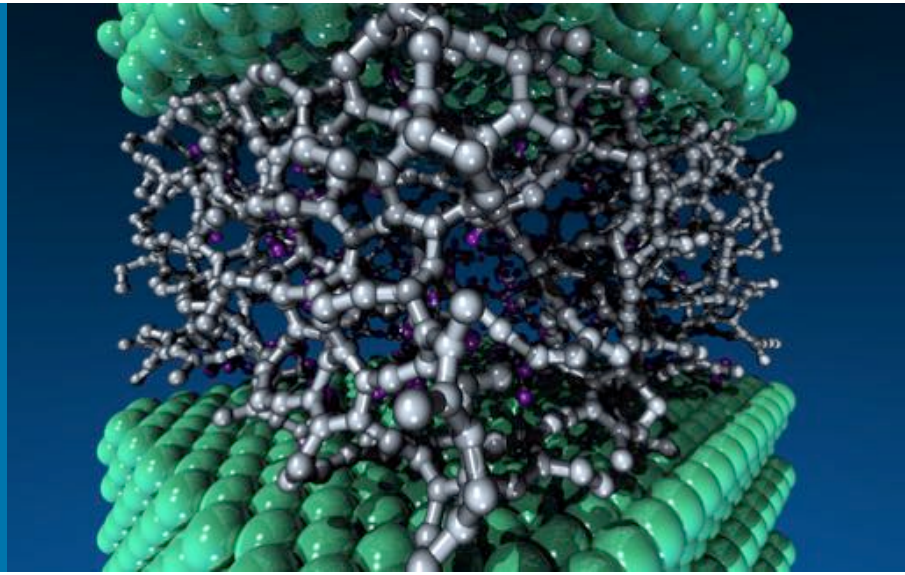
Argonne NATIONAL LABORATORY

VISUALIZATION INTRODUCTION



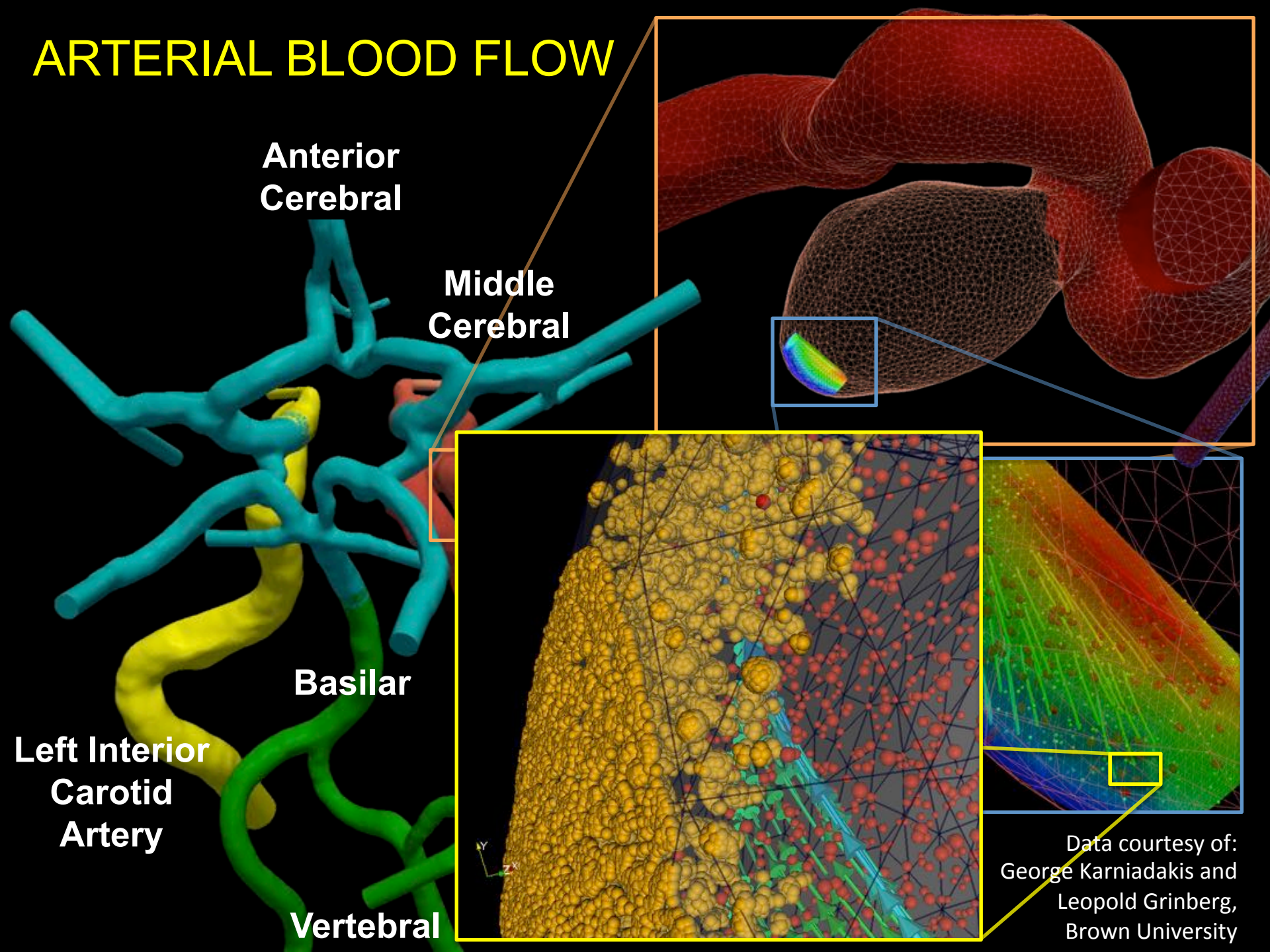
JOSEPH INSLEY
August 10, 2016

ATPESC 2016

HERE'S THE PLAN…

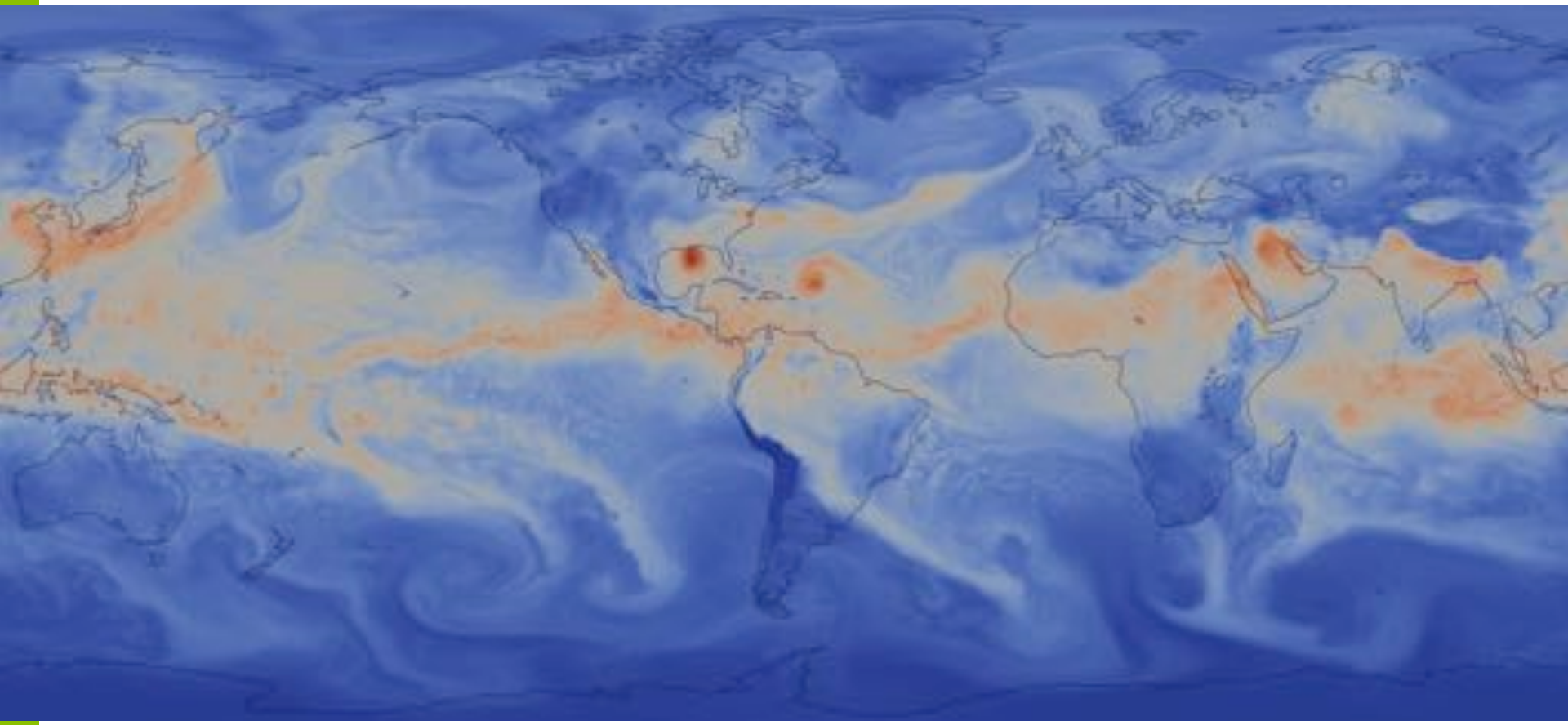
- Examples of visualizations
- Visualization resources
- Data and transformations
- Visualization tools and formats
- Data representations
- Visualization in production and development

ARTERIAL BLOOD FLOW



Data courtesy of:
George Karniadakis and
Leopold Grinberg,
Brown University

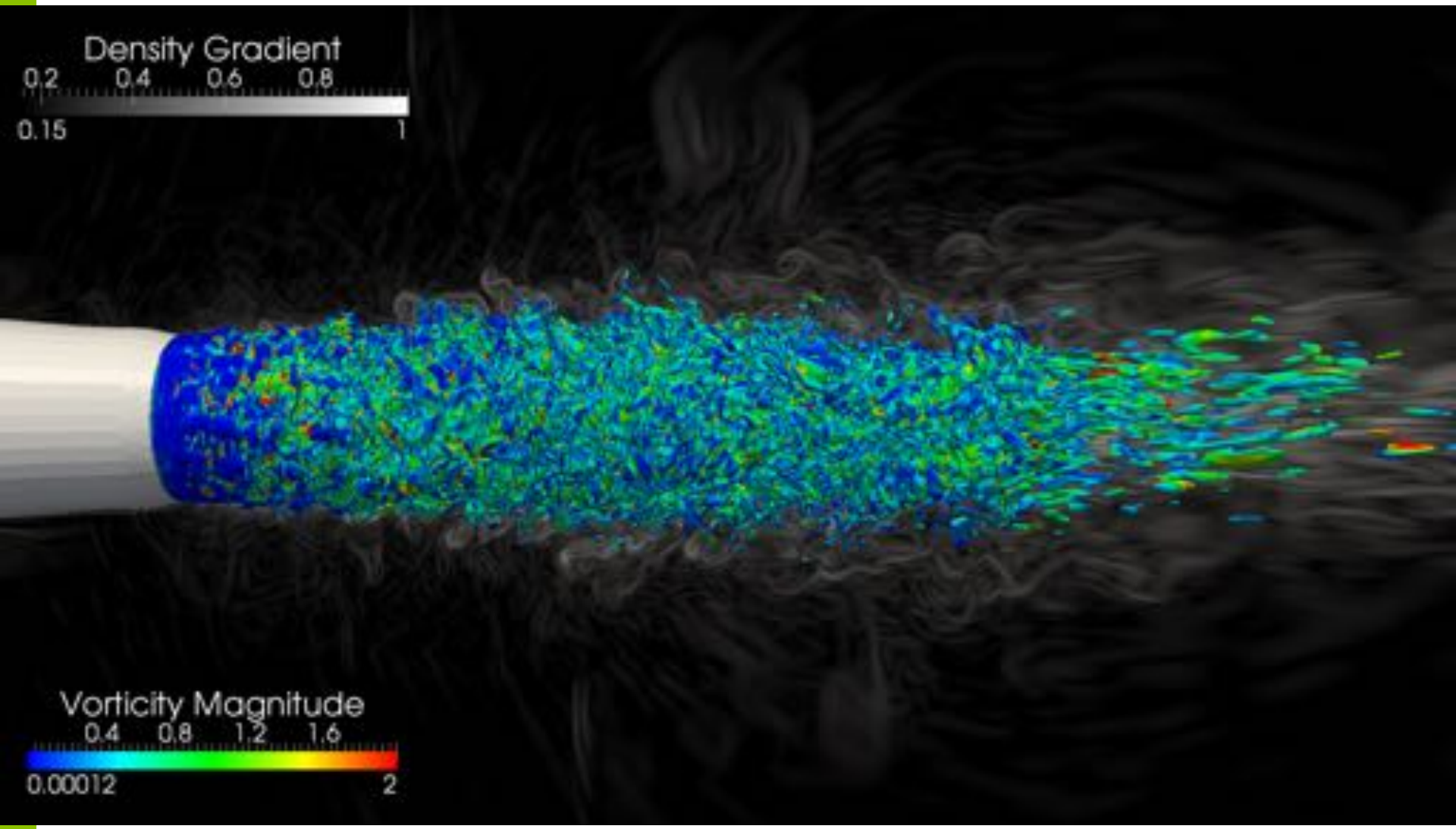
CLIMATE



Data courtesy of: Mark Taylor, Sandia National Laboratory; Rob Jacob, Argonne National Laboratory; Warren Washington, National Center for Atmospheric Research

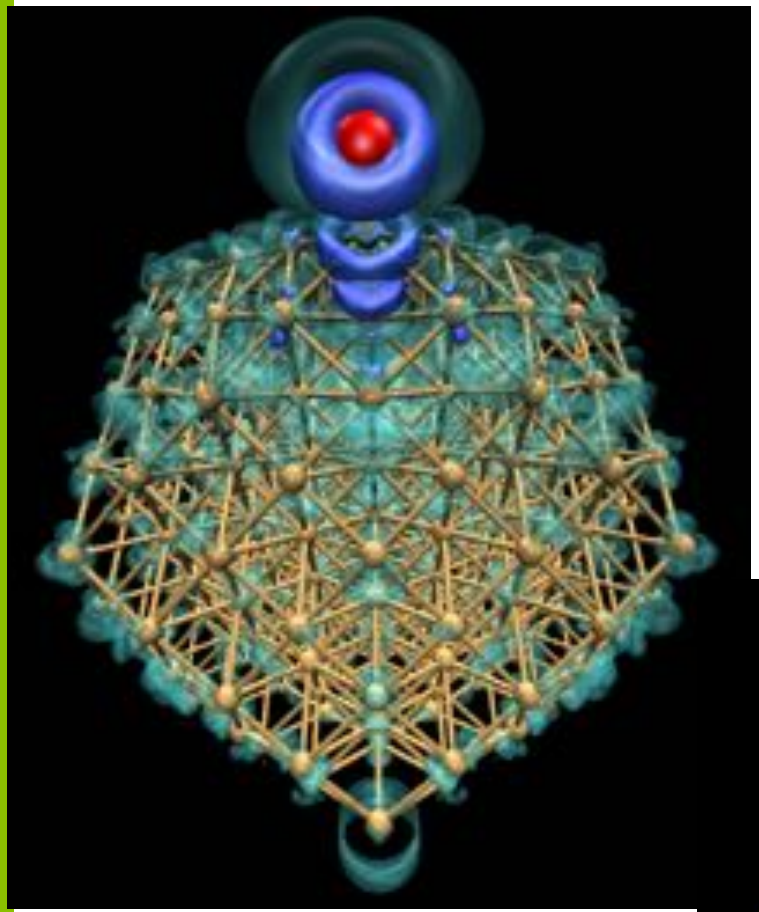
Argonne NATIONAL LABORATORY

AEROSPACE (JET NOZZLE NOISE)



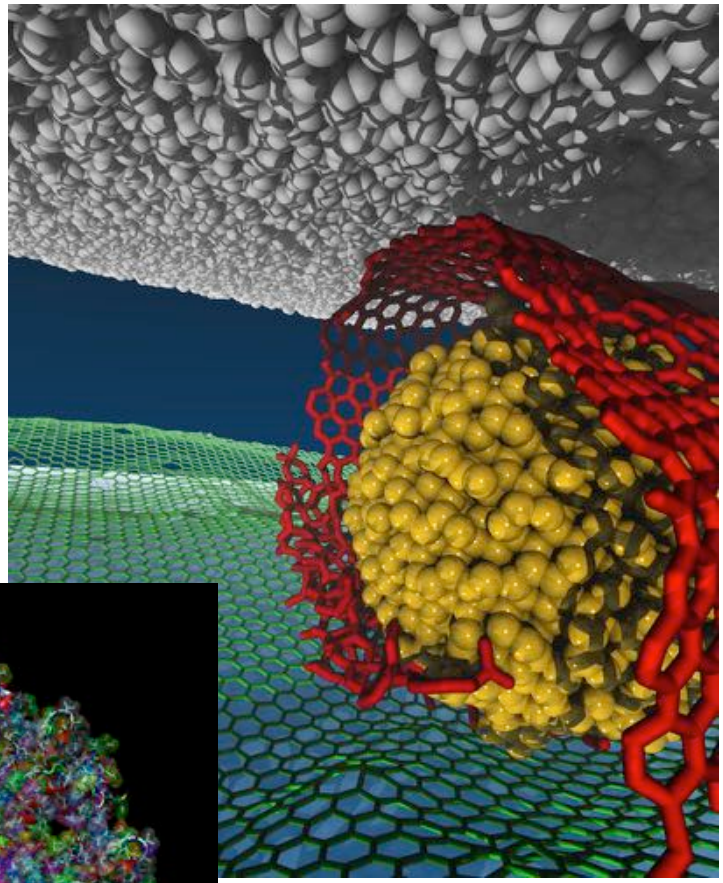
Data courtesy of: Anurag Gupta and Umesh Paliath, General Electric Global Research

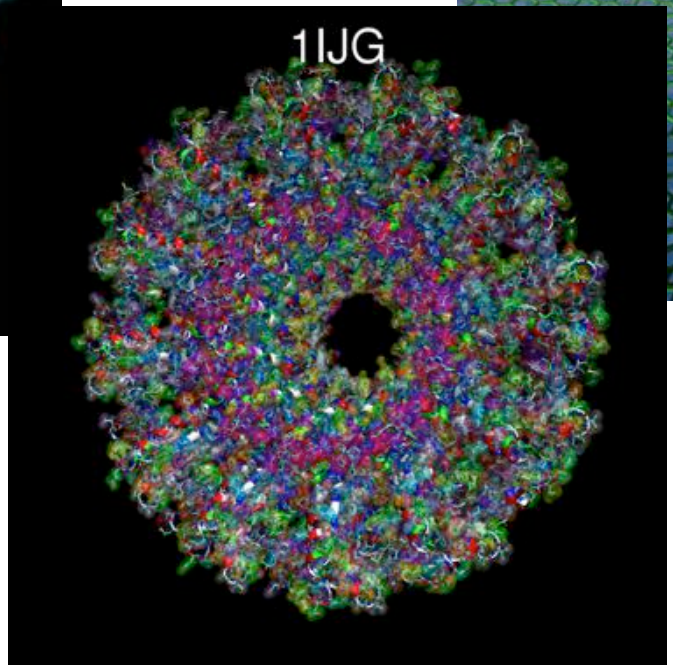
MATERIALS SCIENCE / MOLECULAR



Data courtesy of: Jeff Greeley, Nichols Romero, Argonne National Laboratory

Data courtesy of: Subramanian Sankaranarayanan, Argonne National Laboratory

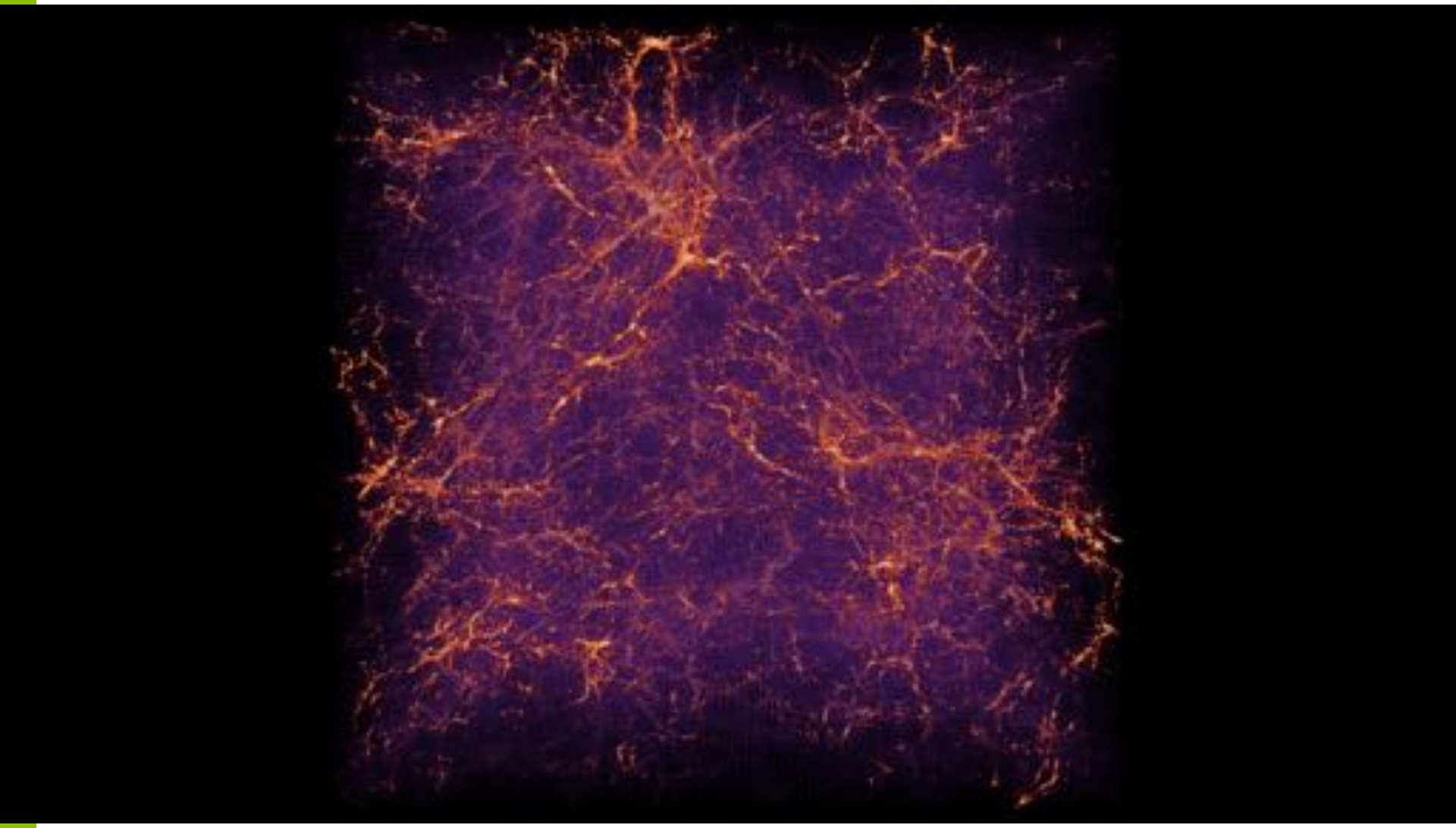




Data courtesy of: Advanced Photon Source, Argonne National Laboratory

Argonne NATIONAL LABORATORY

COSMOLOGY

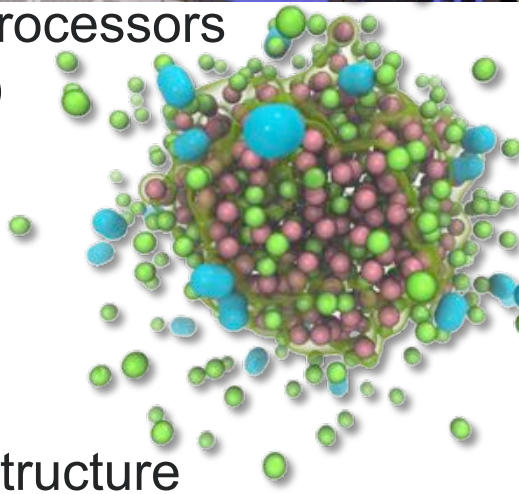


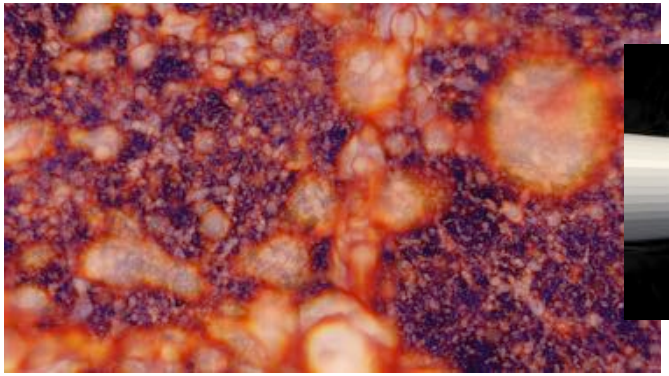
Data courtesy of: Salman Habib, Katrin Heitmann, Argonne National Laboratory

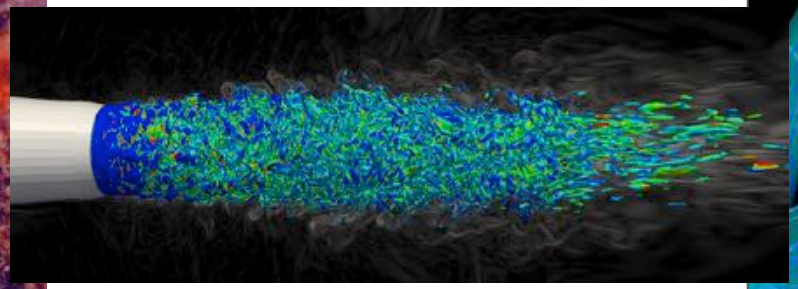
Argonne NATIONAL LABORATORY

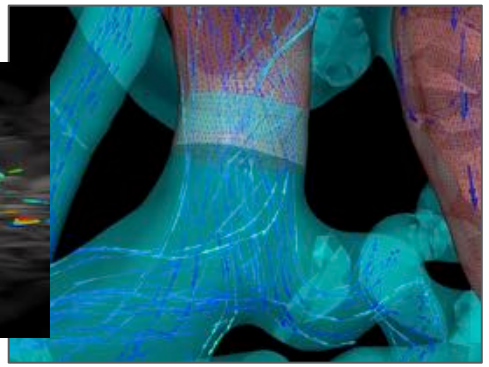
COOLEY

- Analytics/Visualization cluster
- Peak 223 TF
- 126 nodes; each node has
  - Two Intel Xeon E5-2620 Haswell 2.4 GHz 6-core processors
  - NVIDIA Telsa K80 graphics processing unit (24GB)
  - **384 GB of RAM**
- Aggregate RAM of 47 TB (vs. ~6TB for Tukey)
- Aggregate GPU memory of~3TB (vs. ~1.1TB for Tukey)
- Cray CS System
- 216 port FDR IB switch with uplinks to our QDR infrastructure
- Mounts the same GPFS file systems as Mira, Cetus







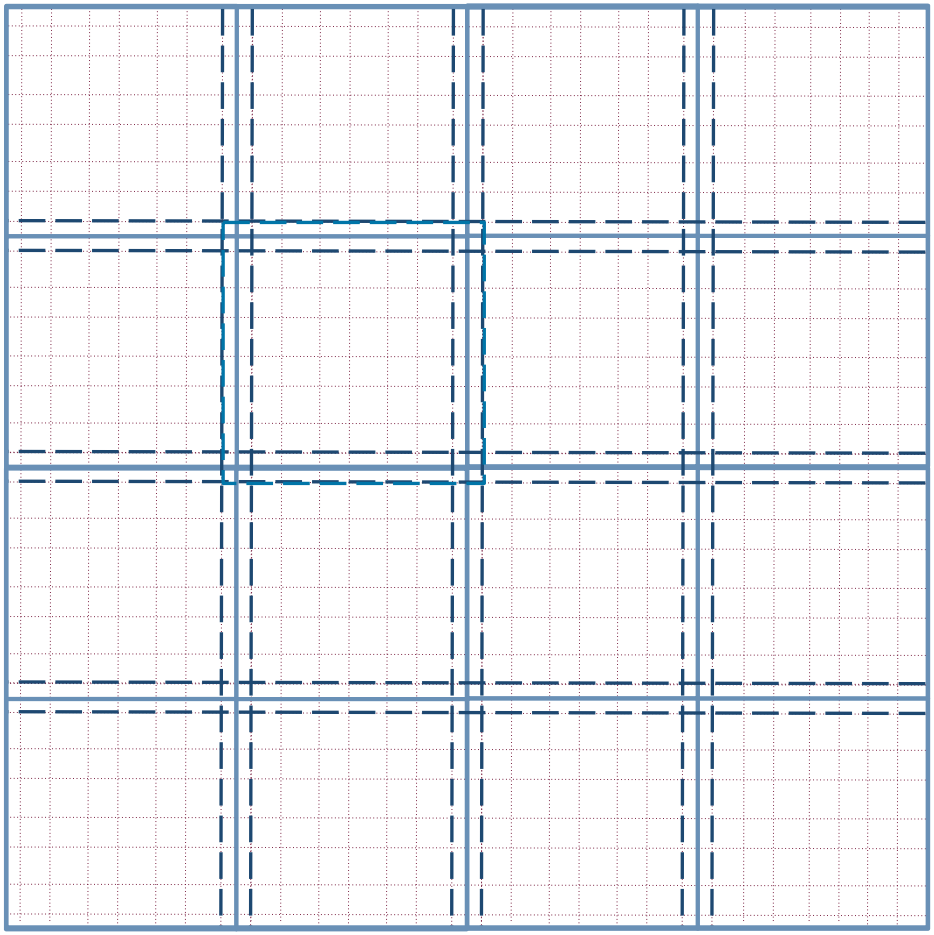


VISUALIZATION ALGORITHMS (TRANSFORMATIONS)

- Structure
  - Geometric
    - Translate, rotate, scale coordinates
    - Topology remains unchanged
  - Attribute
    - Transform or create new data attributes
  - Combined
- Type
  - Scalar (single value)
  - Vector (array of values)
  - Tensor (matrix of values)
  - Combined

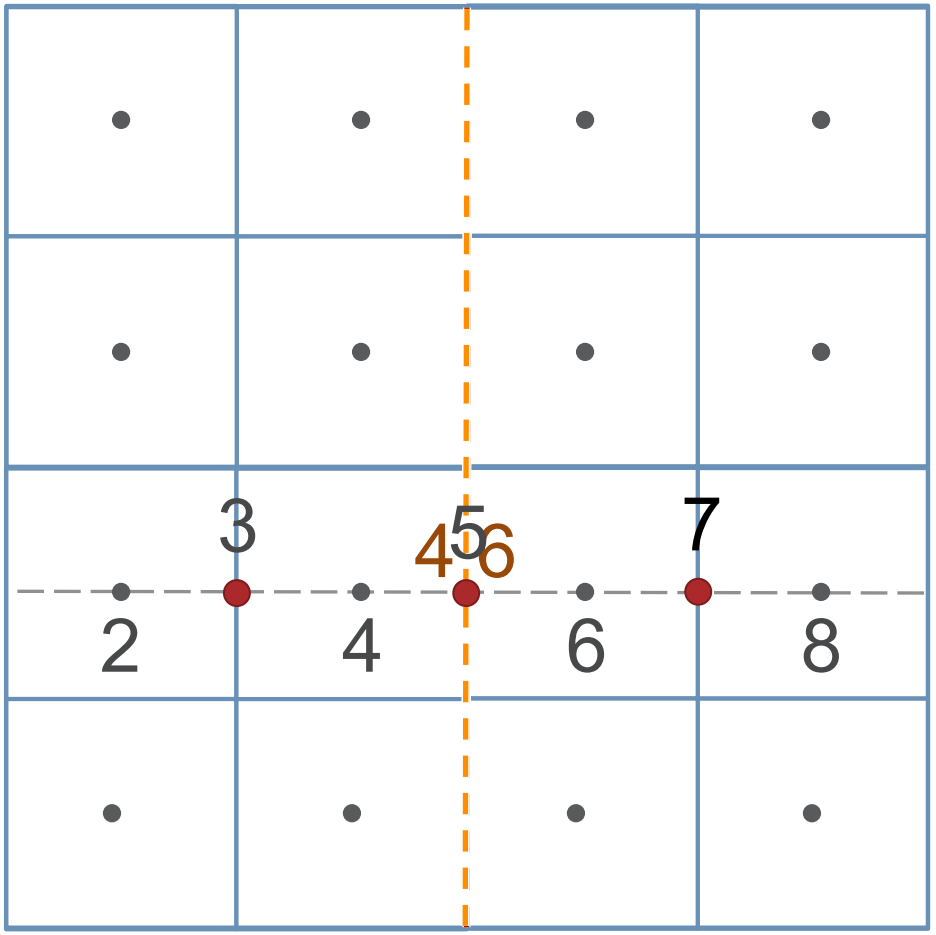
DATA DOMAIN DECOMPOSITION: REGULAR GRID

- Regularly sized/spaced grid of cells, each holds a single value (per variable)
- Data domain is divided among available processes
- Additional “ghost” cells are required to ensure accuracy



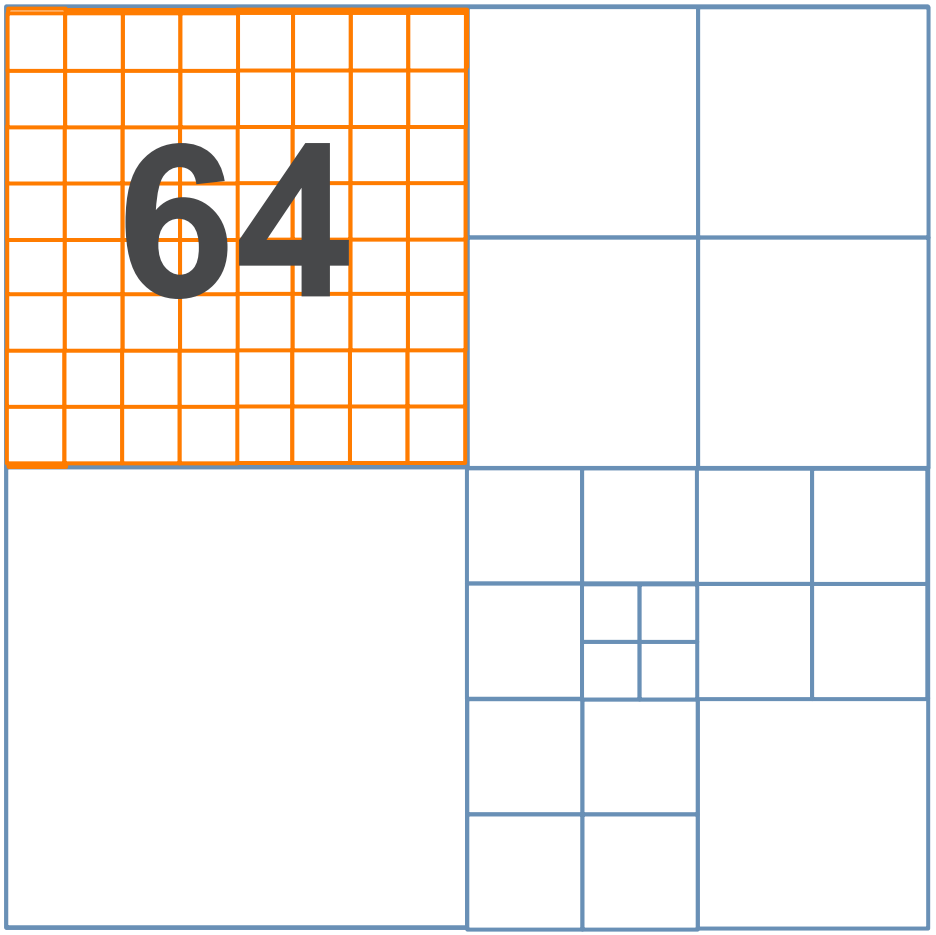
DATA DOMAIN DECOMPOSITION: REGULAR GRID

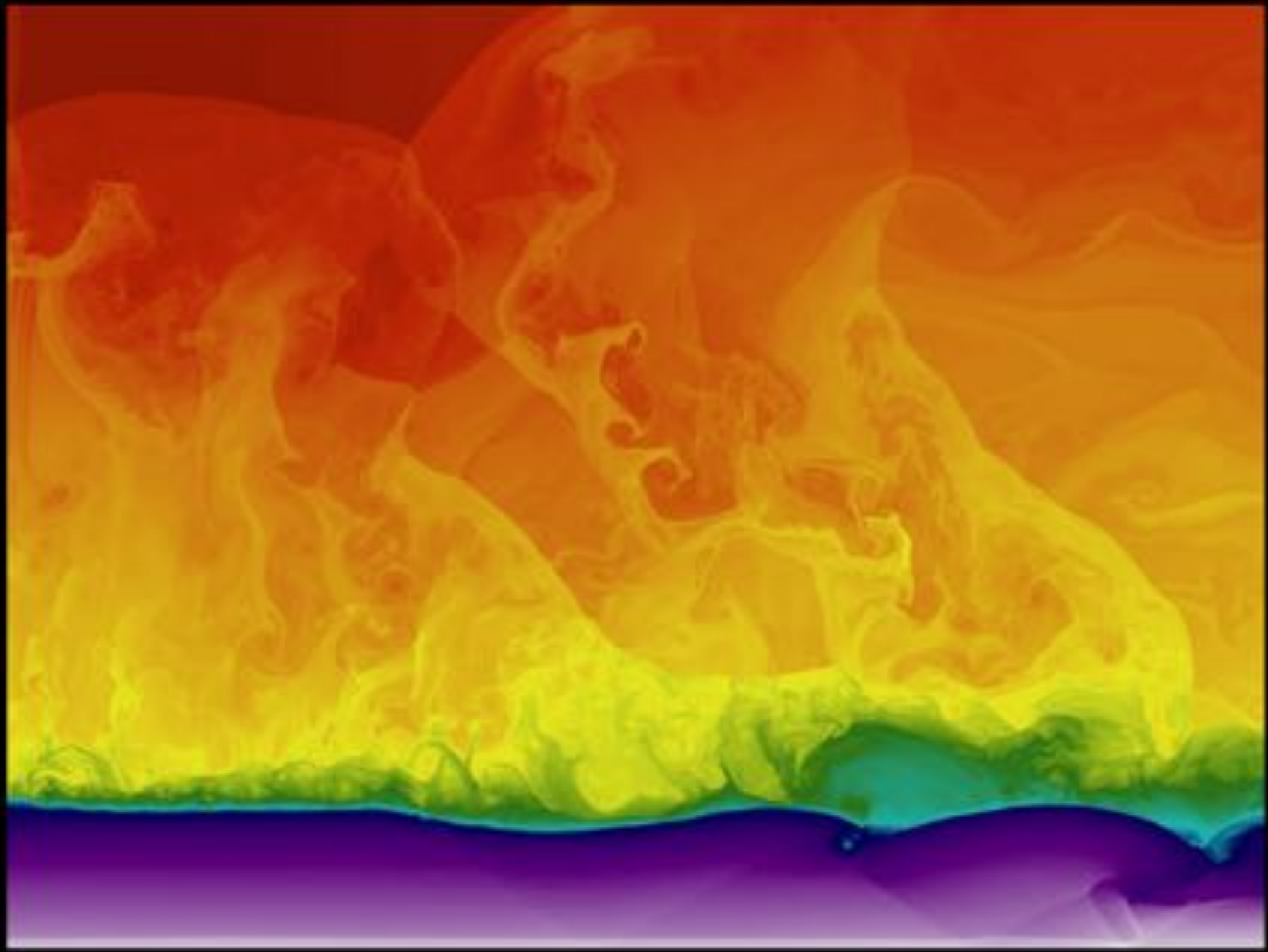
- Regularly sized/spaced grid of cells, each holds a single value (per variable)
- Data domain is divided among available processes
- Additional “ghost” cells are required to ensure accuracy



DATA DOMAIN DECOMPOSITION: ADAPTIVE MESH REFINEMENT (AMR)

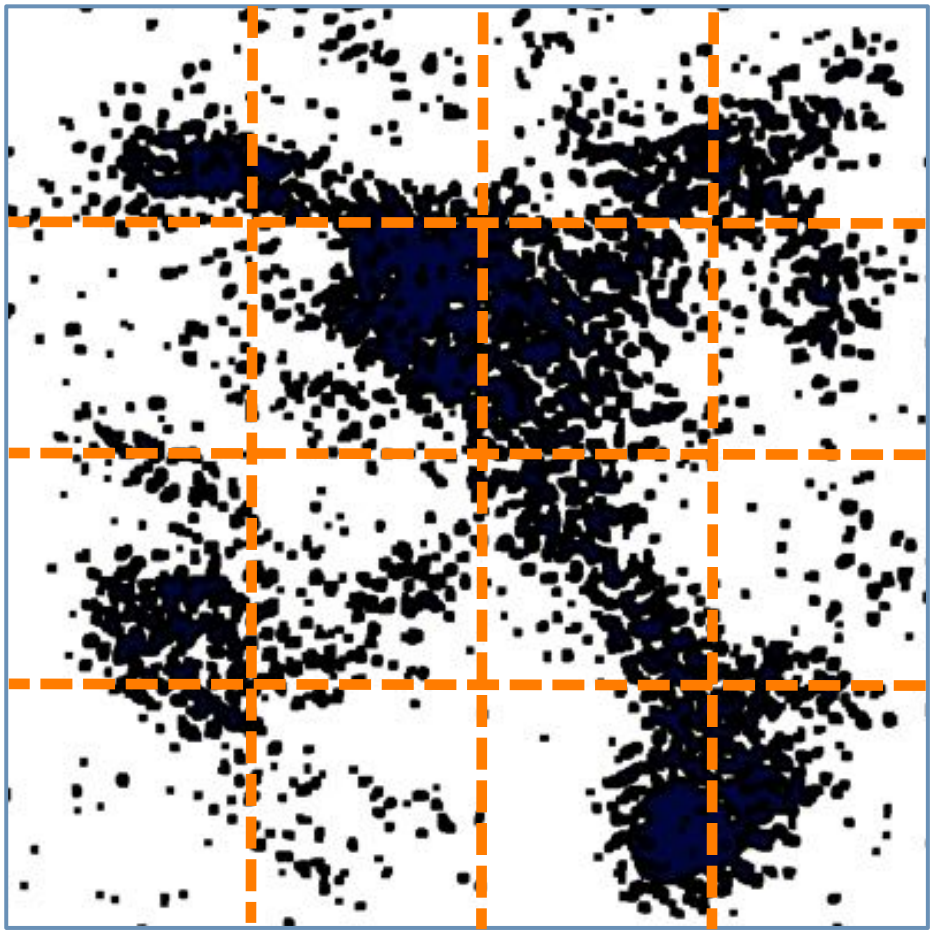
- Puts increased detail in regions where things are changing more rapidly.
- Can increase computational performance
- Results in smaller data sets





DATA DOMAIN DECOMPOSITION: PARTICLE-BASED

- Keep track of individual particles
- Decomposition could be based on particles, or spatial extents
- Can project them onto a grid
  - combine (e.g. average) all particles in each grid cell



ALL SORTS OF TOOLS

- Visualization Applications
  - VisIt
  - ParaView
  - EnSight
- Domain Specific
  - VMD, PyMol, RasMol
- APIs
  - VTK: visualization
  - ITK: segmentation & registration
- GPU performance
  - vl3: shader-based volume rendering
- Analysis Environments
  - Matlab
  - Parallel R
- Utilities
  - GnuPlot
  - ImageMagick
- Visualization Workflow
  - VisTrails

PARAVIEW & VISIT VS. VTK

- ParaView & VisIt
  - General purpose visualization applications
  - GUI-based
  - Scriptable
  - Extendable
  - Built on top of vtk (largely)

- vtk
  - Programming environment / API
  - Additional capabilities, finer control
  - Smaller memory footprint
  - Requires more expertise (build custom applications)

DATA FILE FORMATS (PARAVIEW & VISIT)

- VTK
- Parallel (partitioned) VTK
- VTK MultiBlock (MultiGroup, Hierarchical, Hierarchical Box)
- Legacy VTK
- Parallel (partitioned) legacy VTK
- EnSight files
- EnSight Master Server
- Exodus
- BYU
- XDMF
- PLOT2D
- PLOT3D
- SpyPlot CTH
- HDF5 raw image data
- DEM
- VRML
- PLY
- Polygonal Protein Data Bank
- XMol Molecule
- Stereo Lithography
- Gaussian Cube
- Raw (binary)
- AVS
- Meta Image
- Facet
- PNG
- SAF
- LS-Dyna
- Nek5000
- OVERFLOW
- paraDIS
- PATRAN
- PFLOTRAN
- Pixie
- PuReMD
- S3D
- SAS
- Tetrad
- UNIC
- VASP
- ZeusMP
- ANALYZE
- BOV
- GMV
- Tecplot
- Vis5D
- Xmdv
- XSF

DATA WRANGLING

- XDMF
  - XML wrapper around HDF5 data
  - Can define
    - data sets
    - subsets
    - hyperslabs
- vtk
  - Could add to your simulation code
  - Can write small utilities to convert data
    - Use your own read routines
    - Write vtk data structures
  - C++ and Python bindings

DATA ORGANIZATION

- Format
  - Existing tools support many flavors
  - Use one of these formats
  - Use (or write) a format converter
  - Write a custom reader for existing tool
  - Write your own custom vis tool

- Serial vs. Parallel/Partitioned
  - Single big file vs. many small files: middle ground generally best
    - vtk data types
    - XDMF for HDF5 (VisIt and ParaView)
    - Custom

DATA ORGANIZATION

- Serial vs. Parallel/Partitioned
  - Performance trade-offs
    - vtk/paraview: serial files all data read on head node, partitioned and distributed
    - vtk/paraview: parallel files: serial files partitioned

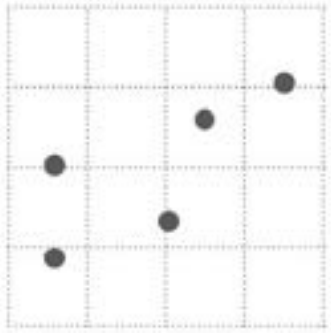
Performance example:

- Single serial .vtu file (unstructured grid)
  - Data size: ~3.8GB
  - Read time on 64 processes: > 15 minutes
    - most of this was spent partitioning and distributing
- Partitioned .pvtu file (unstructured grid)
  - Data size: ~8.7GB (64 partitions)
  - Read time on 64 processes: < 1 second

VISUAL CUES

Position

Where in space the data is

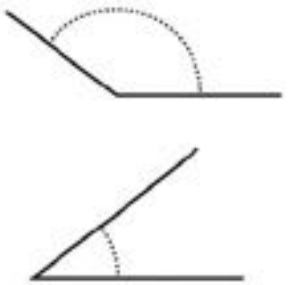


Length

How long the shapes are

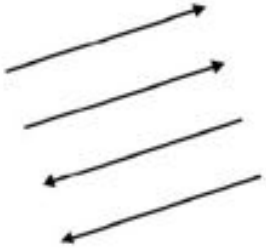
Angle

Rotation between vectors



Direction

Slope of a vector in space



Shapes

Symbols as categories

Area

How much 2-D space

Volume

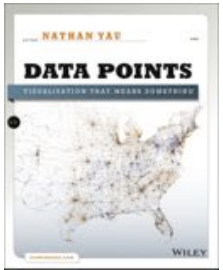
How much 3-D space

Color Saturation

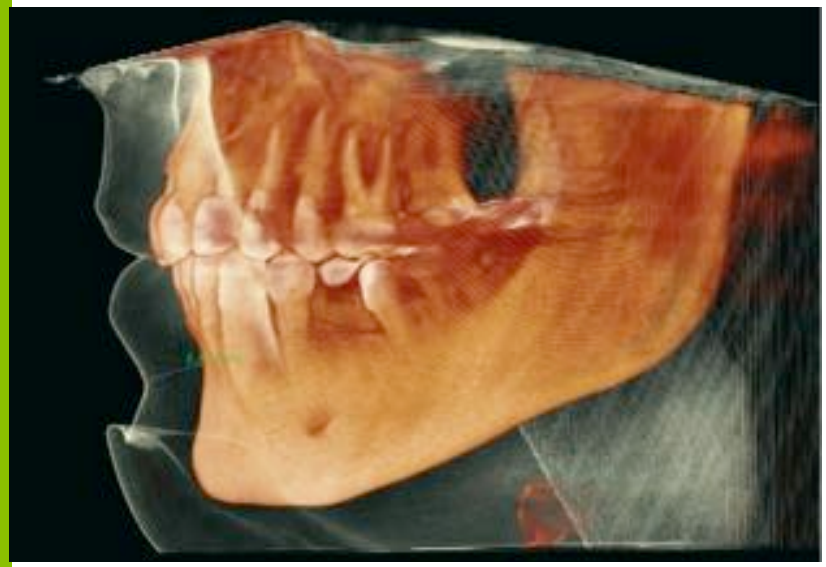
Intensity of a color hue

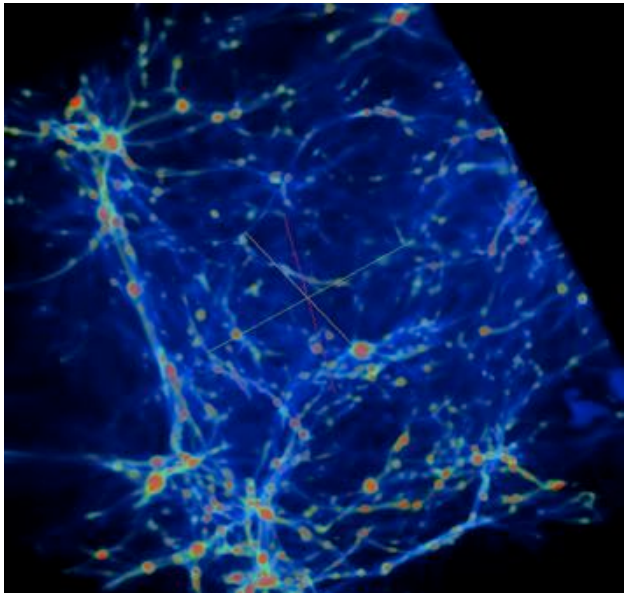
Color Hue

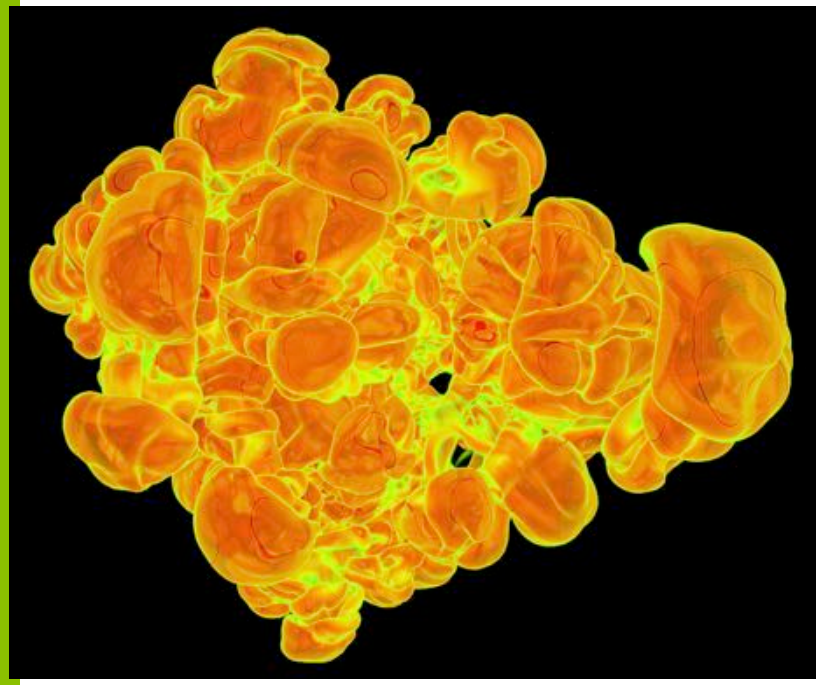
Usually referred to as color

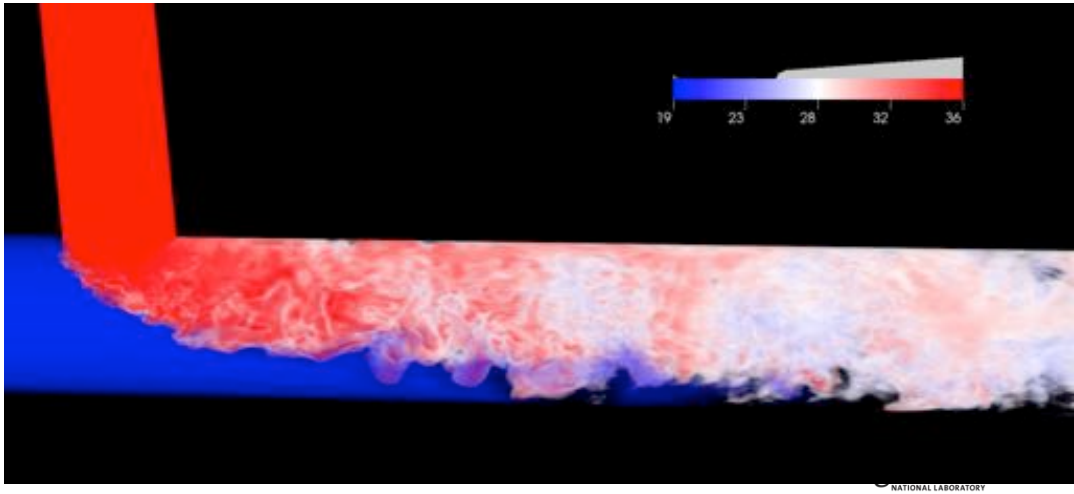


DATA REPRESENTATIONS: VOLUME RENDERING



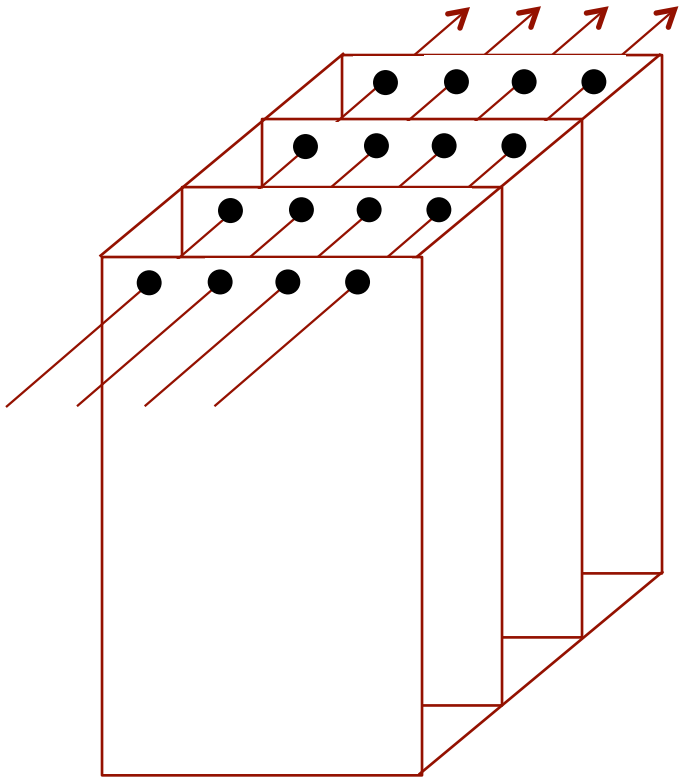


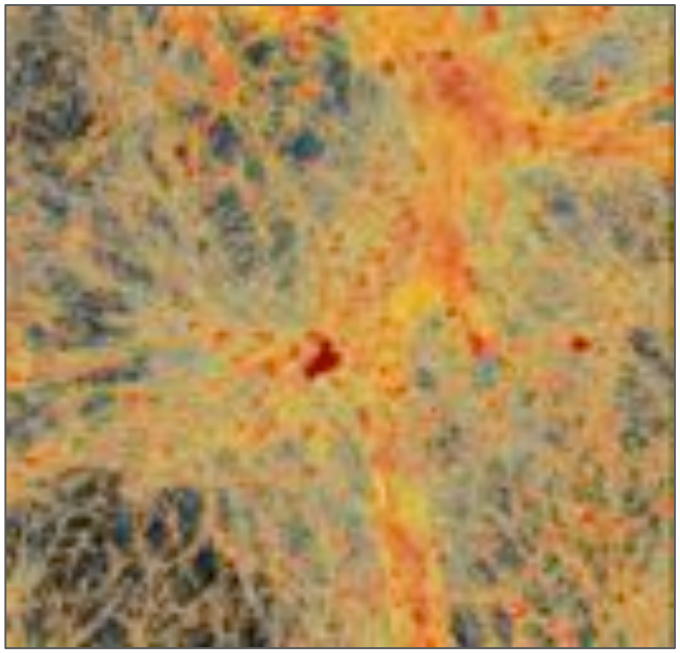




DATA REPRESENTATIONS: VOLUME RENDERING

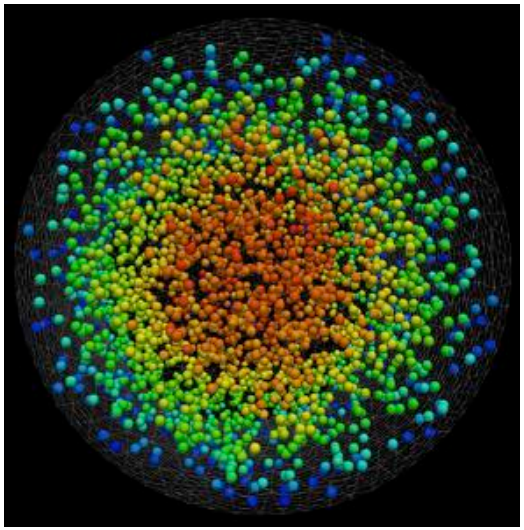
- Turn 2- and 3-dimensionsal datasets into 2D images
- Approximation: Volume ray casting

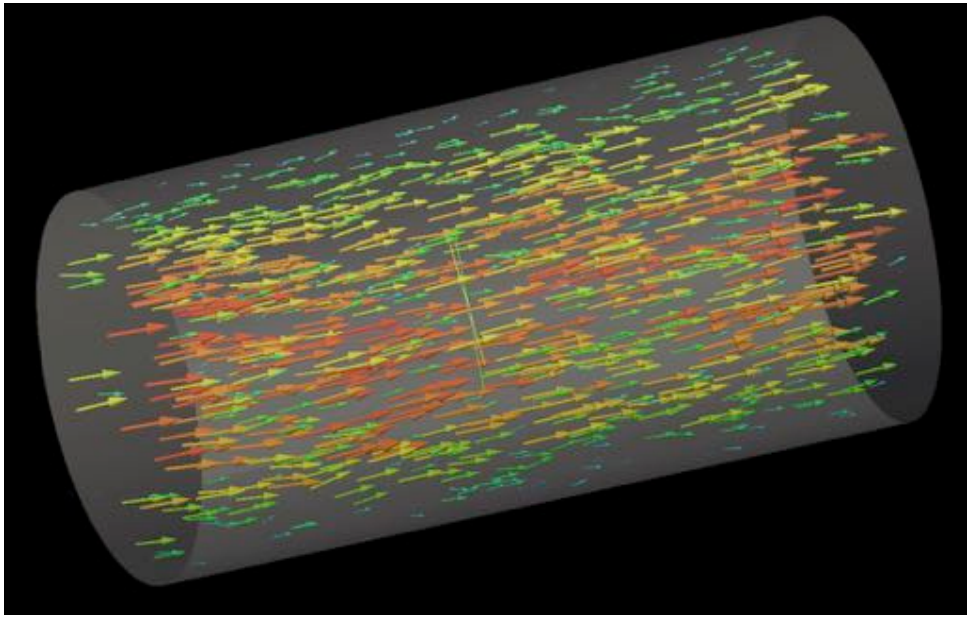


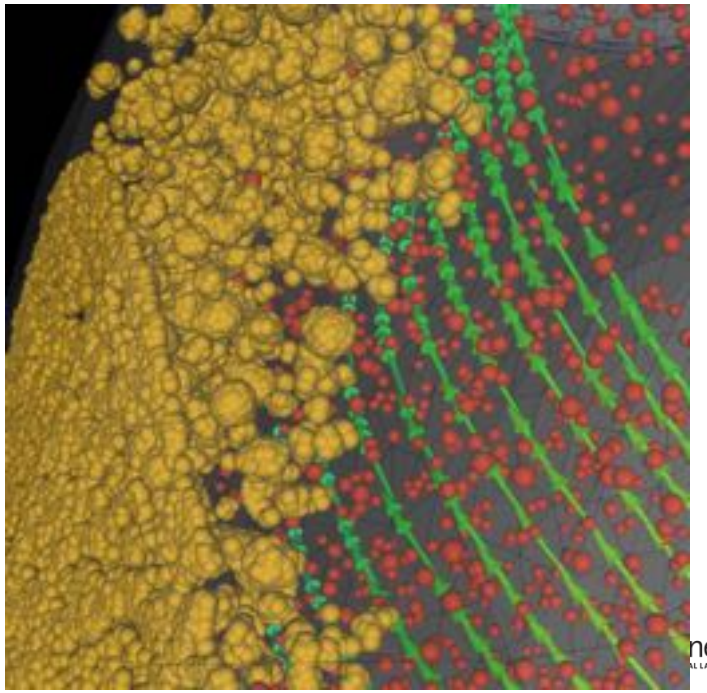


DATA REPRESENTATIONS: GLYPHS

- 2D or 3D geometric object to represent point data
- Location dictated by coordinate
  - 3D location on mesh
  - 2D position in table/graph
- Attributes graphical entity dictated by attributes of a data
  - color, size, orientation

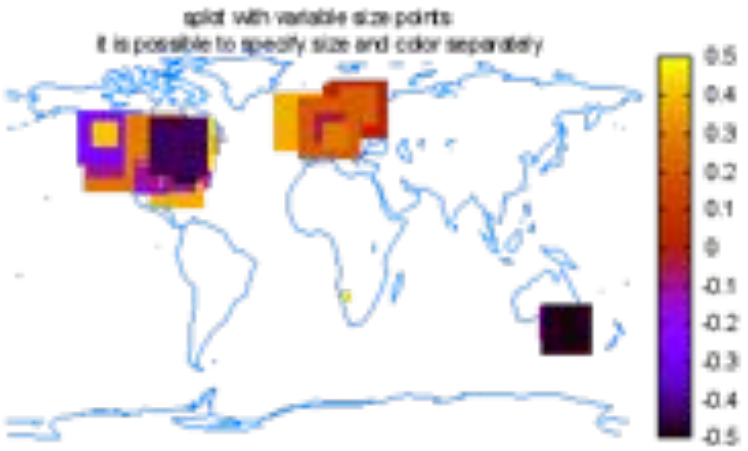


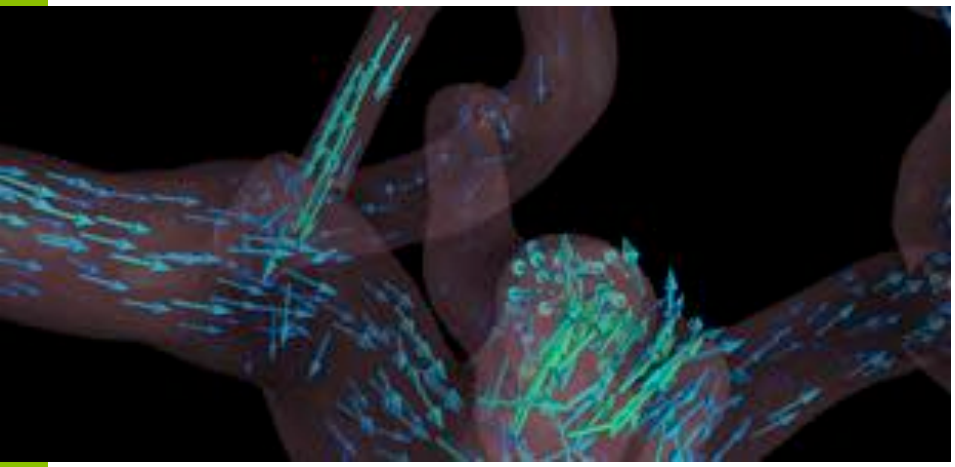


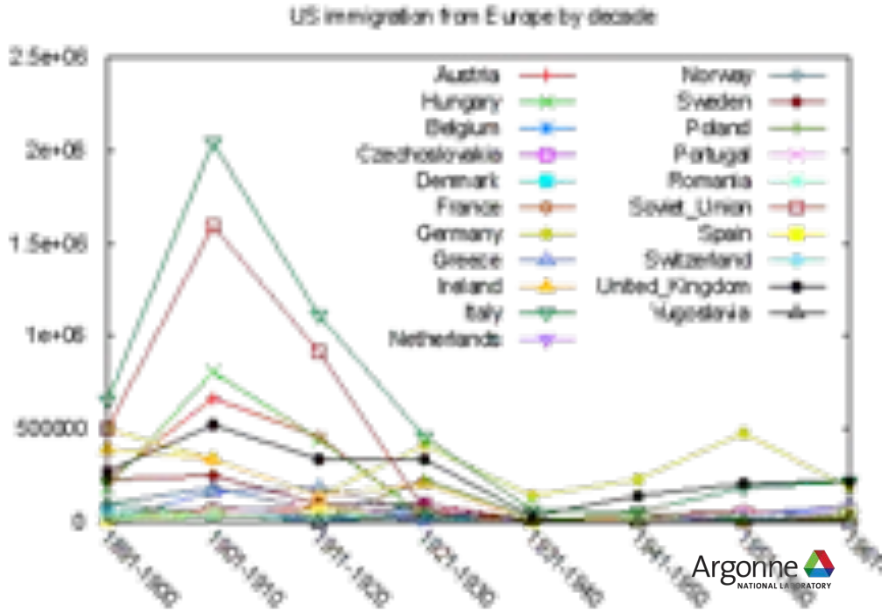


DATA REPRESENTATIONS: GLYPHS

- VisIt & ParaView:
  - good at this
- vtk:
  - same, but again requires more effort
- gnuplot:
  - good at 2D plots, tables of numbers

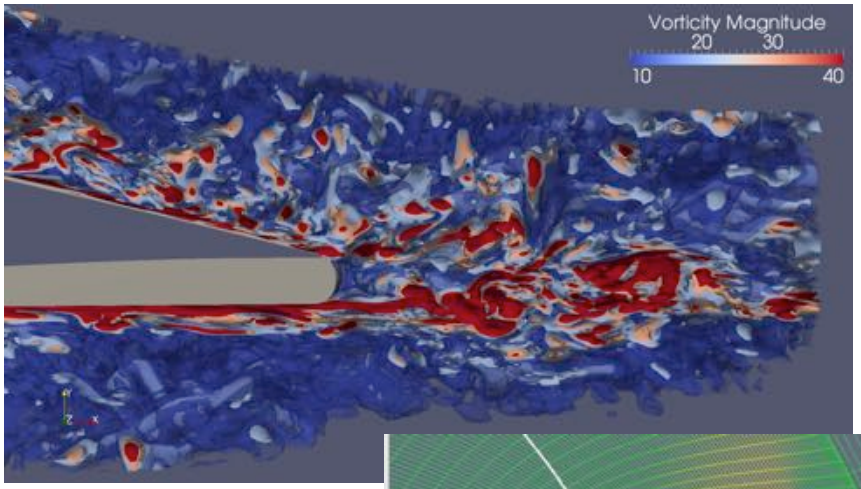


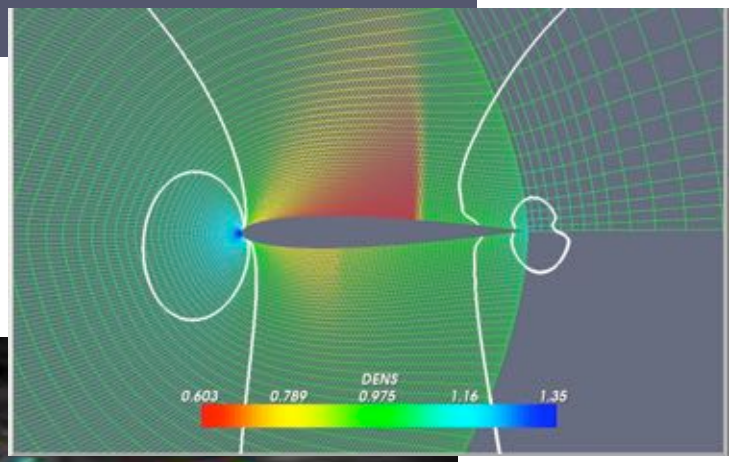


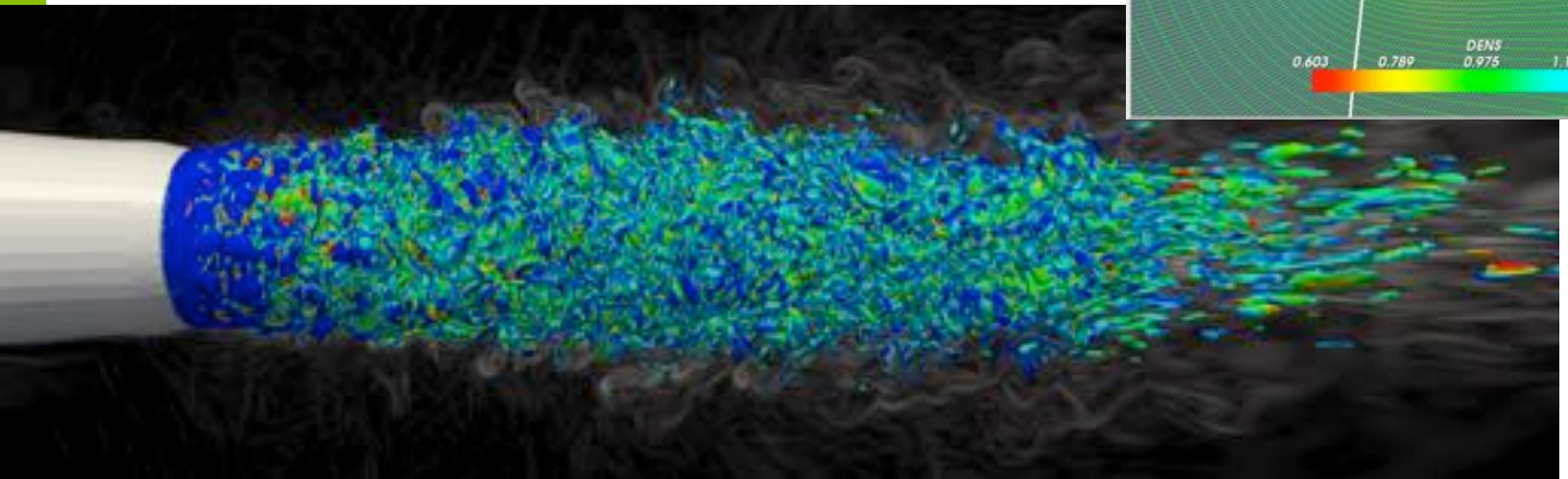


DATA REPRESENTATIONS: CONTOURS (ISOSURFACES)

- A Line (2D) or Surface (3D), representing a constant value
- VisIt & ParaView:
  - good at this
- vtk:
  - same, but again requires more effort

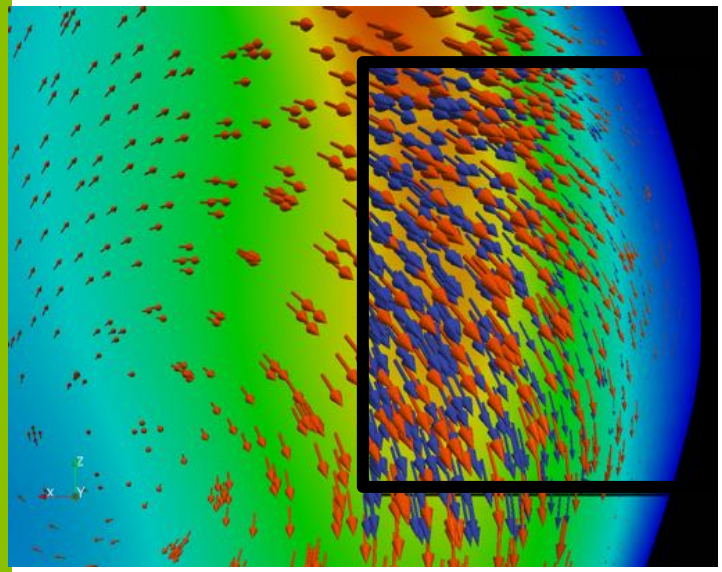


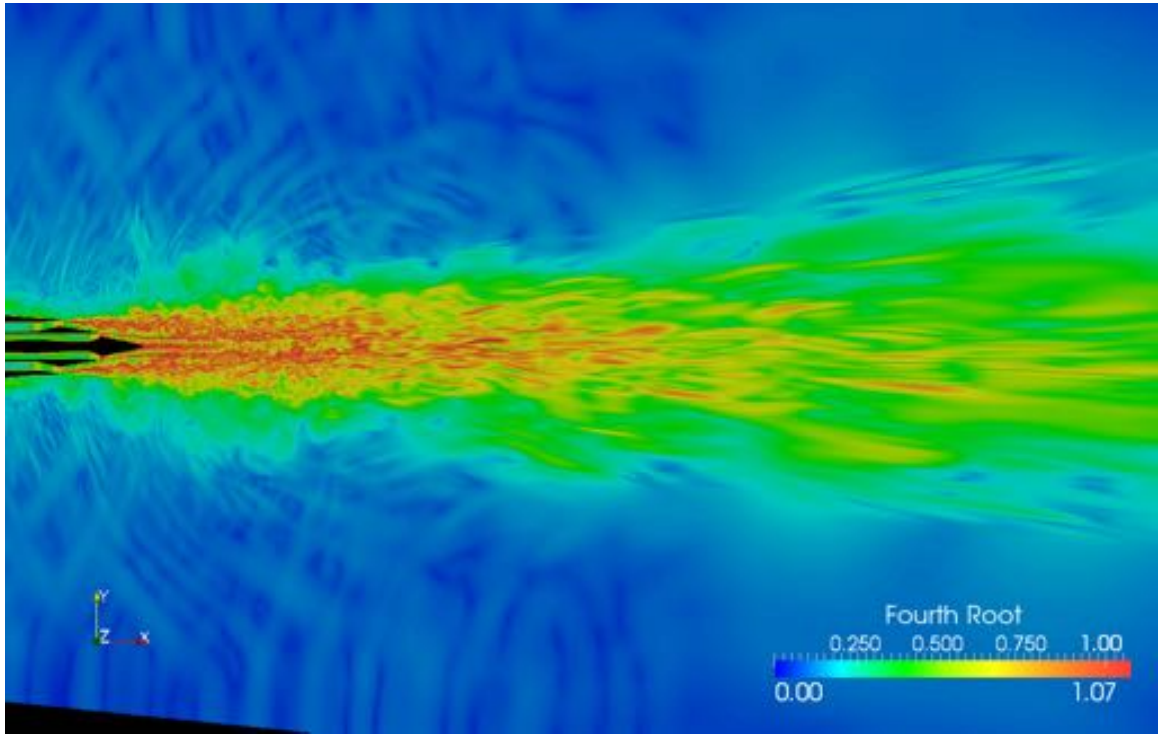




DATA REPRESENTATIONS: CUTTING PLANES

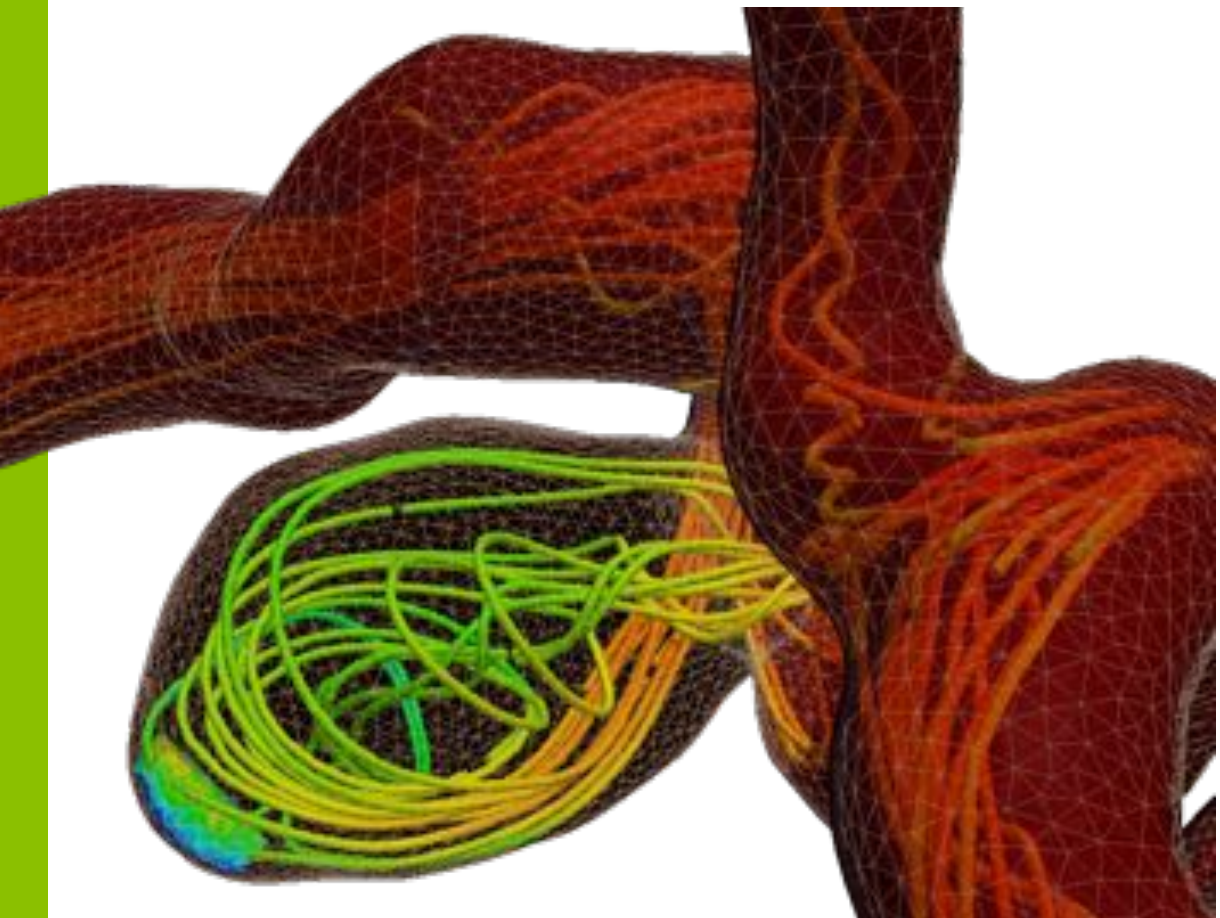
- Slice a plane through the data
  - Can apply additional visualization methods to resulting plane
- VisIt & ParaView & vtk good at this
- VMD has similar capabilities for some data formats

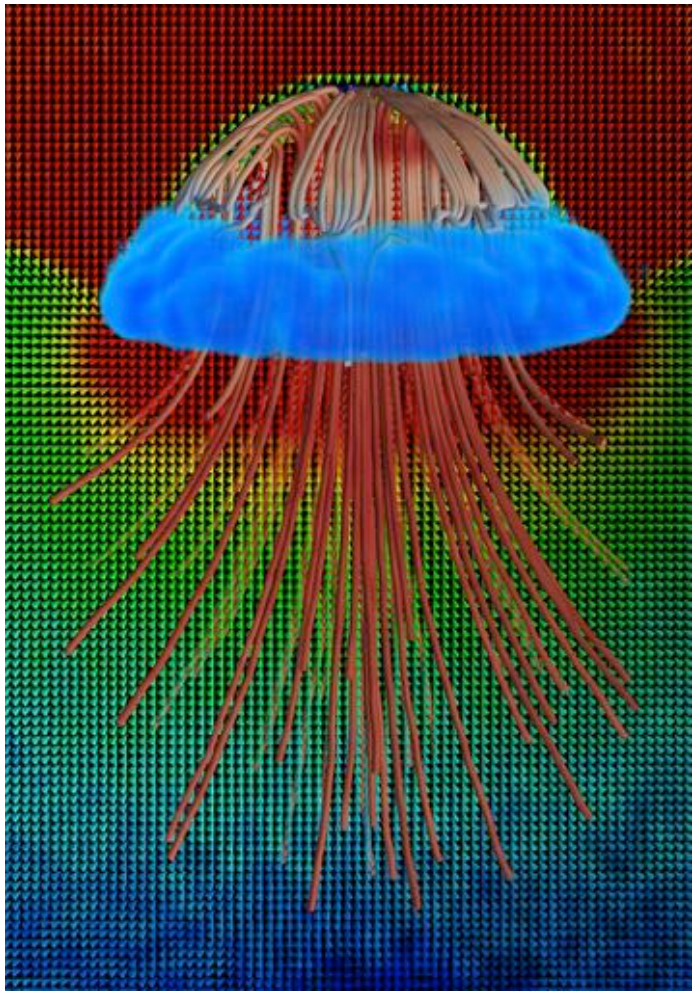




DATA REPRESENTATIONS: STREAMLINES

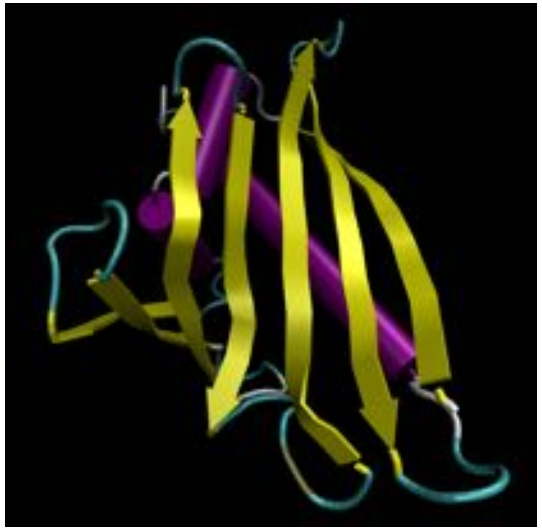
- From vector field on a mesh (needs connectivity)
  - Show the direction an element will travel in at any point in time.
- VisIt & ParaView & vtk good at this

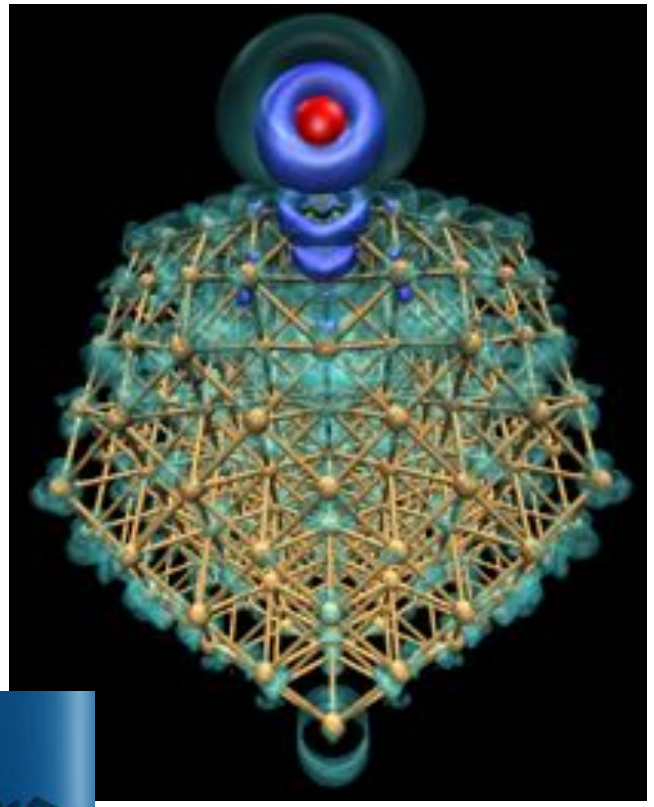


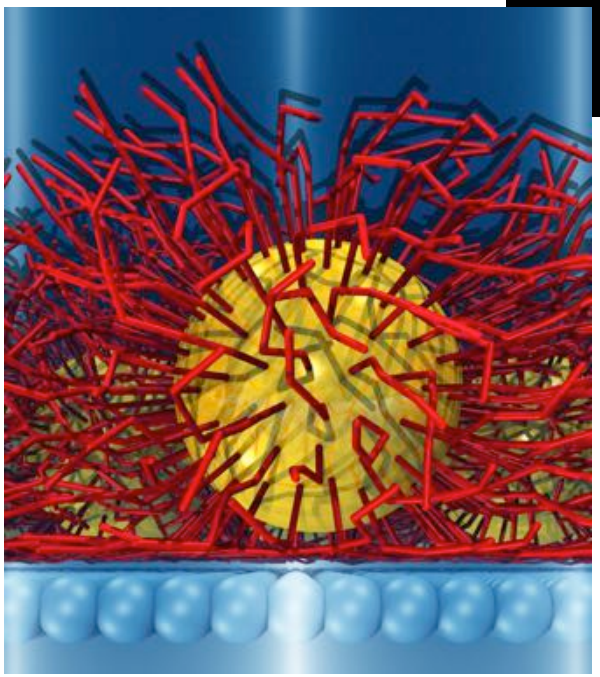


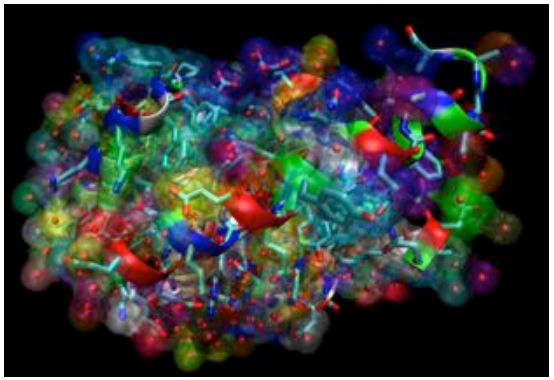
MOLECULAR DYNAMICS VISUALIZATION

- VMD:
  - Lots of domain-specific representations
  - Many different file formats
  - Animation
  - Scriptable
  - Not parallel
- VisIt & ParaView:
  - Limited support for these types of representations
- VTK:
  - Anything's possible if you try hard enough



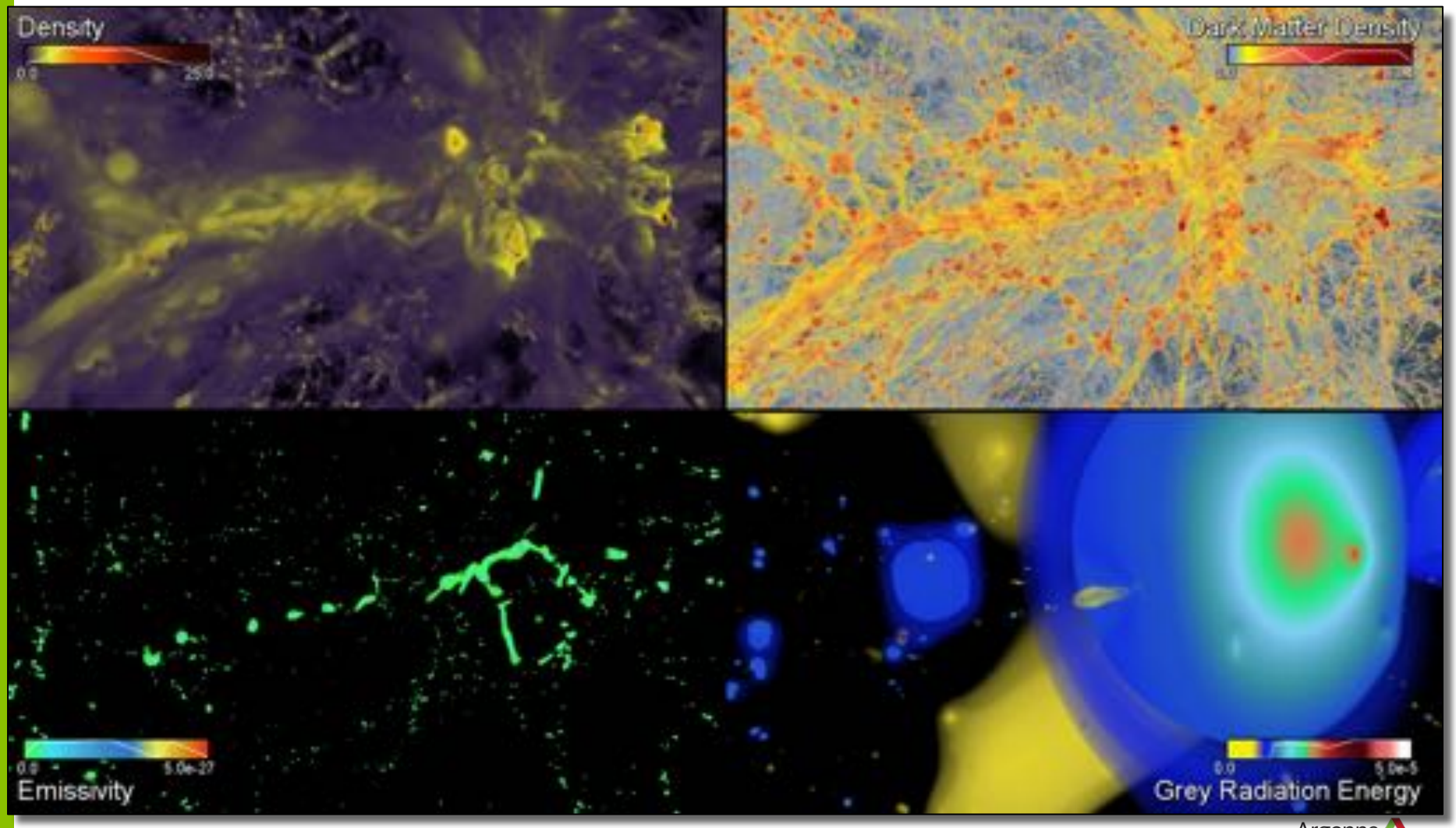






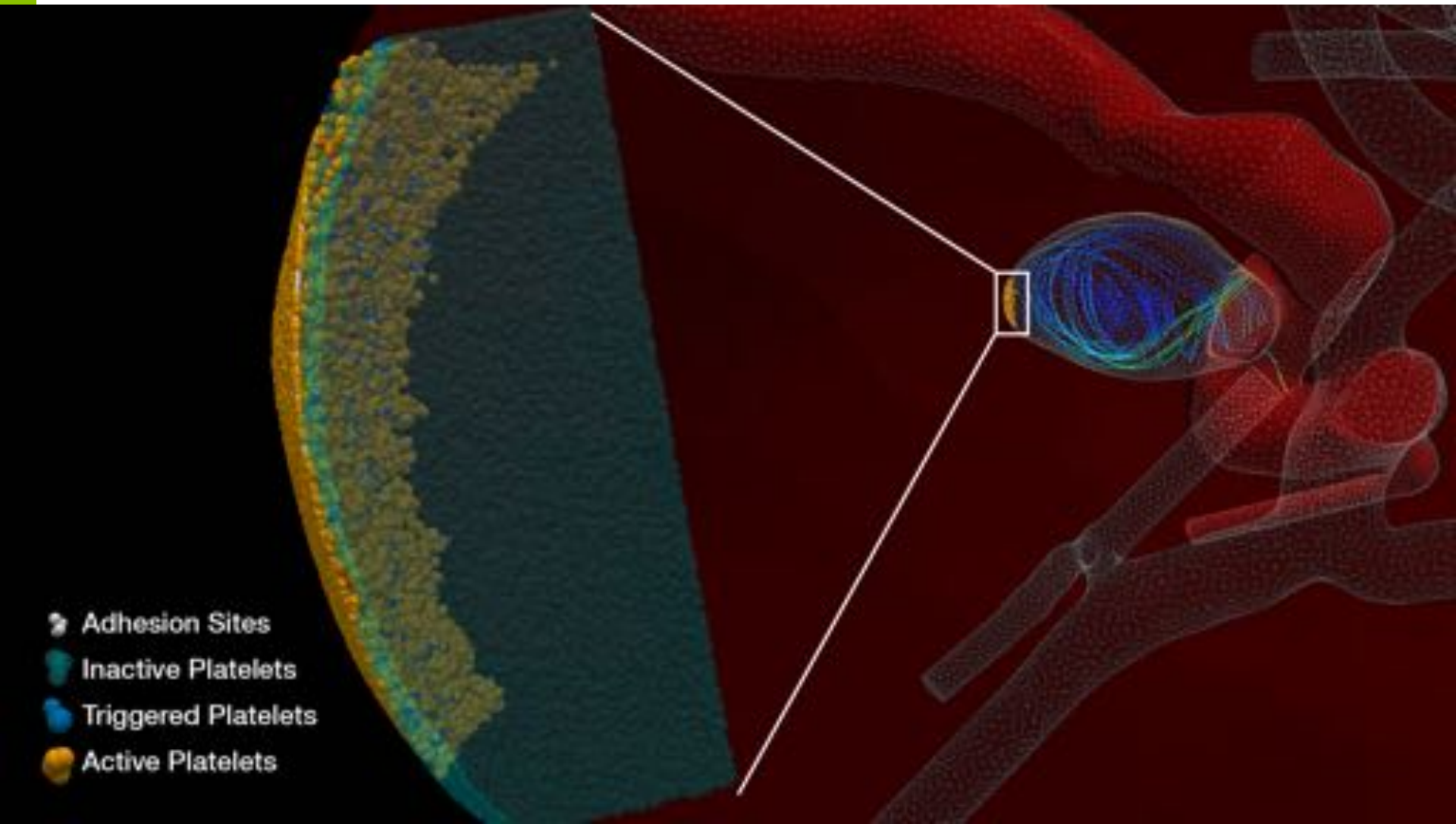
ANNOTATION, COMPOSITING, SCALING…

- ImageMagick
  - convert, composite, montage, etc.



ANNOTATION, COMPOSITING, SCALING…

- ImageMagick
  - scale, fade



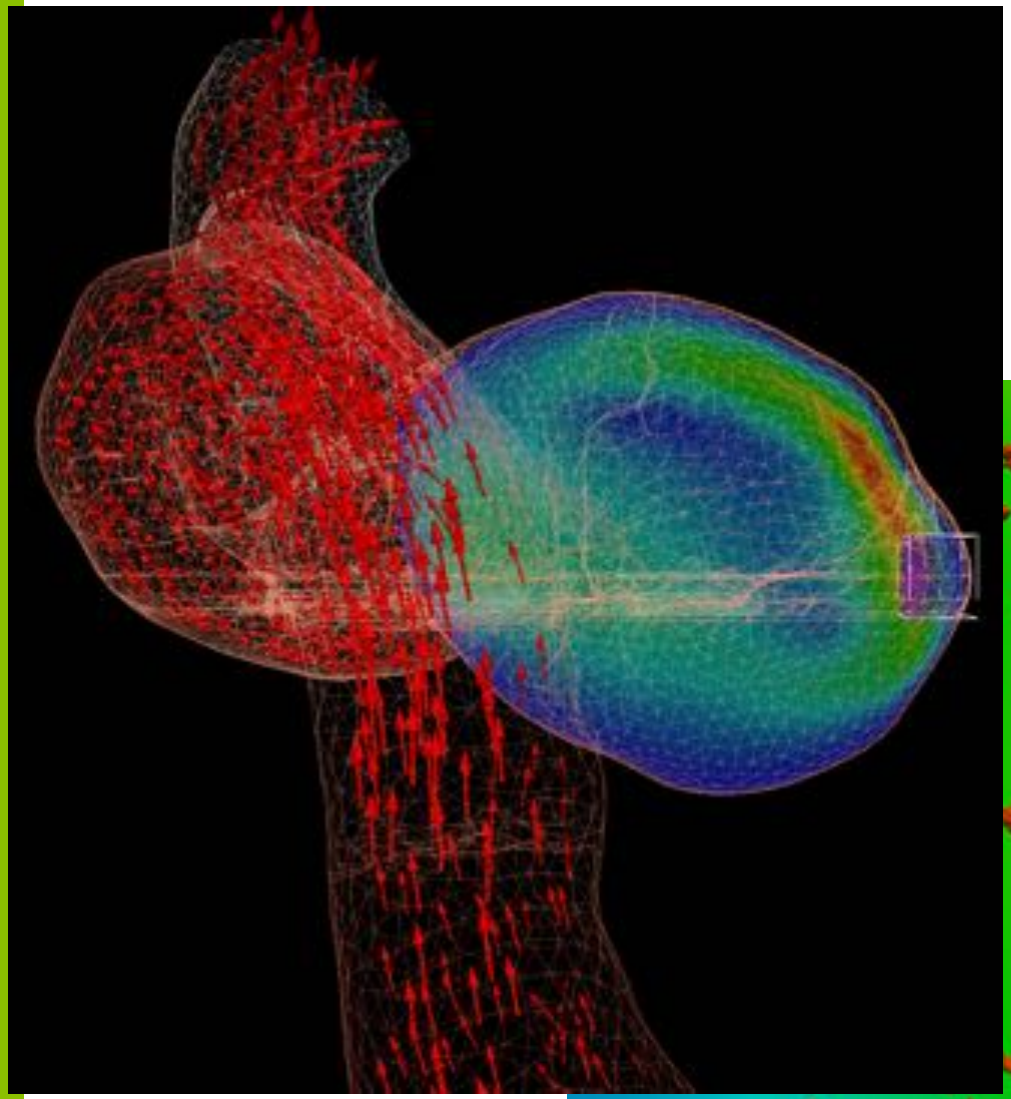
MOVIE CREATION

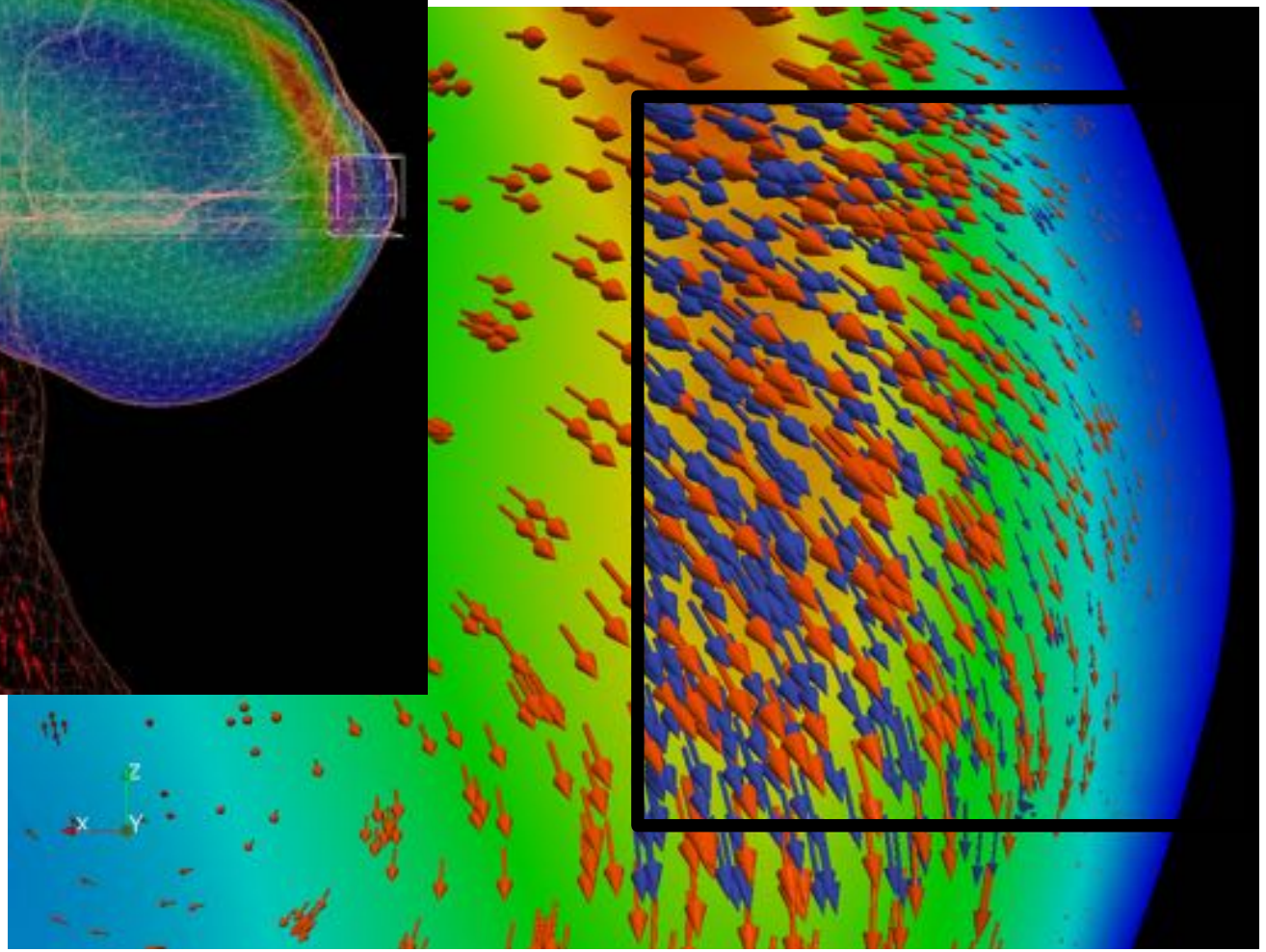
- VisIt and ParaView can spit out a movie file (.avi, etc.)
  - can also spit out individual images

- Combine multiple segments of frames
  - Create a directory of symbolic links to all frames in order

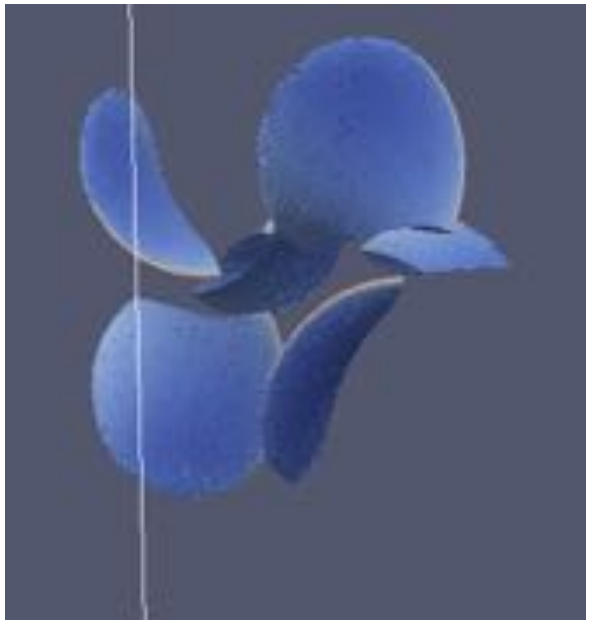
- ffmpeg: Movie encoding
  - ffmpeg –sameq –i frame.%04d.png movie.mp4

VISUALIZATION FOR VERIFICATION

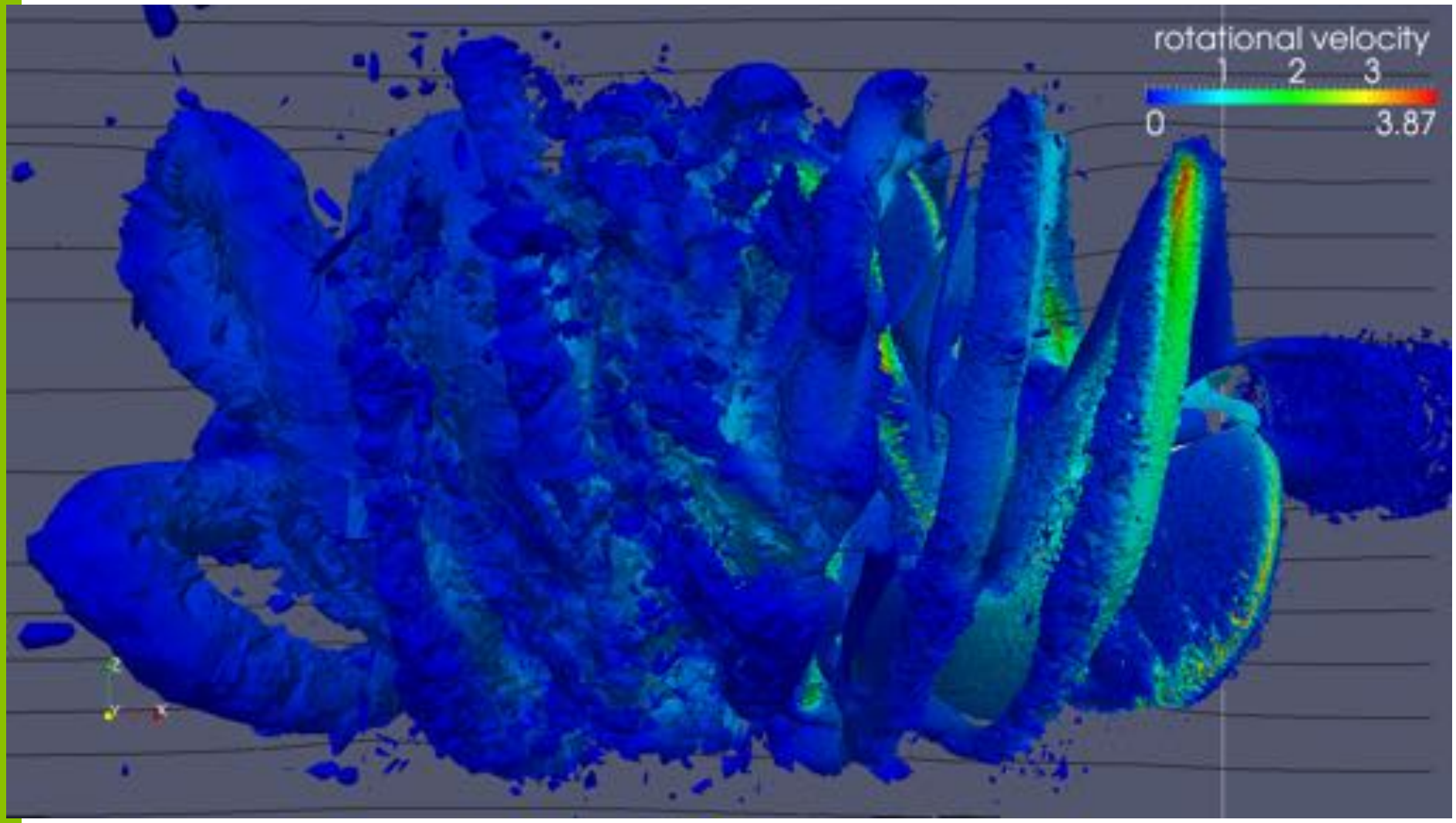




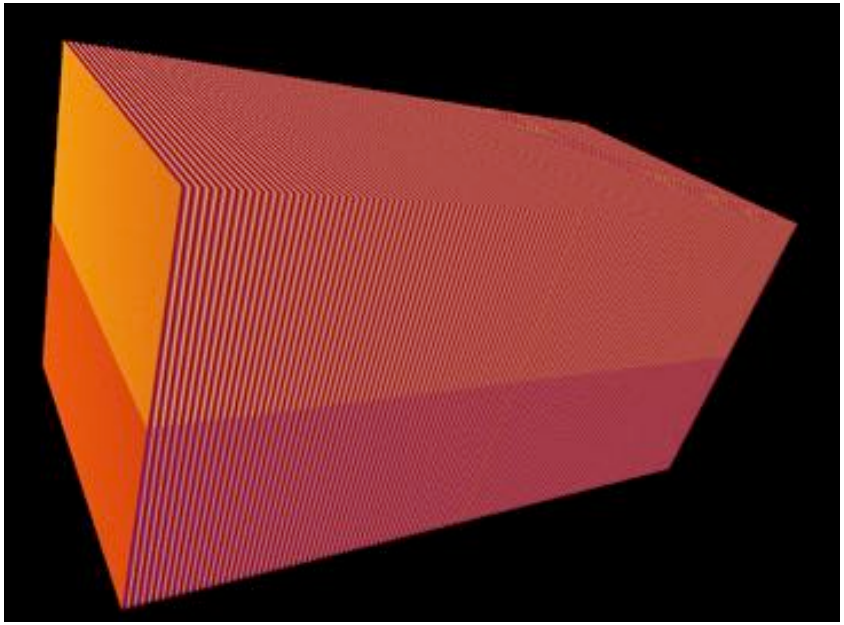
VISUALIZATION FOR DEBUGGING

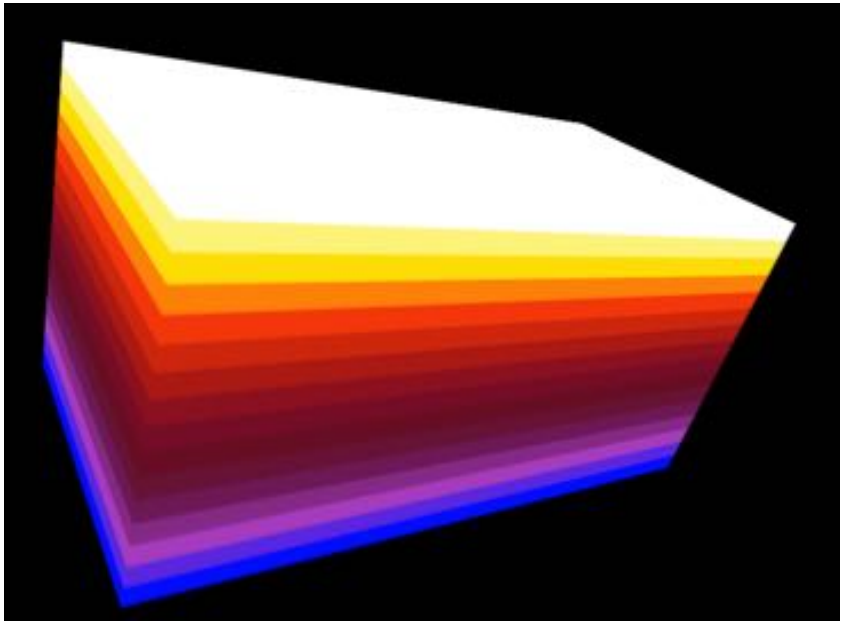



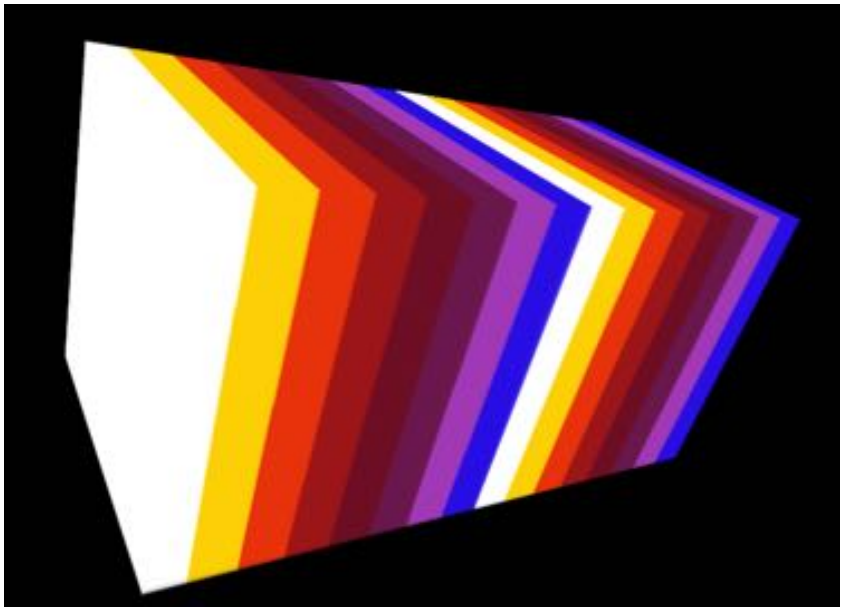

VISUALIZATION FOR DEBUGGING

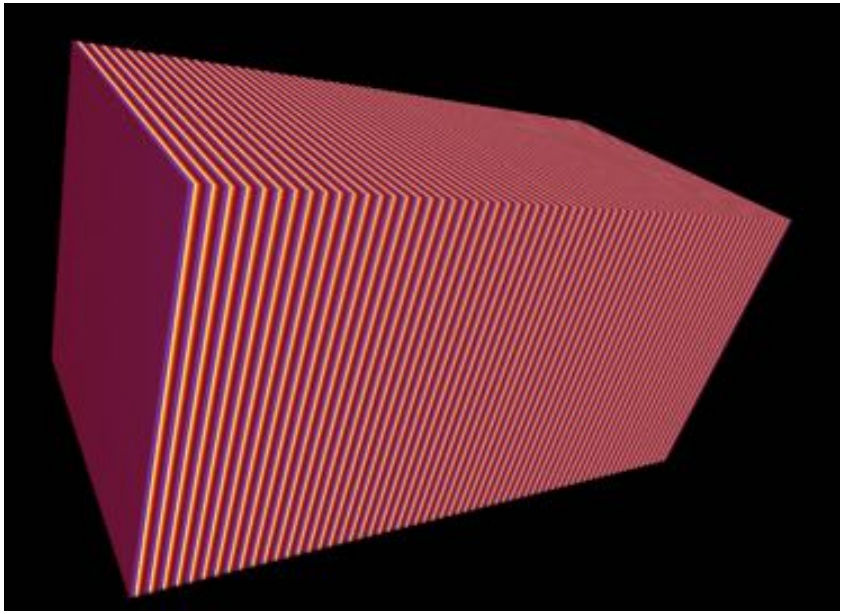


VISUALIZATION AS DIAGNOSTICS: COLOR BY THREAD ID









MORE INFO…

- www.alcf.anl.gov/user-guides/cooley
- www.visitusers.org
- paraview.org/Wiki/ParaView
- www.imagemagick.org