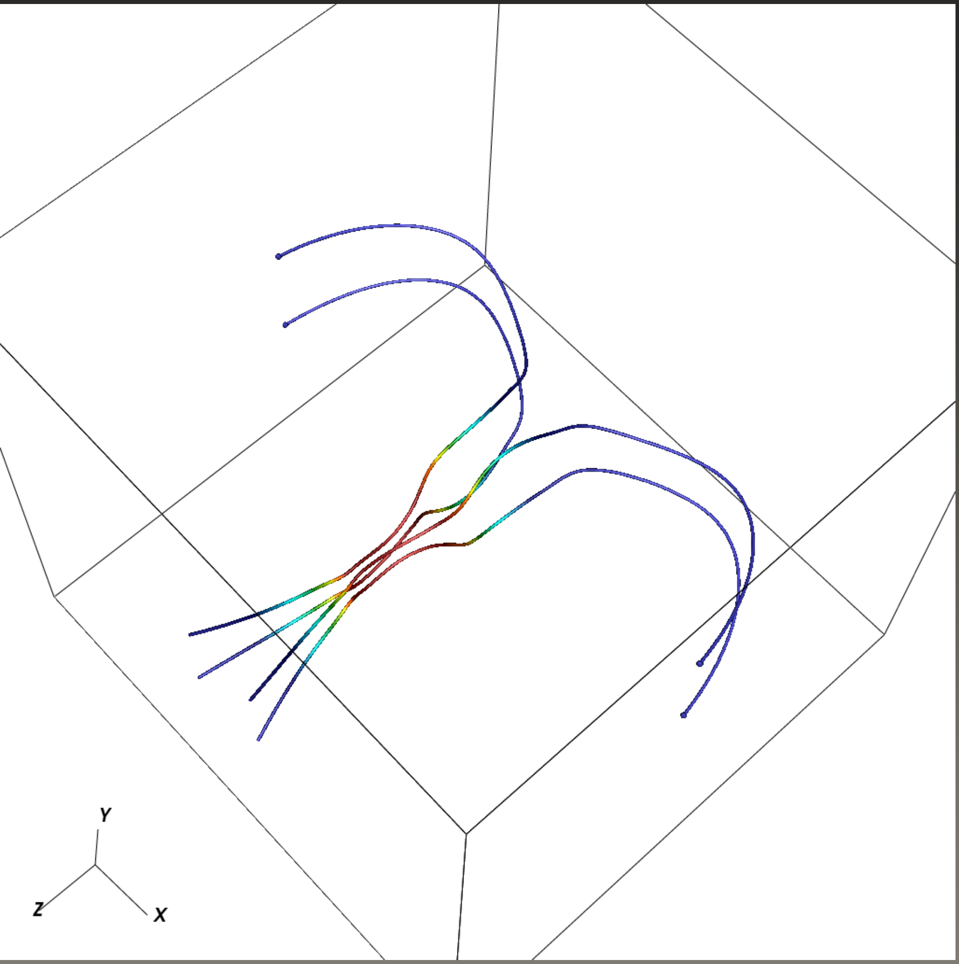
Joint work with
Eduard Deines[1], Christoph Garth[1], Brian Van Straalen[2], Sergey Borovikov[3], Daniel F. Martin[2] and Ken I. Joy[1,2]

[1] Institute for Data Analysis and Visualization, University of California, Davis
[2] Computational Research Division, Lawrence Berkeley National Laboratory
[3] Department of Physics, University of Alabama, Huntsville

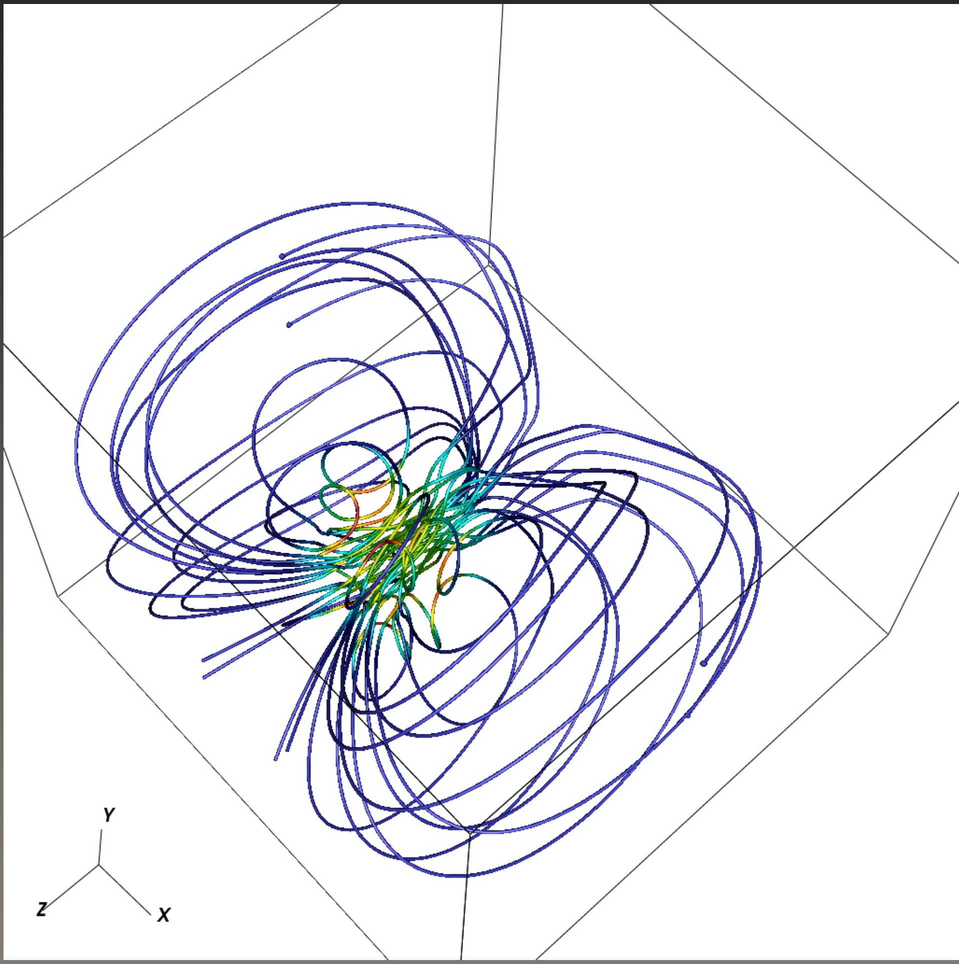
Motivation

- VisIt streamlines one year ago:
  - Streamlines cannot cross boundaries between multiple grids of multi-block data
- More recently:
  - Fully parallel implementation that "hands off" streamlines between grids [Pugmire et al, 2009]
  - Publicly available in VisIt 1.11 (enhancements in upcoming 1.12 release)
- However:
  - Streamlines still not "AMR" aware (refined cells)

Motivating Example – Vortex Core Merger



Stay in Level 0



“Descend” into AMR hierarchy

Integral Curves and Steamlines – Definition

- Integral curves in visualization
  - Streamlines
  - Streaklines
  - Pathlines, etc.

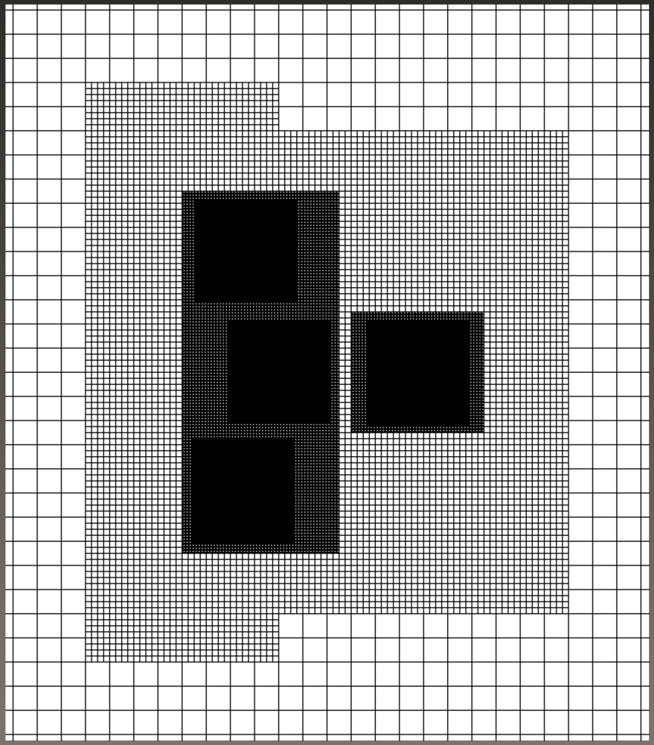
- Essential visualization tool providing intuitive understanding of flow data

- Obtained Numerical solution of ordinary differential equation

$$S(t)' = F(S(t)) \text{ or } S(t)' = F(S(t),t)$$

Challenges Posed by Block-structured AMR data sets

- Multiple refinement levels
- Several domains represented as rectilinear grids
- Data in finer levels replaces data in coarser levels
- Cell-centered data



Proper Handling of AMR Hierarchies (1/2)

1. Evaluate vector field at each integration step in finest available level:

   – Discontinuity handling by the integration method

   – Higher number of integration steps

   – Higher error after passing the discontinuity (level boundary)

   – All domains needed during the integration

Proper Handling of AMR Hierarchies (2/2)

2. Stop at domain boundaries and restart computation in next domain:

   – Reduces error

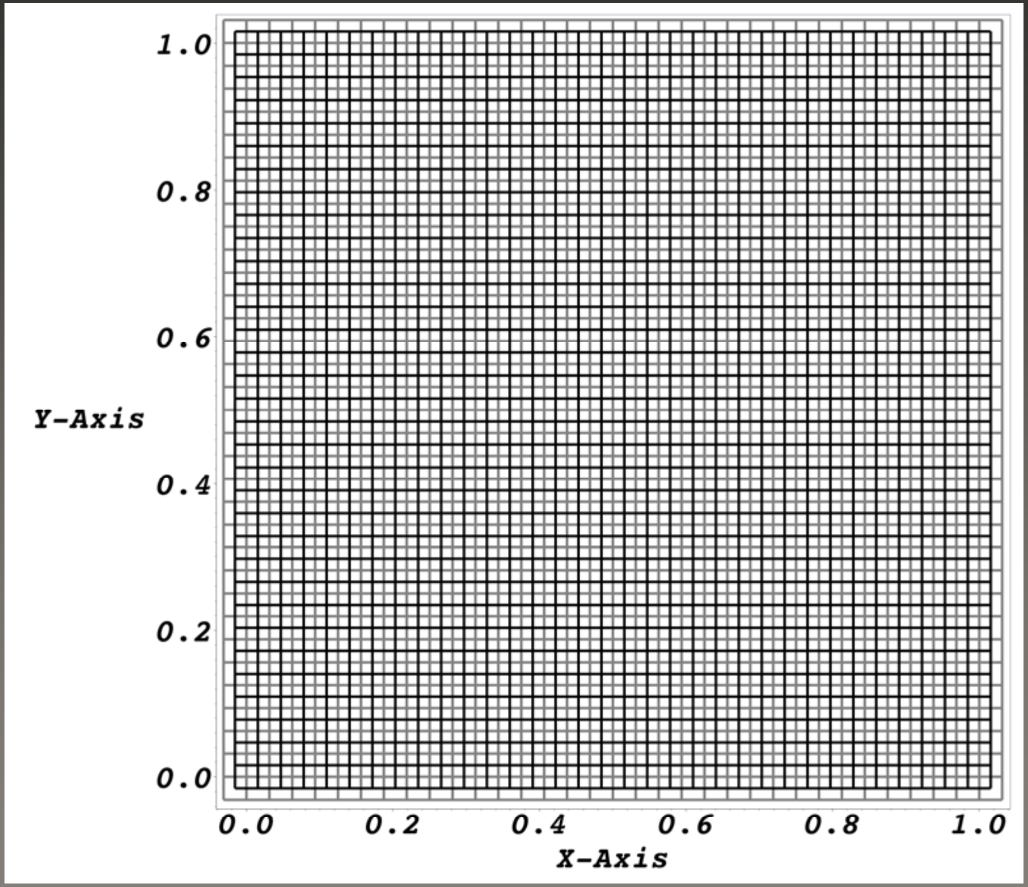
   – Reduces number of integration steps

   – Process curve integration in each domain separately (parallel processing possible)

   – Intersection point calculation with domain bounding box using Newton method

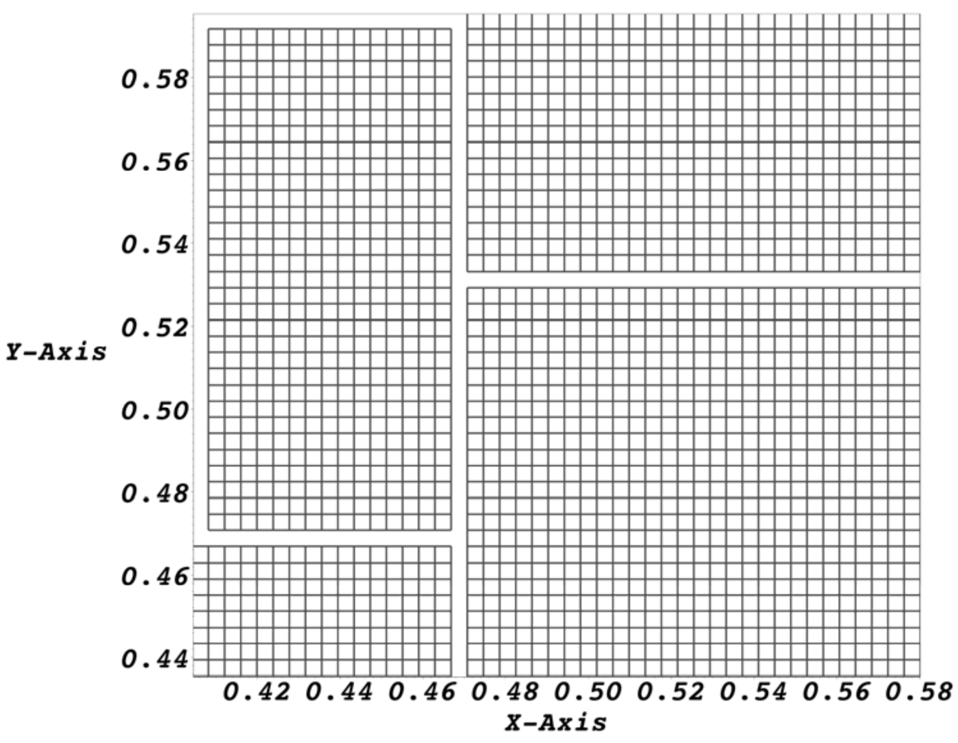
Proper Handling of Cell-centered Data

Dual-mesh representation

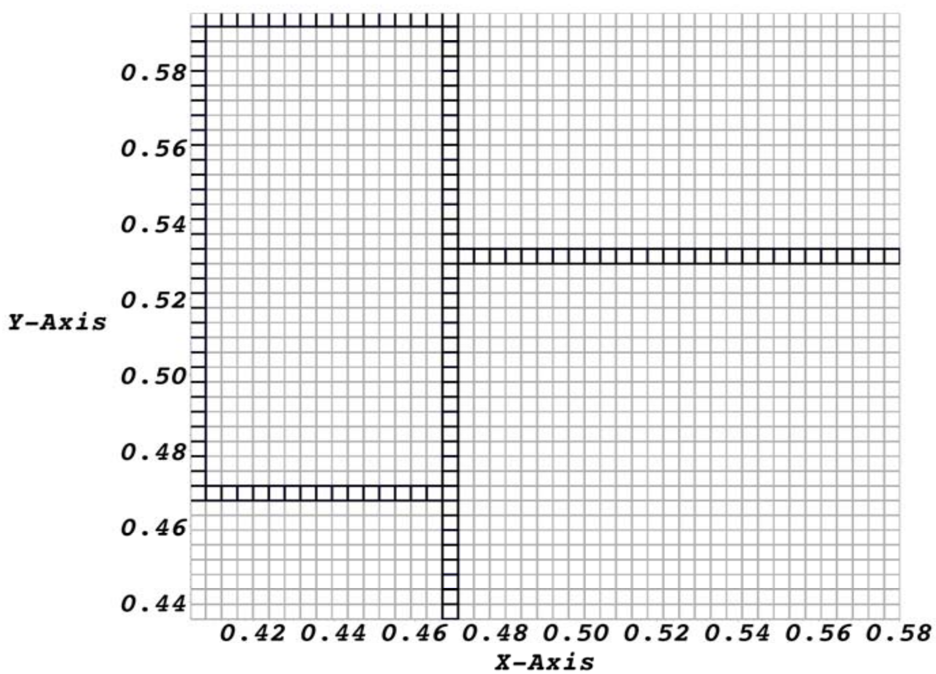


Proper handling of cell-centered data

“Gaps” between domains



Dual grid using “ghost” cells



Algorithm

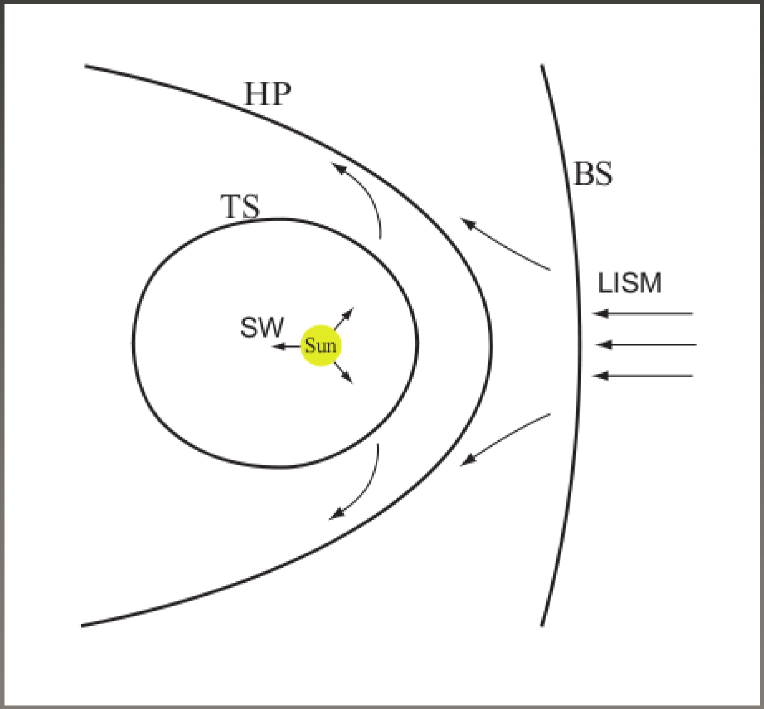
- Start in domain in finest possible level
- Build dual mesh
- Advance integration step
- If step inside nested domains (finer level)
  - Intersect with the bounding box of the finer domain
  - Restart the algorithm inside the finer domain
- If outside domain
  - Intersect with the domain bounding box
  - Restart in the next domain

Solar System Simulation

- Interaction of solar wind with interstellar medium

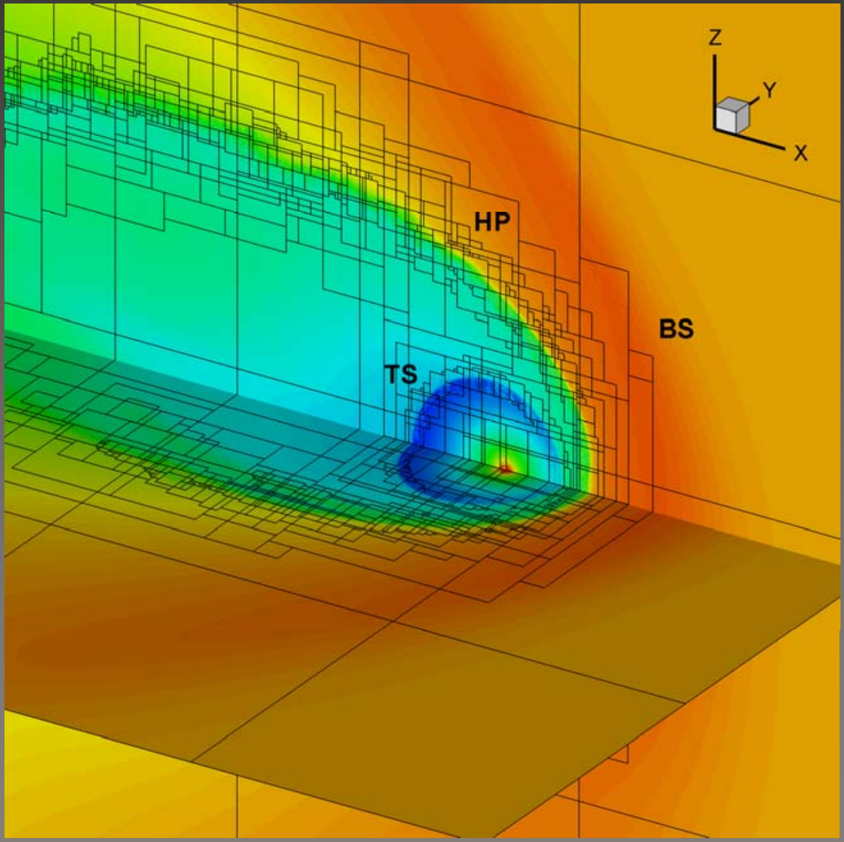
- Termination shock
  - First boundary of solar system
  - End of strong sphere of influence of solar wind

- Heliopause
  - Collision of solar wind with local interstellar medium
  - End of direct sphere of influence of solar wind

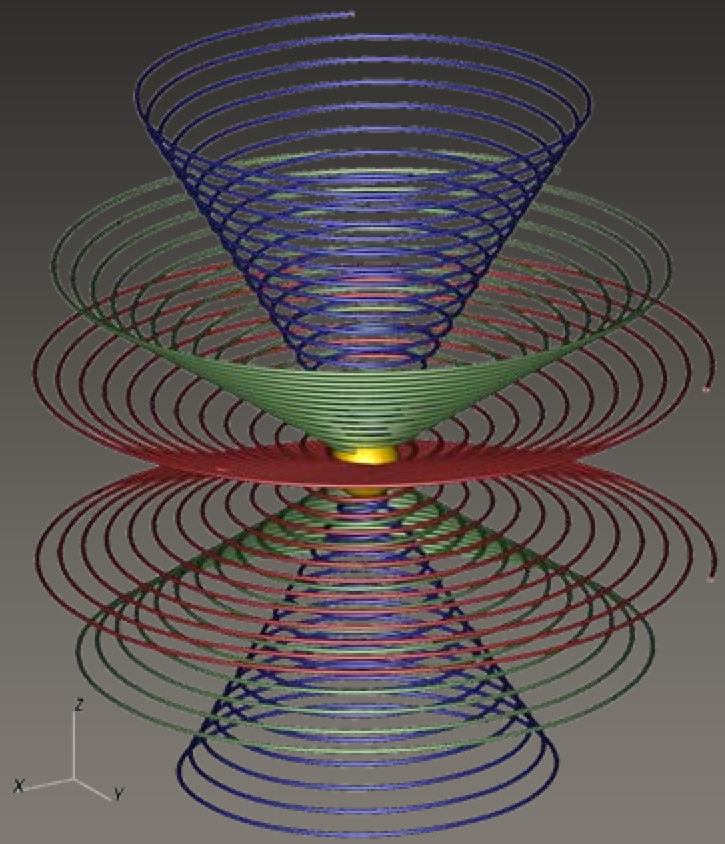


Solar System Simulation

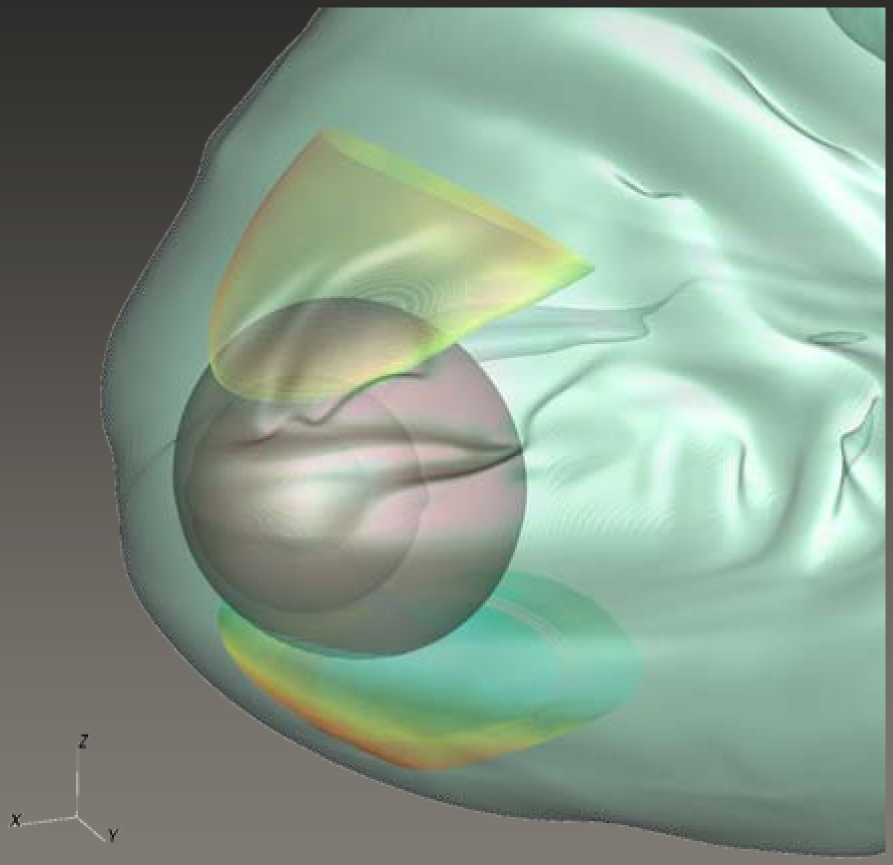
- Computational region about 1000 AU
- Plasma fluctuations 0.01 AU
- To fine to be modeled without AMR
- AMR Mesh
  - Five refinement levels
  - 20037 domains



Interplanetary Magnetic Field Lines



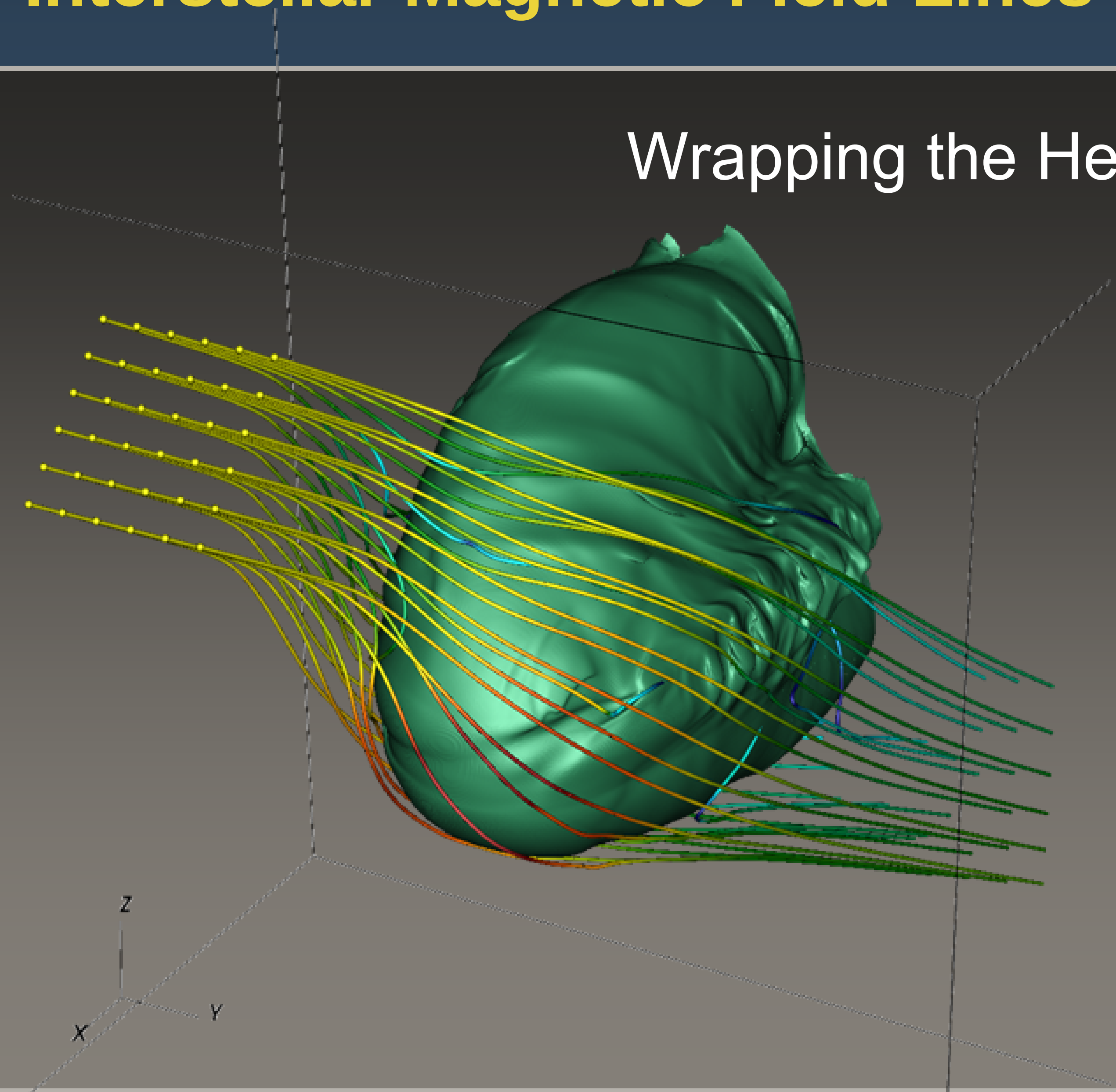
Parker Spiral



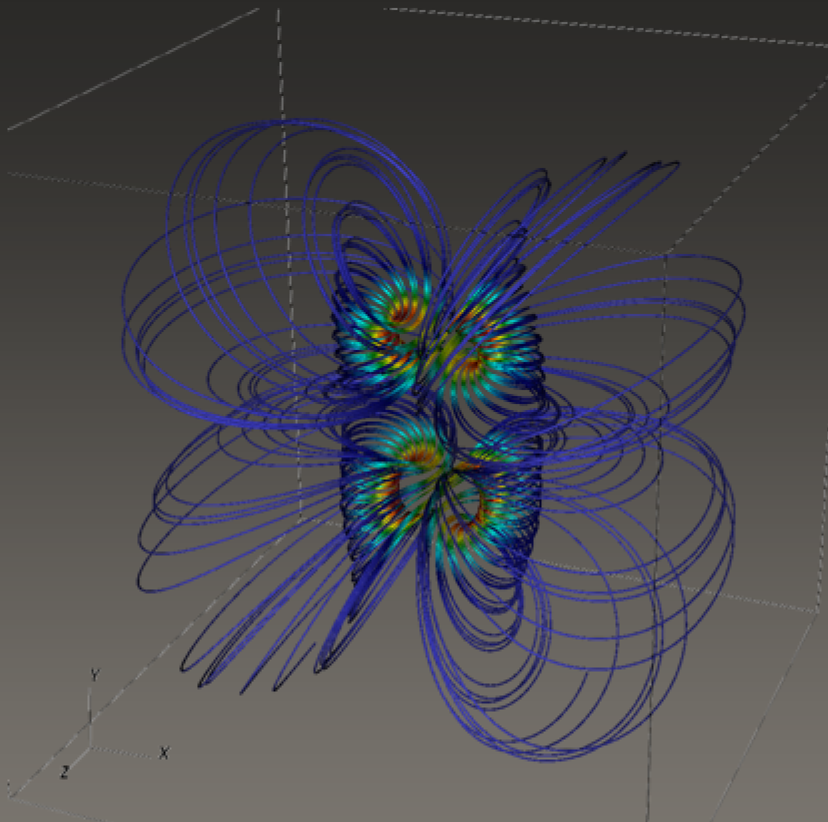
Past the Termination Shock

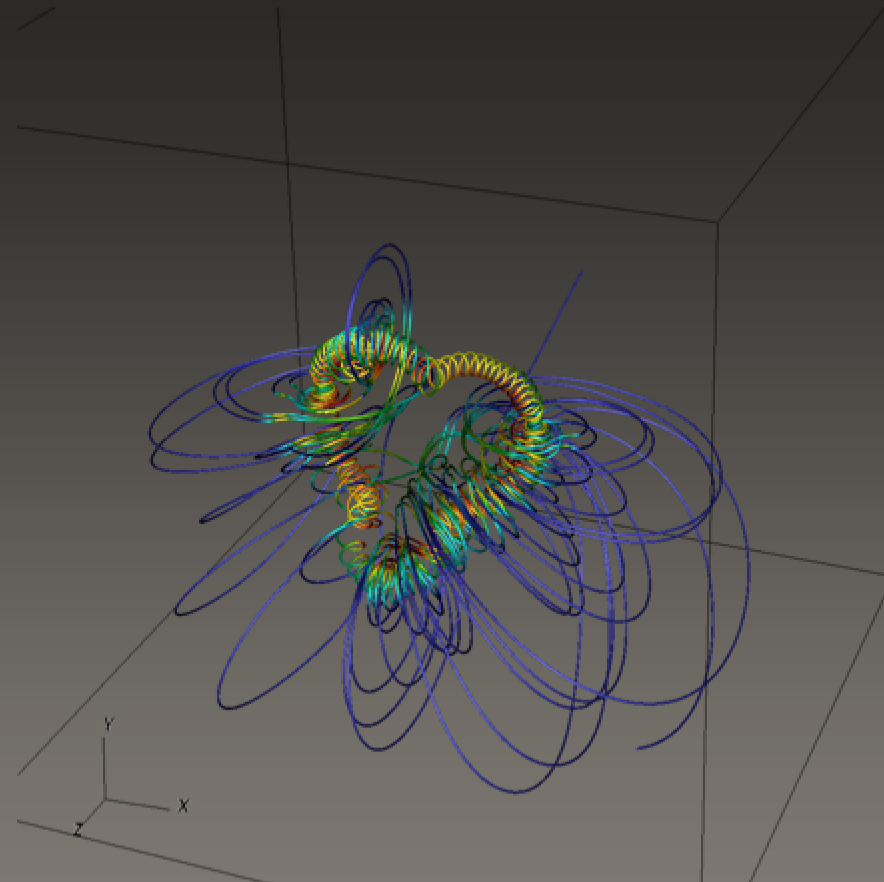
Interstellar Magnetic Field Lines

Wrapping the Heliopause



Two Incompressible Viscous Vortex Cores





Two distinct vortex rings   Merged into a single flow structure

Future work

- Distribute in future VisIt version
- Proper handling of:
  - Embedded boundaries
  - Mapped grids
- Parallelize
- Pathlines, etc.

Analysis of Large-Scale Laser Wakefield Particle Acceleration Simulation

Joint work with

Oliver Rübel[1,2,3], Prabhat[1], Kesheng Wu[1], Hank Childs[1], Jeremy Meredith[4], Cameron G. R. Geddes[5], Estelle Cormier-Michel[5], Sean Ahern[4], Peter Messmer[6], Hans Hagen[2], Bernd Hamann[3,2,1] and E. Wes Bethel[1,3]

[1] Computational Research Division, Lawrence Berkeley National Laboratory
[2] Internatl. Research Training Group 1131, TU Kaiserslautern, Germany
[3] Institute for Data Analysis and Visualization, University of California, Davis
[4] Oak Ridge National Laboratory
[5] LOASIS program of Lawrence Berkeley National Laboratory
[6] Tech-X Corporation

Motivation:
Laser Wakefield Particle Acceleration

[Image courtesy of http://worldwakesurfingchampionships.com]

- **Advantages:**
  Can achieve electric fields thousands of times stronger than in conventional accelerators

➔ Can achieve high acceleration over very short distance.

Laser Wakefield Particle Acceleration

Simulation

- Performed over 2D and 3D domains using the VORPAL code
- Simulations restricted to window covering only a plasma subset in x direction in beam vicinity
- Simulation window moves along local x axis
- Produces particle and field data (at typically 40-100 timesteps)

Particle data

- Scattered data with particle location, momentum and identifier

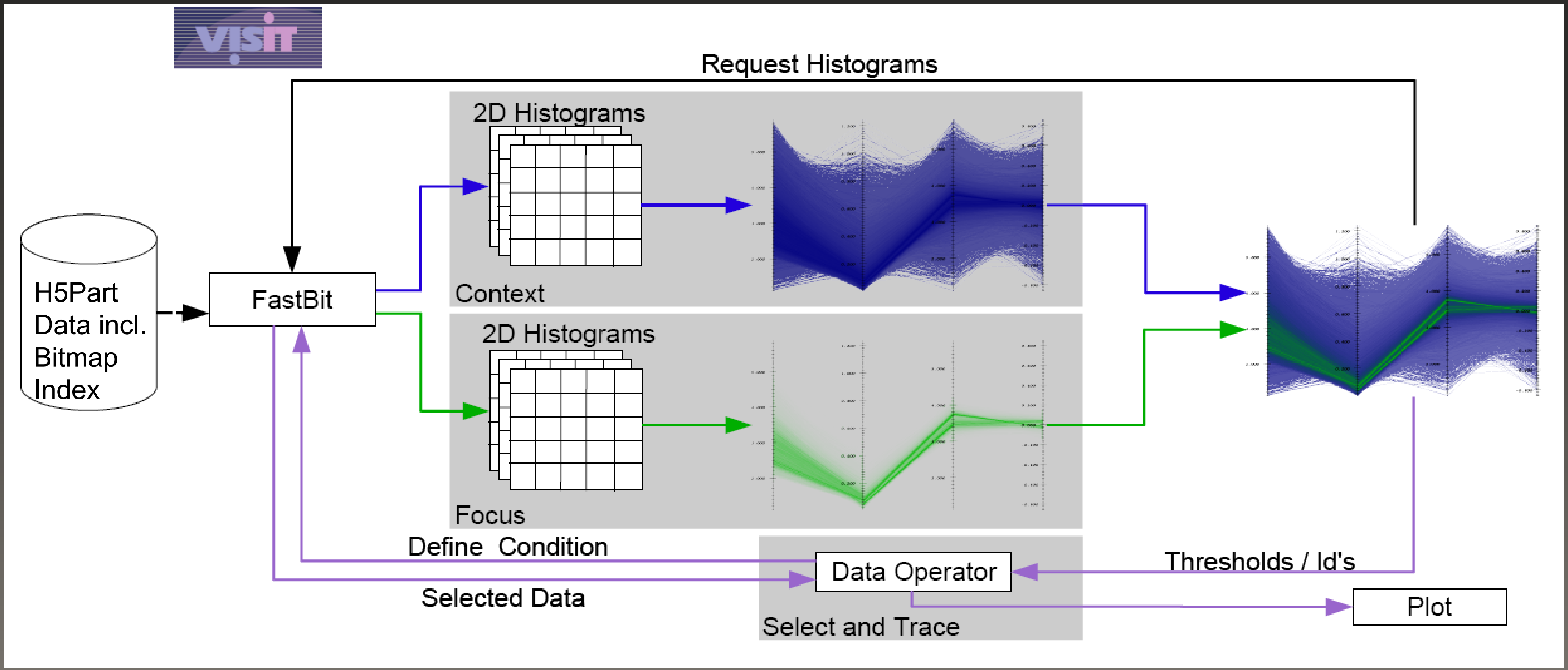
  ~ $0.4*10^6$ – $30*10^6$ (in 2D) and ~$80*10^6$ – **$200 *10^6$** (in 3D) per time step

  ➔ Total size: ~1.5GB – >30GB (in 2D) and ~100GB – **>1TB** (in 3D)

Field data

- Electric field, magnetic field, and RhoJ (regular grid)
  - Resolution: Typically ~0.02-0.03μm longitudinally, and ~ 0.1-0.2μm transversely
  - Total size: ~3.5GB - >70GB (in 2D) and ~200GB - >2TB (in 3D)

System Design



References:

- VisIt is available from https://wci.llnl.gov/codes/visit/
- FastBit is available from https://codeforge.lbl.gov/projects/fastbit
- H5Part is available from http://h5part.web.psi.ch/ or http://vis.lbl.gov/Research/AcceleratorSAPP/

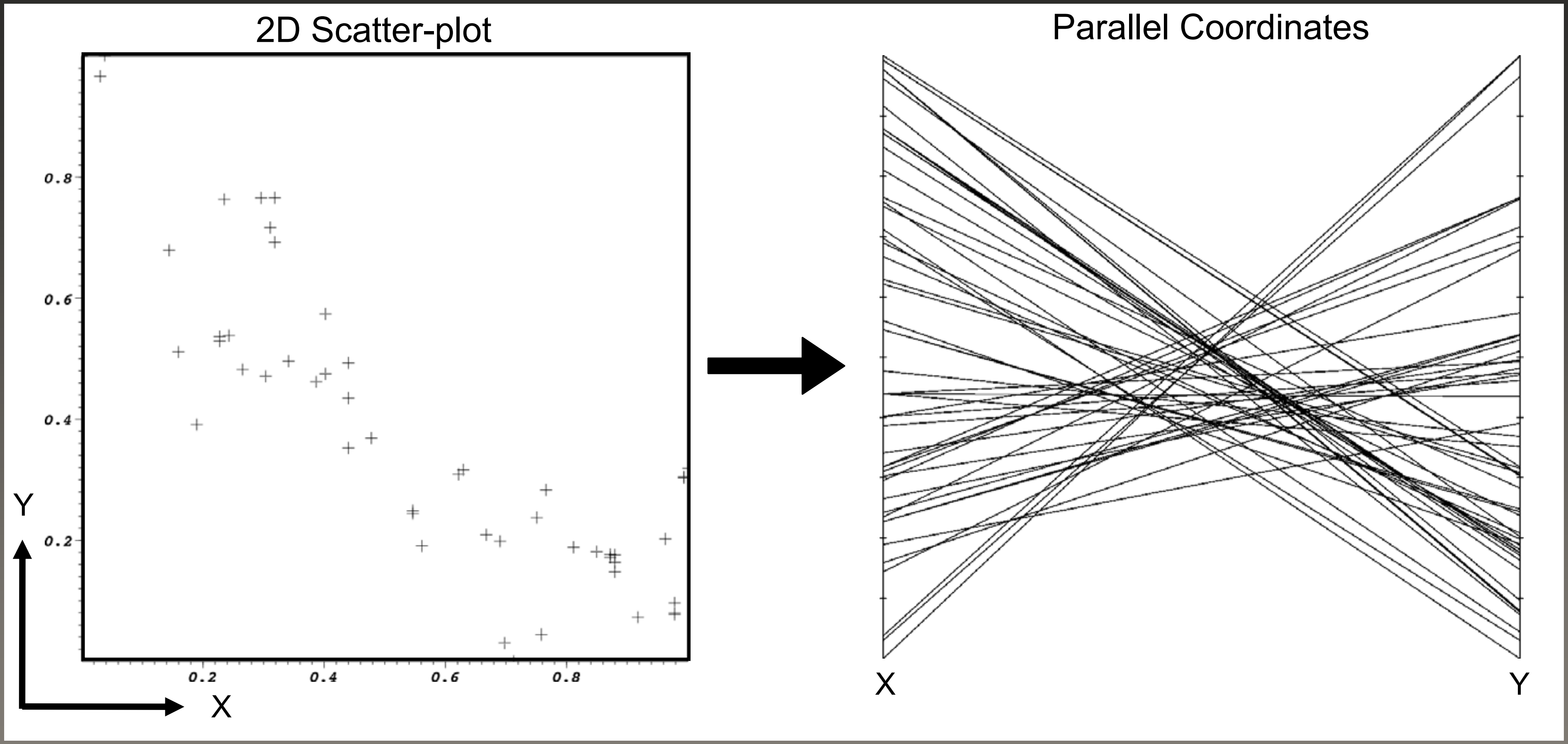
Data Selection via FastBit

Value	b_0	b_1	b_2	b_3	b_4	b_5
0	**1**	0	0	0	0	0
1	0	**1**	0	0	0	0
5	0	0	0	0	0	**1**
3	0	0	0	**1**	0	0
1	0	**1**	0	0	0	0
2	0	0	**1**	0	0	0
4	0	0	0	0	1	0
…						
	=0	**=1**	**=2**	**=3**	**=4**	**=5**

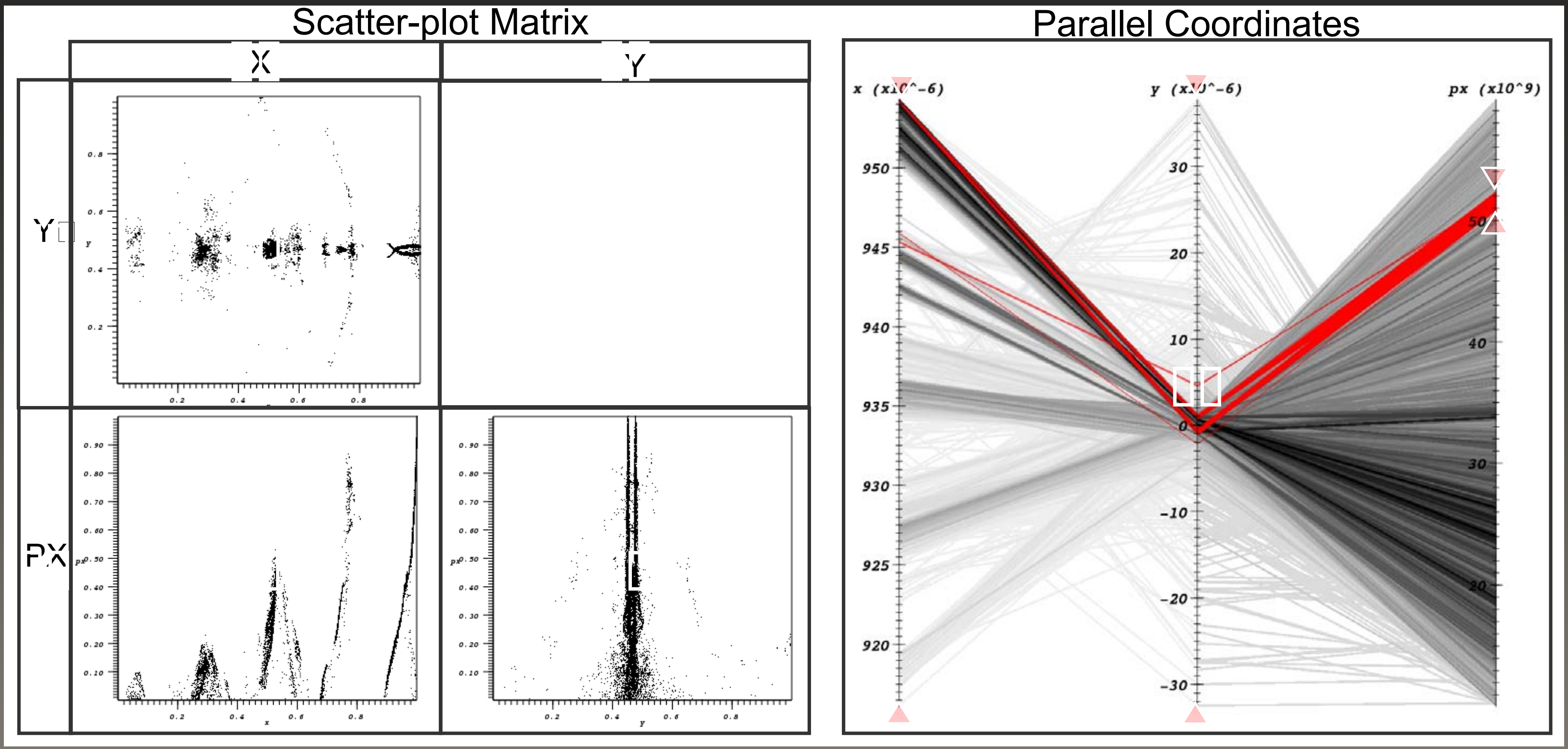
- Use FastBit to accelerate:
  - Computation of conditional histograms for parallel coordinate rendering
  - Multi-dimensional threshold queries for particle of interest identification
  - ID-queries for tracing of particles over time:

Reference: K. Wu, E. Otoo, and A. Shoshani, "Compressing bitmap indexes for faster search operations", ACM Transactions on Database Systems, vol 31, pp. 1-38, 2006

Introduction to Parallel Coordinates



Introduction to Parallel Coordinates, cont.



Advantages:

Parallel display of many data dimensions

Easy interface for data thresholding

Immediate feedback during data selection

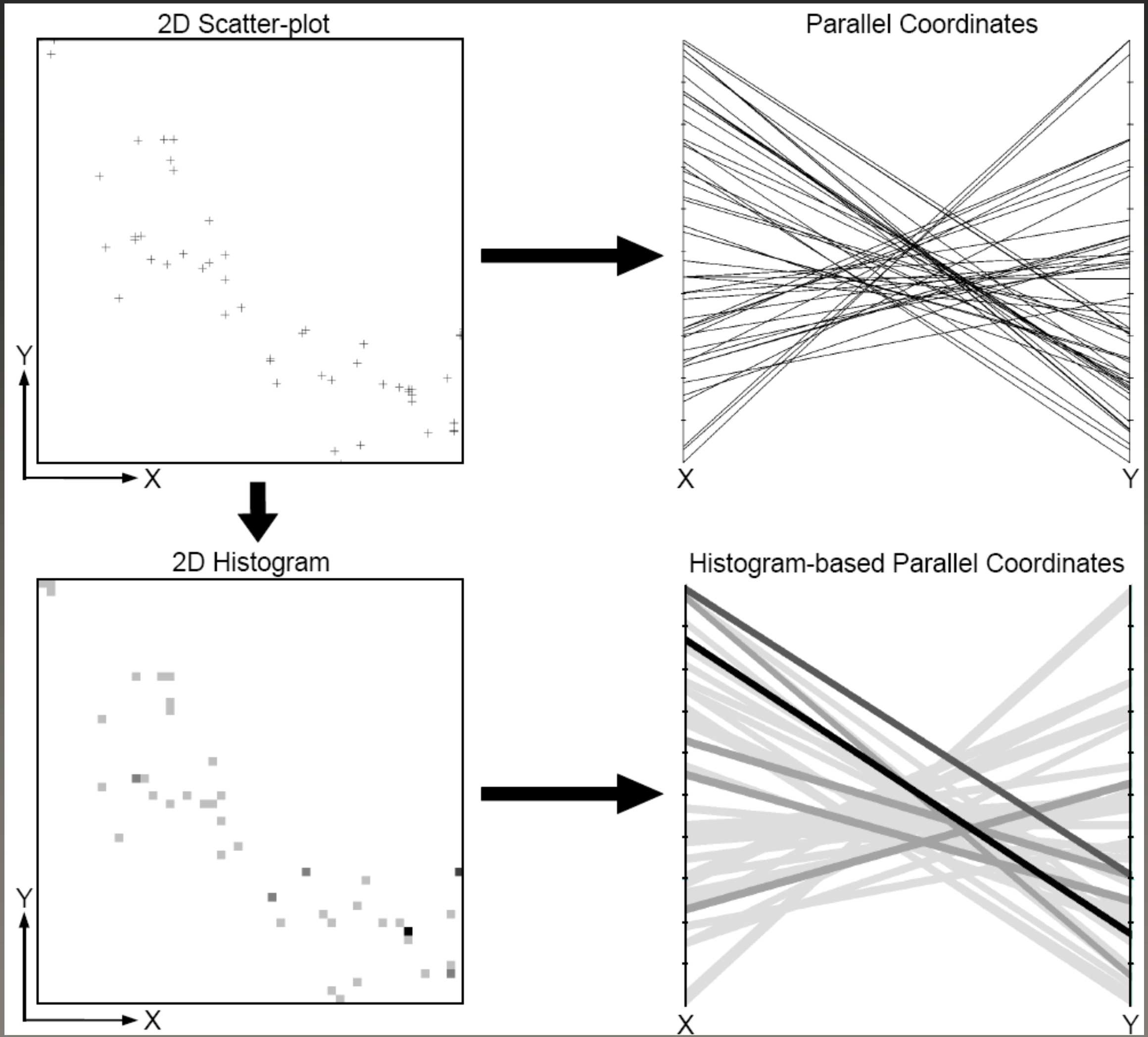
Disadvantages:

Order dependent visualization

Clutter, Occlusion

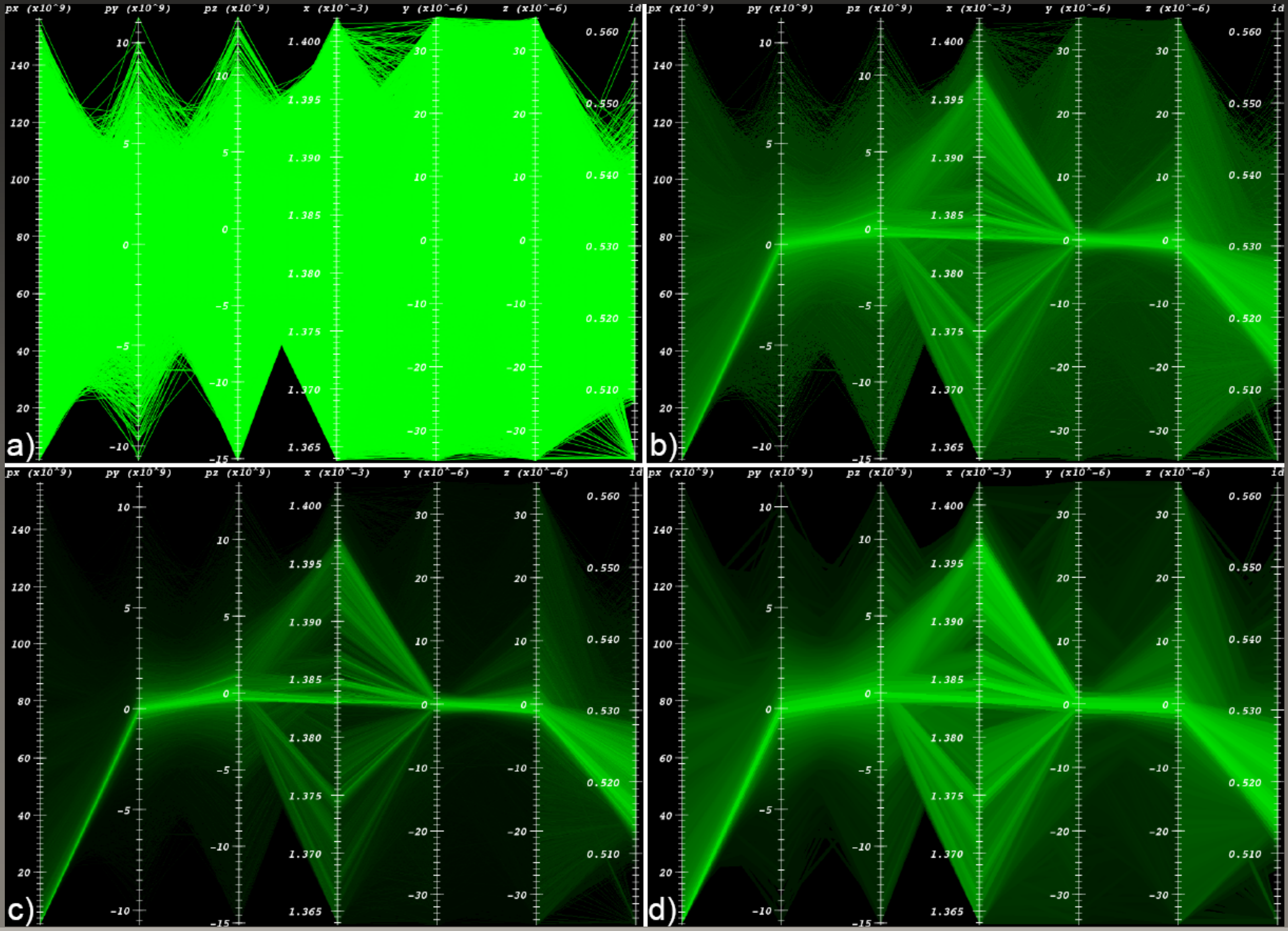
Comp. complexity proportional to data size

Histogram-based Parallel Coordinates



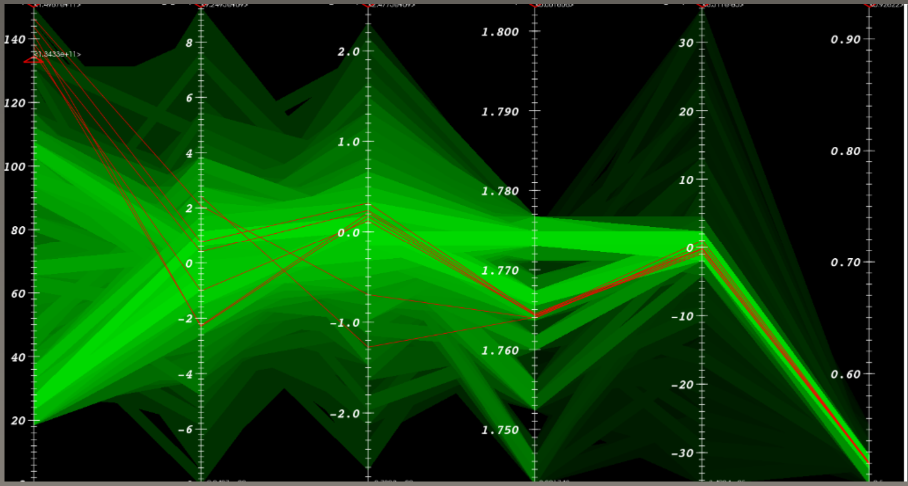
Reference: M. Novotny and H. Hauser, “Outlier-preserving focus+context visualization in parallel coordinates,” *IEEE Transactions on Visualization and Computer Graphics,* vol. 12, no. 5, pp. 893-900. 2006.

Histogram-based Parallel Coordinates cont.

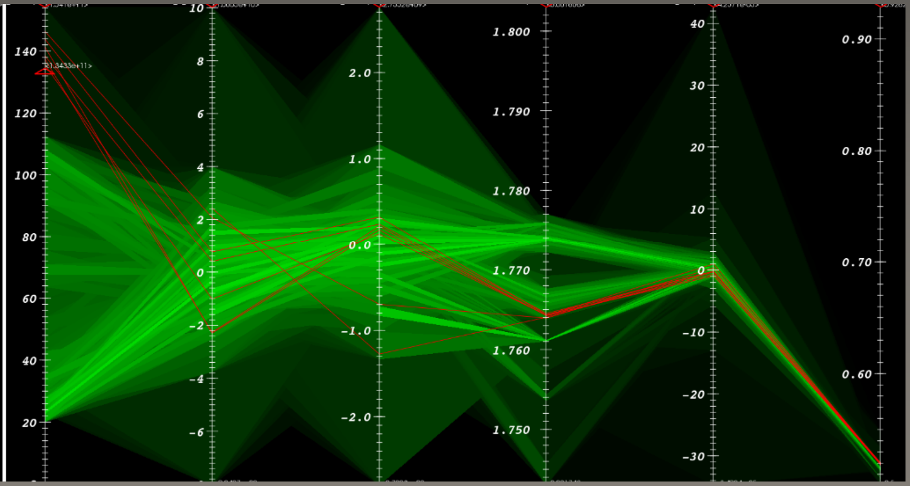


Histogram-based Parallel Coordinates, cont.

- Histograms computed on request:
  - Rendering of data subsets using histograms
  - Close zoom-ins and smooth drill-downs into the data
  - Rendering with arbitrary number of bins
- Support adaptively binned histograms:
  - More accurate representation in lower-level-of-detail views



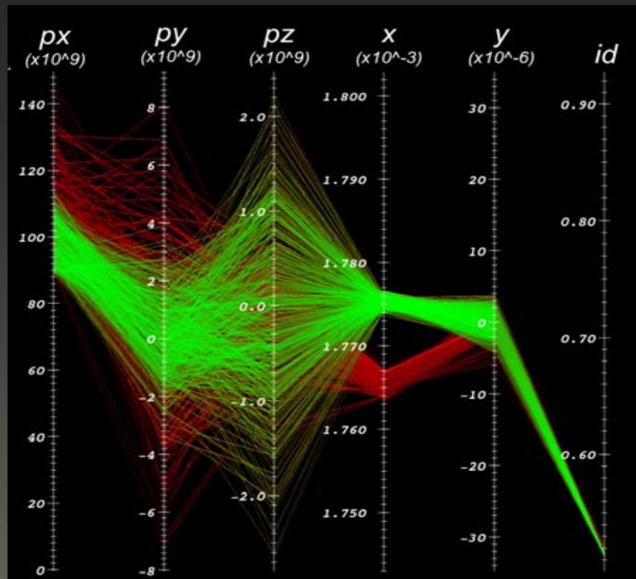
32x32 uniform binning

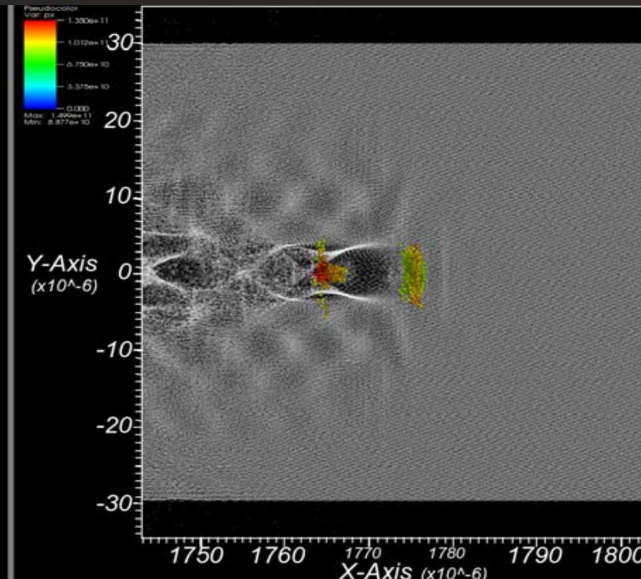


32x32 adaptive binning

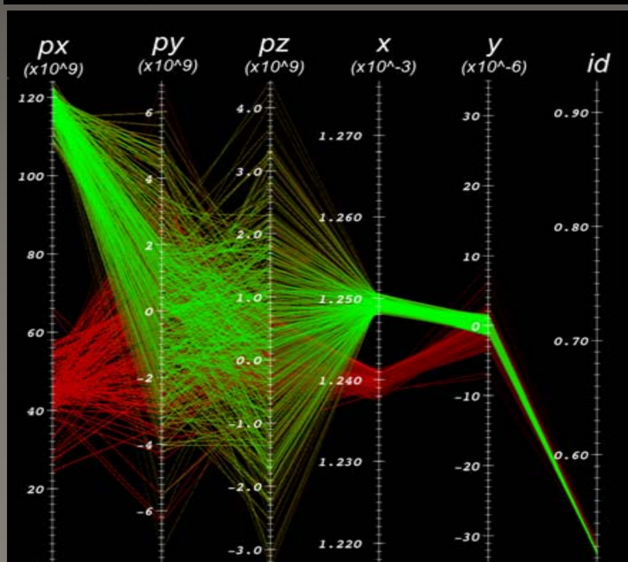
Beam Selection and Assessment

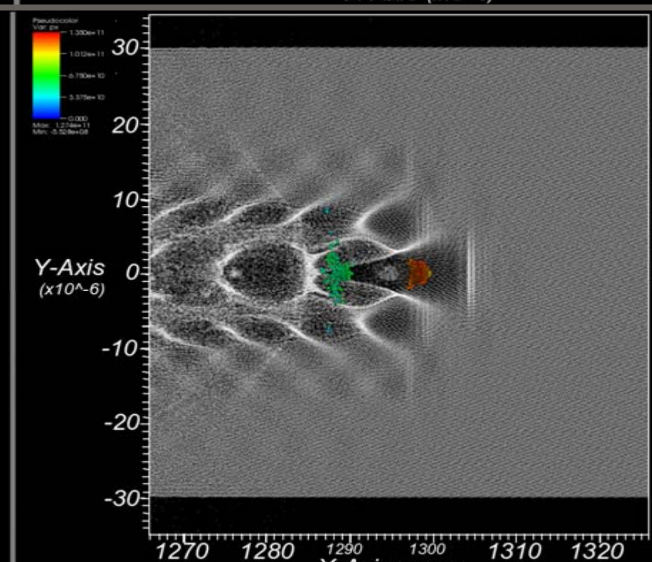
t=37



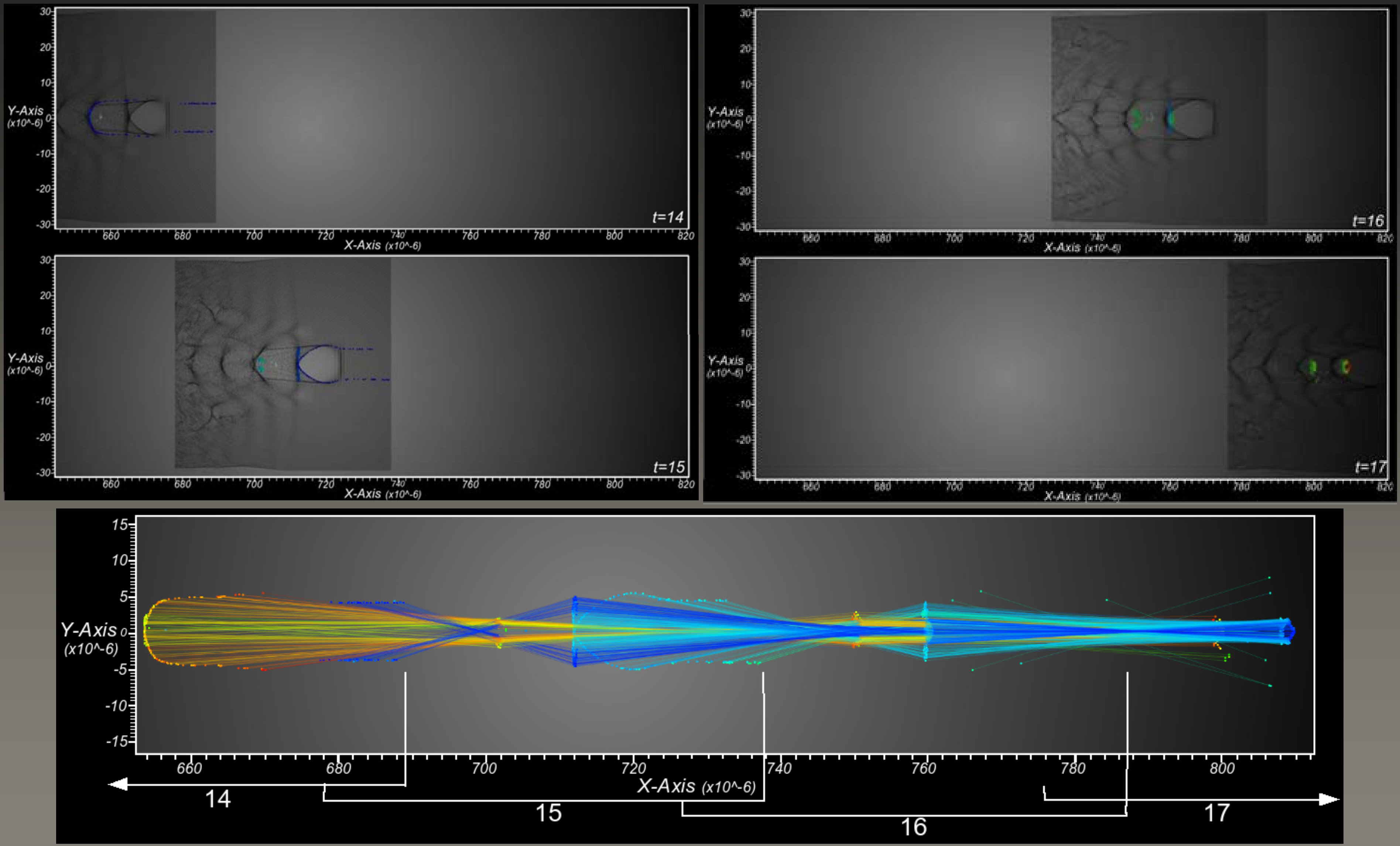


t=27

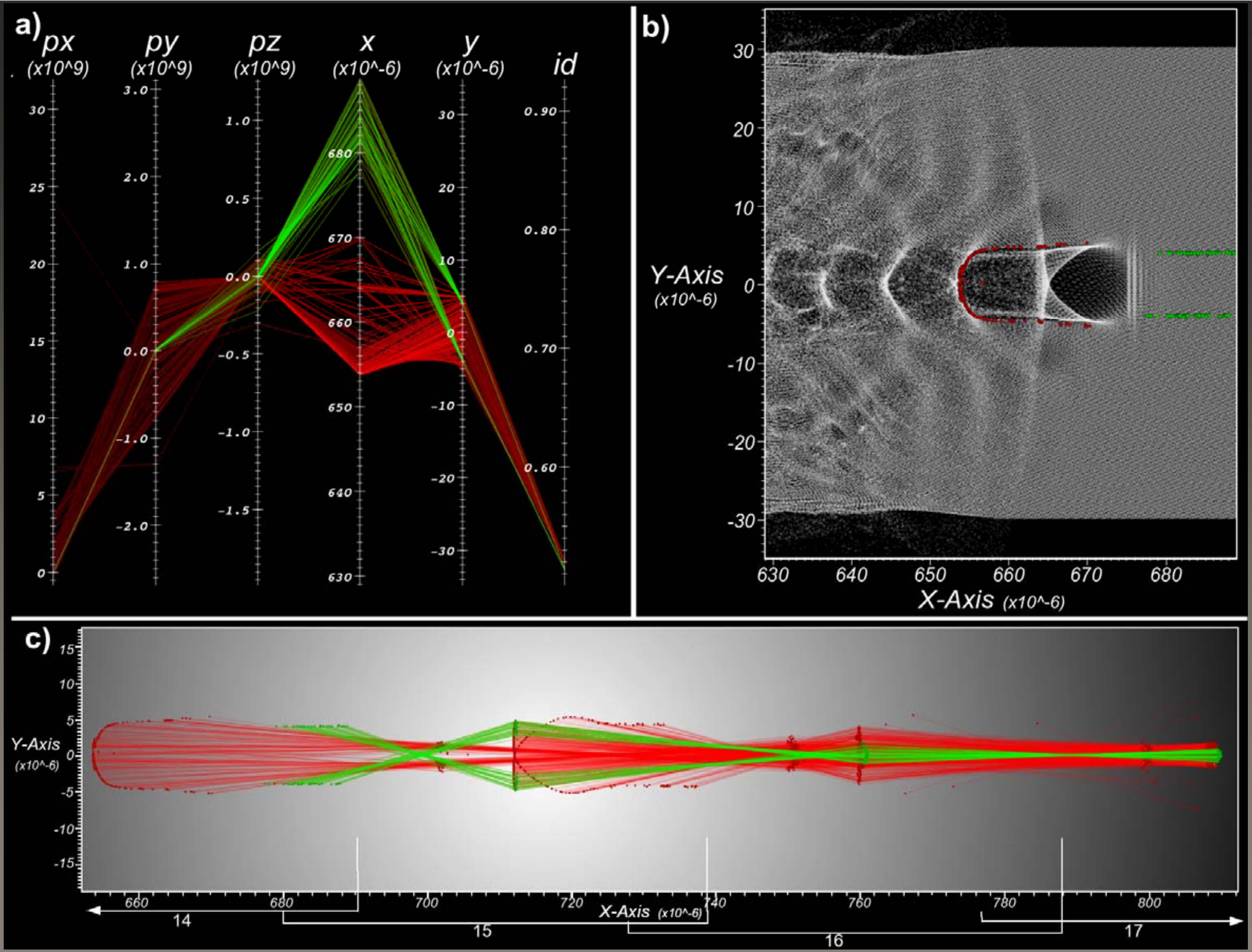




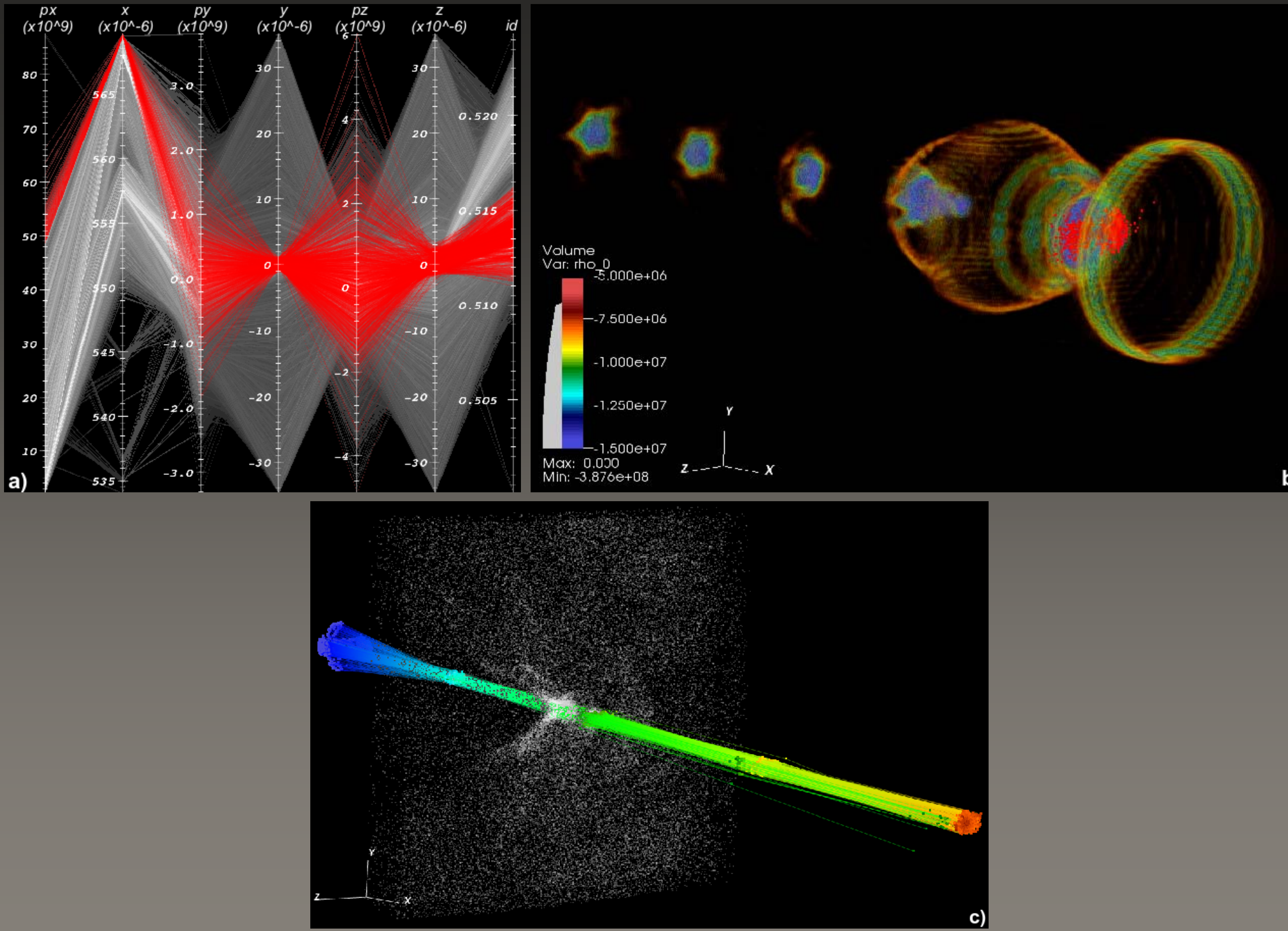
Beam Formation



Beam Refinement



3D Analysis Example



Laser Wakefield Acceleration Simulations – Conclusions and Future Work

- **Conclusions**
  - Rapid knowledge discovery from large, complex, multivariate, time-varying data
  - New approach for quickly generating histogram-based parallel coordinates
  - Case study on how system can be used to analyze Laser Wakefield particle acceleration data effectively
- **Future Work**
  - Distribute in public VisIt version (1.12)
  - Support for more file formats, e.g., particles in Chombo files
  - Explore parallelizing most expensive system parts
  - Improve integration of field and particle data
  - Couple with other traditional data analysis methods, e.g., clustering

Acknowledgements

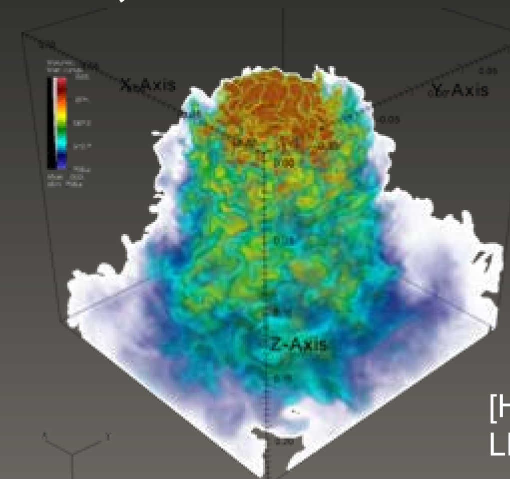
- **This work was supported by**
  - Director, Office of Advanced Scientific Computing Research, Office of Science, U.S. Department of Energy under Contract No. DE-AC02-05CH11231 through the Scientific Discovery through Advanced Computing (SciDAC) program's Visualization and Analytics Center for Enabling Technologies (VACET).
  - National Energy Research Scientific Computing Center, supported by Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.
- **We would like to thank**
  - Members of the VisIt team
  - Members of VACET
  - Members of LBNL Vis Group
  - Members of LBNL Center for Computational Sciences and Engineering

Questions?

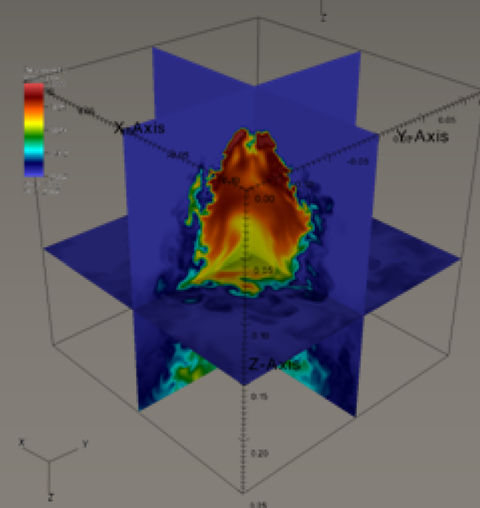
Additional Slides

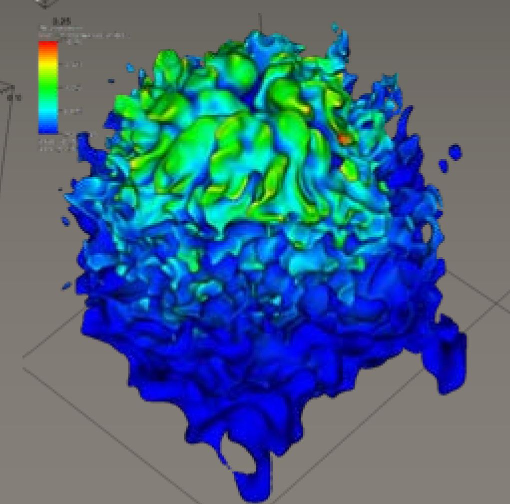
VisIt Concepts (1/2)

- **Databases:** Implementation of various file formats, e.g., Chombo, Boxlib, Enzo, FLASH, …
- **Plots:** Display data on screen (volume rendering, pseudocolor, isosurface, vector glyphs)
- **Operators:** Filter data before it is displayed (slice, isosurface, clip, displace, …)



[Hydrogen flame, LBNL CCSE]





VisIt Concepts (2/2)

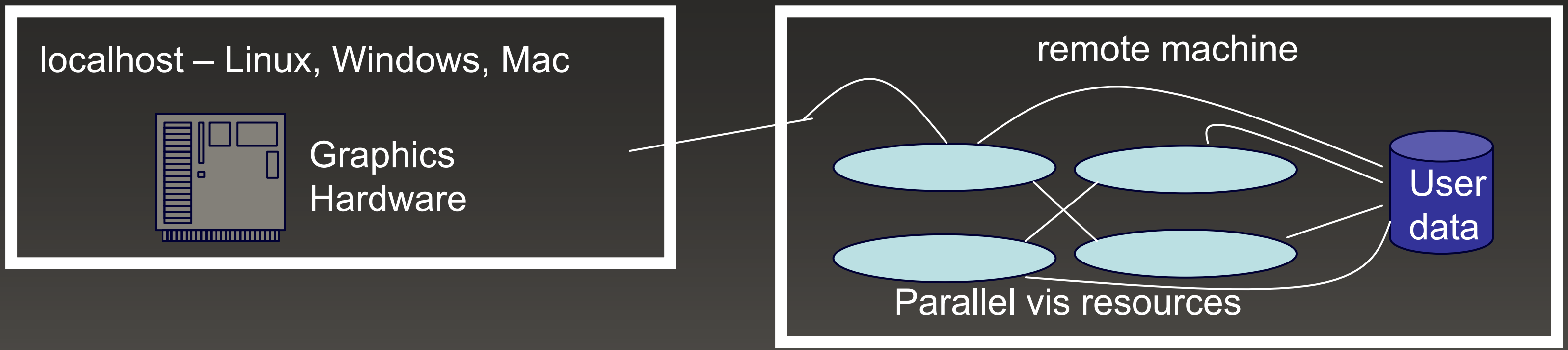
- **Expressions:**
  - Create new variables from existing ones via arithmetic expressions: +, -, *, /, dot product, cross product, …
  - Conditionals and comparisons
  - Import variables from different time steps or simulations
  - Other operations: image processing (smoothing), connected components of isosurfaces, …
- **Picking:** Get information (value, location in mesh) for a given “point” in a visualization
- **Queries:** High-level information (area, volume, integral of variable, number of connected components, …) coupled with ability to create curve of quantity over time

VisIt Capabilities

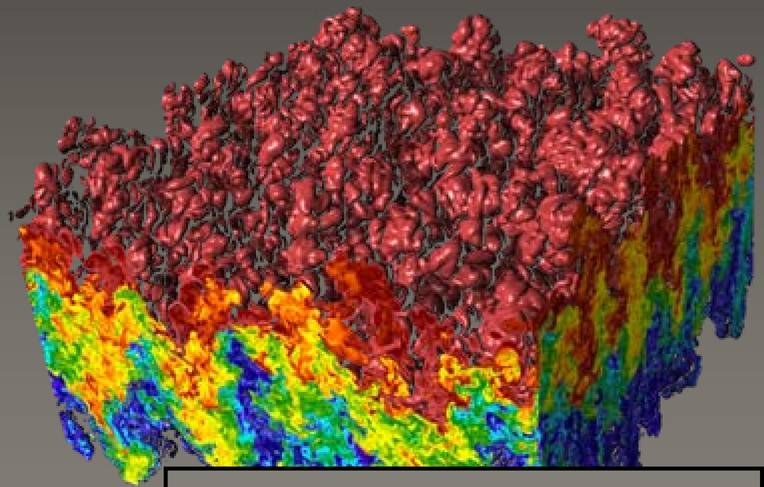
- Meshes: rectilinear, curvilinear, unstructured, point, AMR
- Data: scalar, vector, tensor, material, species
- Dimension: 1D, 2D, 3D, time varying
- Rendering (~15): pseudocolor, volume rendering, hedgehogs, glyphs, mesh lines, etc…
- Data manipulation (~40): slicing, contouring, clipping, thresholding, restrict to box, reflect, project, revolve, …
- File formats (~85)
- Derived quantities: >100 interoperable building blocks

  +,-,*,/, gradient, mesh quality, if-then-else, and, or, not
- Many general features: position lights, make movie, etc
- Queries (~50): ways to pull out quantitative information, debugging, comparative analysis

VisIt's Parallelized Client-Server Architecture.



- Client-server observations:
  - Good for remote visualization
  - Leverages available resources
  - Scales well
  - No need to move data



27B element Rayleigh-Taylor Instability (MIRANDA, BG/L)

VisIt and AMR Data

- Supported as “first-class” data type
- VisIt understands:
  - Nesting of patches / Boundaries between patches
- Strategy:
  - Each patch is processed separately
  - After reading data, subsequent passes to:
    - Identify coarse elements that are refined and mark them as “ghost”
    - Create ghost layers around outer boundaries (needs work)
  - Work on rectilinear grids and skip ghost cells or “remove” results produced in ghost cells later on
- UI:
  - Color by patch, color by level
  - Remove / show patches, levels
  - Have appropriate info returned for picks