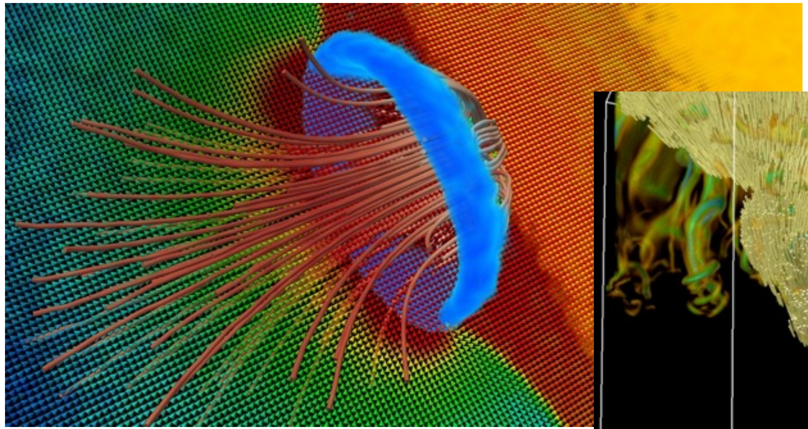
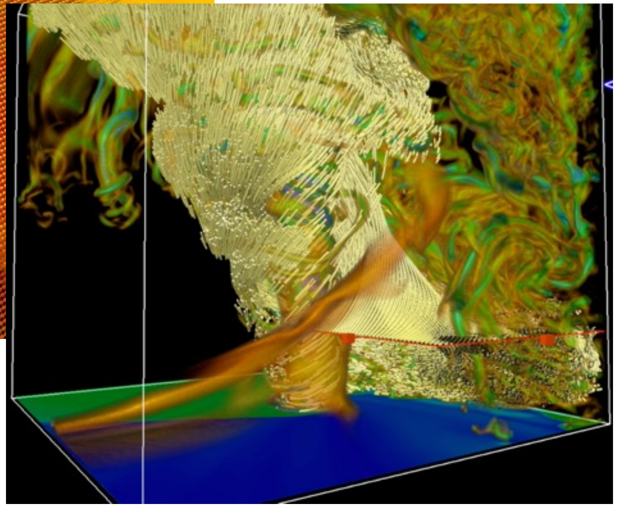
ARGONNE

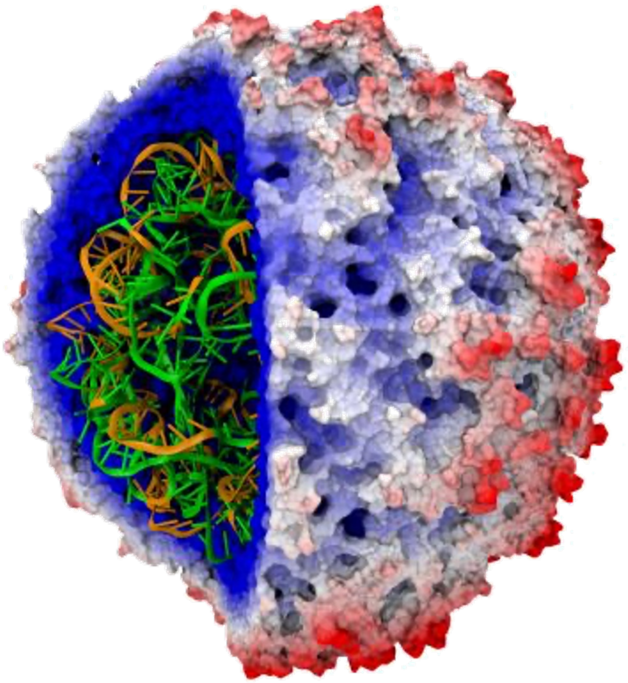
ATPESC2023

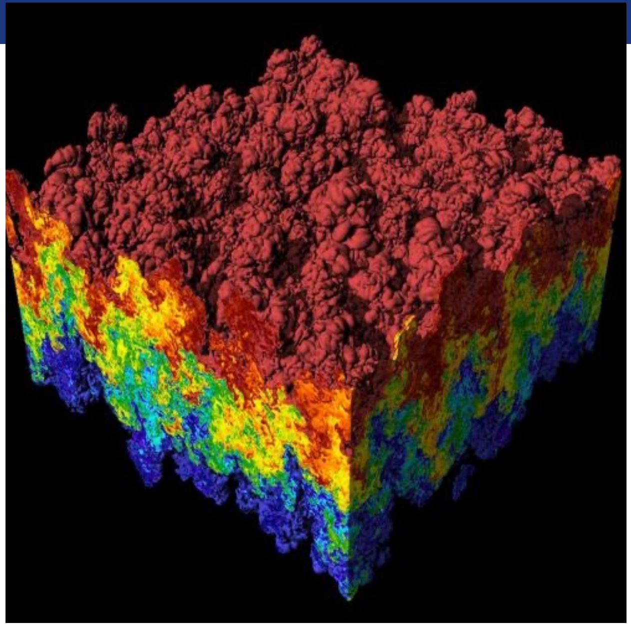
EXTREME-SCALE COMPUTING

Data Analysis and Visualization









extremecomputingtraining.anl.gov

ECP EXASCALE COMPUTING PROJECT

Argonne NATIONAL LABORATORY

Visualization & Data Analysis

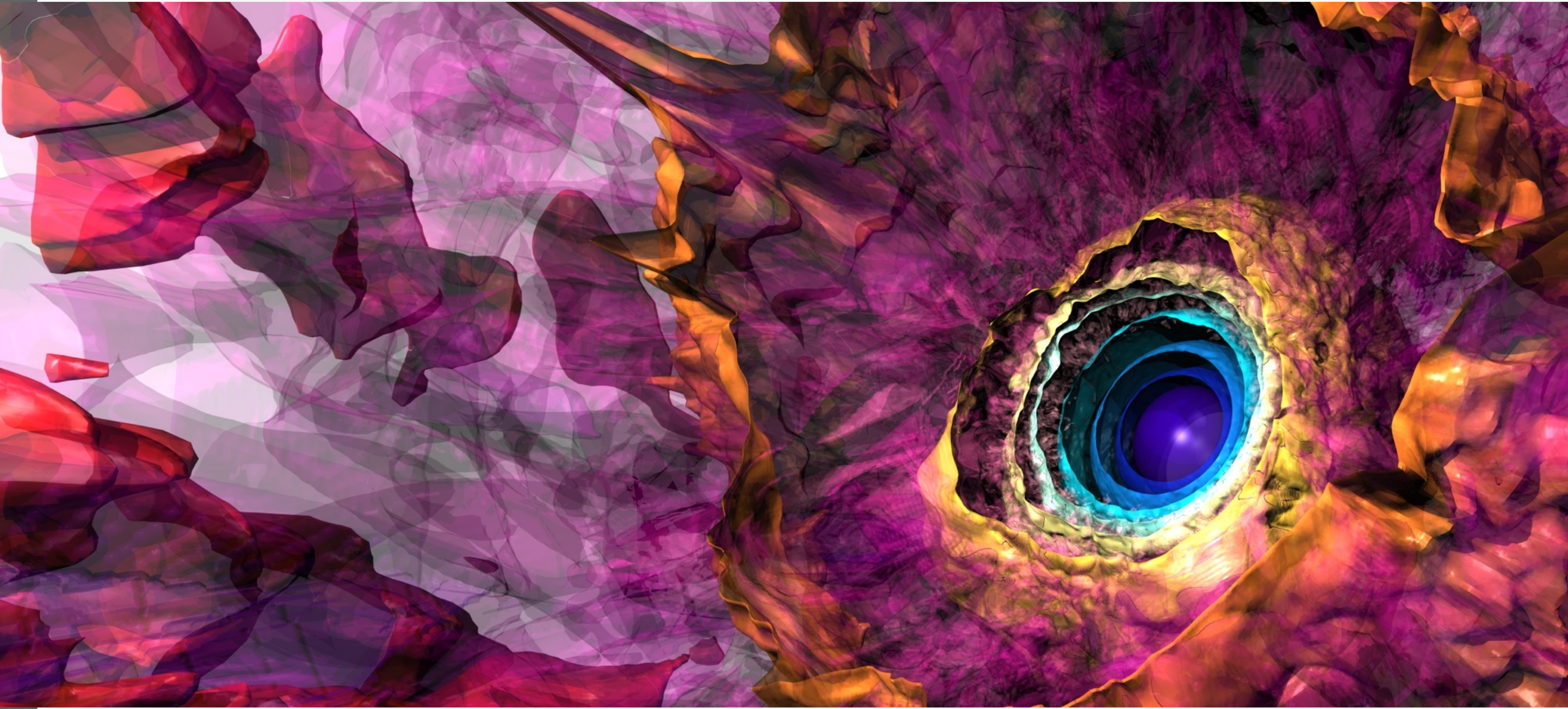
Time	Title of presentation	Lecturer
8:30 am	**Data Analysis and Visualization Introduction**	Joe Insley *ANL/NIU*, Silvio Rizzi *ANL,* Victor Mateevitsi, *ANL*
9:15 am	**Scalable Molecular Visualization and Analysis Tools in VMD**	Alex Bryer *UD*
10:00 am	*Break*	
10:30 am	**Large Scale Visualization with ParaView**	Dan Lipsa *Kitware*
12:00 pm	**Visualization and Analysis of HPC Simulation Data with VisIt**	Cyrus Harrison *LLNL*
12:30 pm	*Lunch*	
1:30 pm	**Visualization and Analysis of HPC Simulation Data with VisIt (Cont.)**	Cyrus Harrison *LLNL*
2:30 pm	**Vapor**	Scott Pearse *NCAR*
3:30 pm	*Break*	
4:00 pm	**Exploring Visualization with Jupyter Notebooks**	• David Koop *NIU* • Cyrus Harrison *LLNL*
5:30 pm	*Hands-on*	All
6:30 pm	*Dinner*	
7:30 pm	*After-dinner talk:* Growing up at Argonne National Lab	Jack Dongarra *UT*

Here’s the plan…

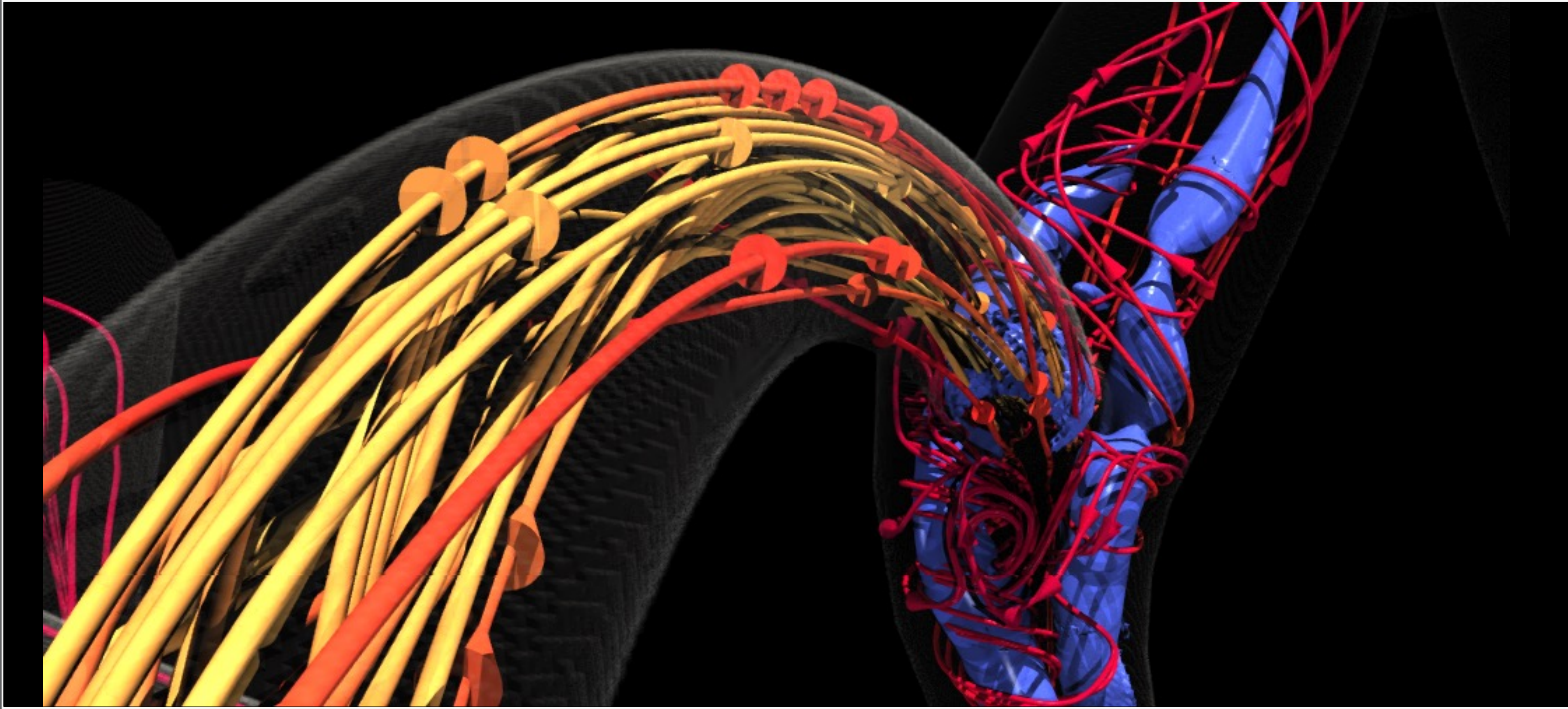
- Examples of visualizations
- Visualization tools and formats
- Data representations
- Visualization for debugging
- Advanced Rendering
- In Situ Visualization and Analysis

Physics: Stellar Radiation

Data courtesy of: Lars Bildsten and Yan-Fei Jiang, University of California at Santa Barbara

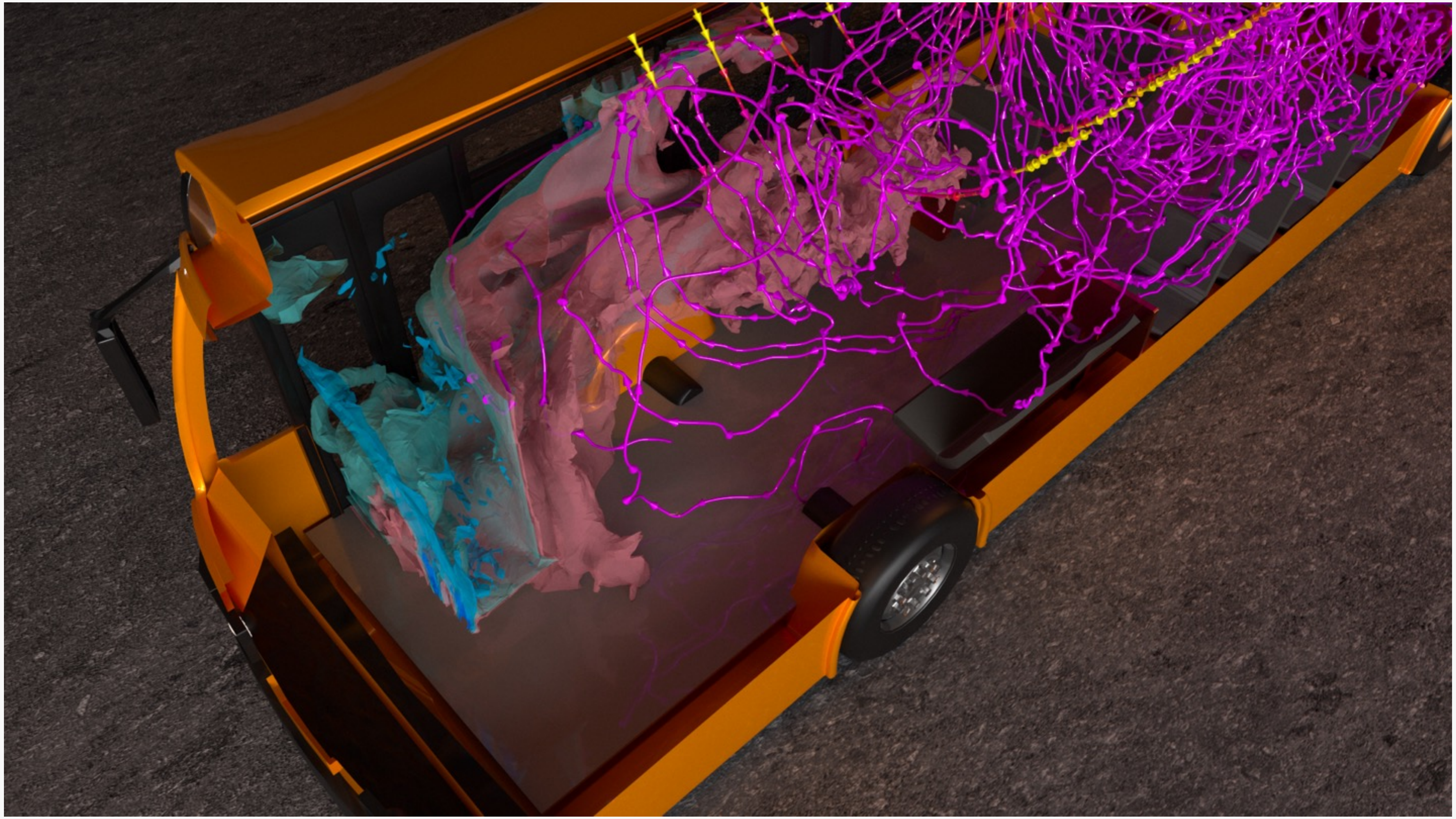


ARTERIAL BLOOD FLOW



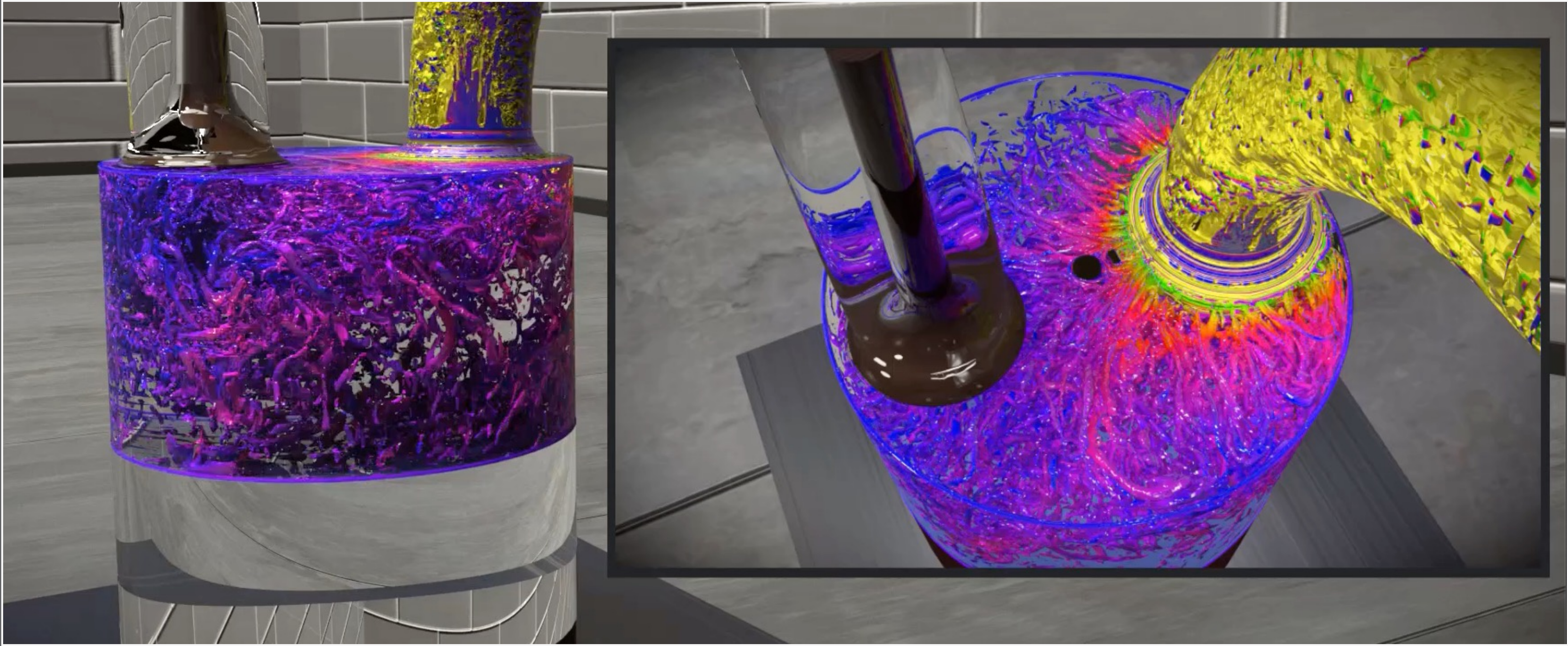
Data courtesy of: Amanda Randles, Duke University

Computational Fluid Dynamics



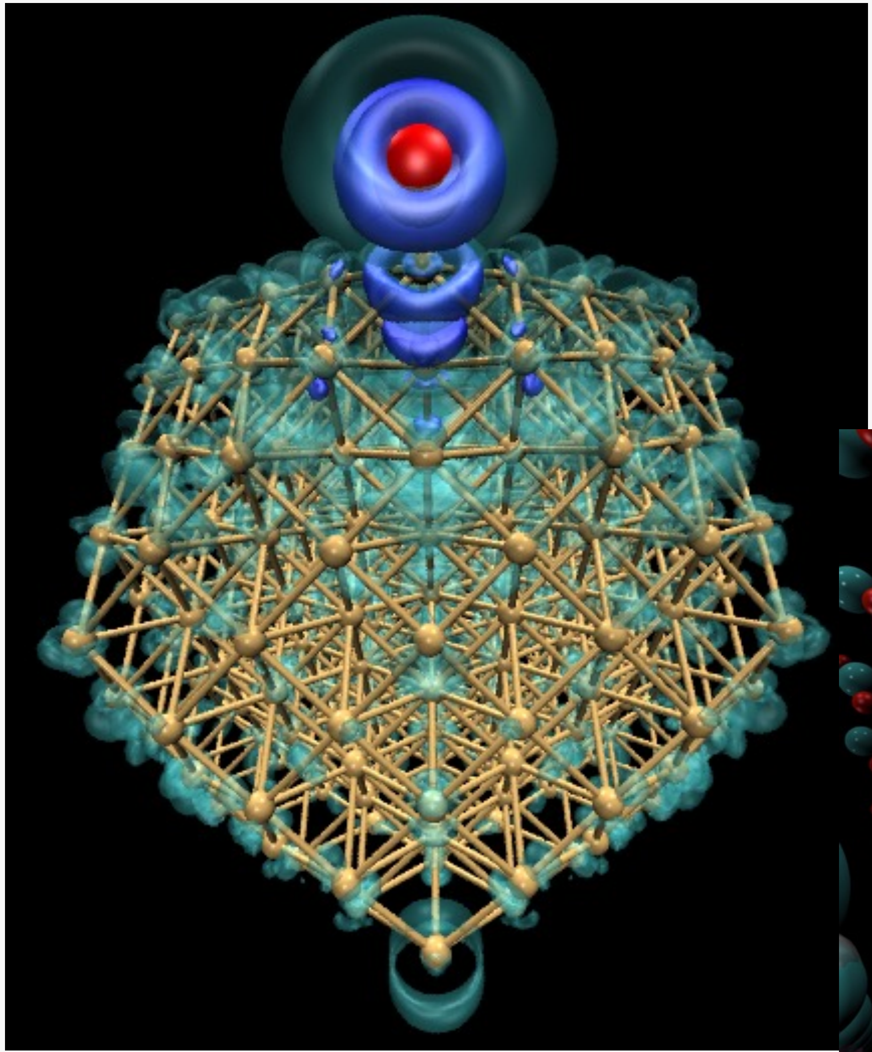
Data courtesy of Rao Kotamarthi, Ramesh Balakrishnan, Aleks Obabko, Argonne National Laboratory

Engineering Technologies: Combustion



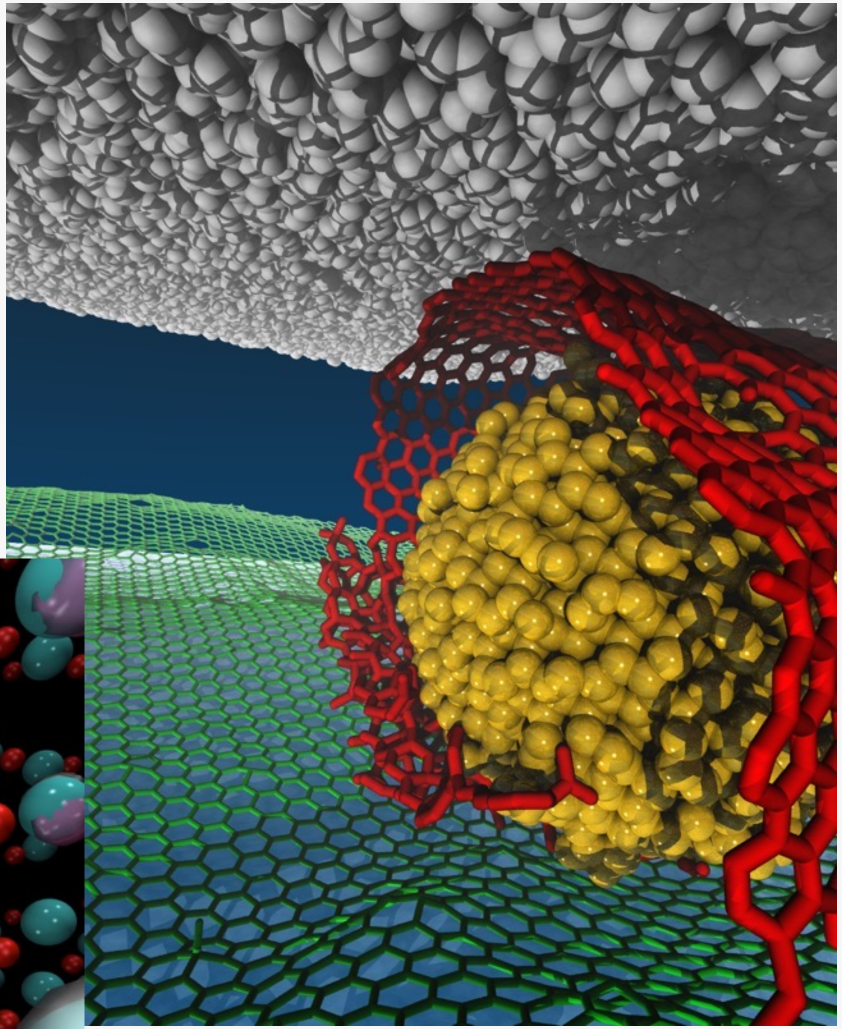
Data courtesy of: Saumil Patel, Muhsin Ameen, Sicong Wu, Argonne National Laboratory;
Tanmoy Chatterjee, GE Global Research

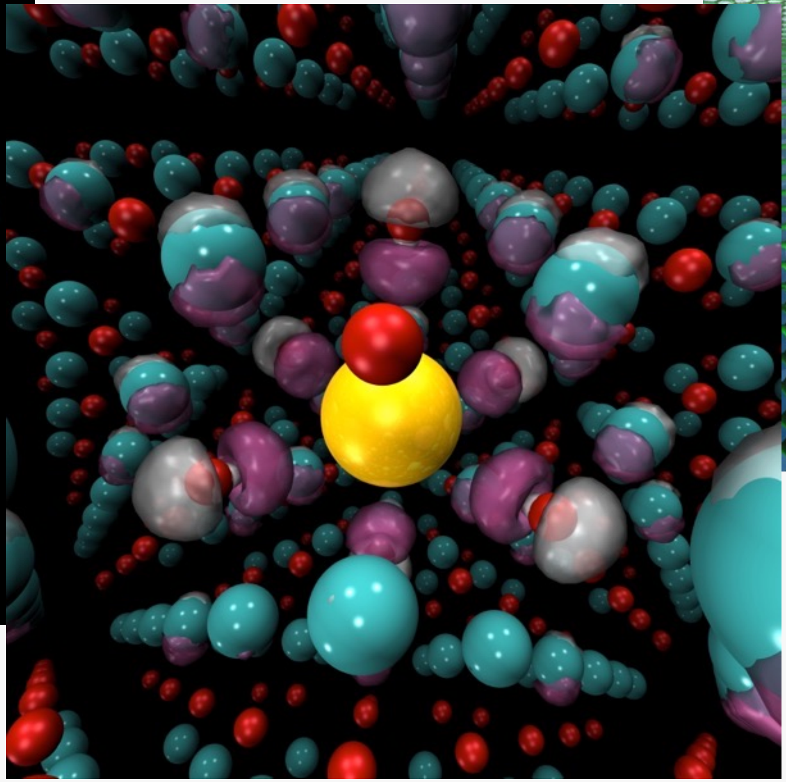
Materials Science / Molecular



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

Data courtesy of: Subramanian Sankaranarayanan, Argonne National Laboratory





Data courtesy of: Paul Kent, Oak Ridge National Laboratory, Anouar Benali, Argonne National Laboratory

Visualization Tools and Data Formats

All Sorts of Tools

Visualization Applications
- VisIt
- ParaView
- EnSight

Domain Specific
- VMD, PyMol, Ovito, Vapor

APIs
- VTK: visualization
- ITK: segmentation & registration

Analysis Environments
- Matlab
- Parallel R

Utilities
- GnuPlot
- ImageMagick

■ Available on Cooley

ParaView & VisIt vs. vtk

ParaView & VisIt

- General purpose visualization applications
- GUI-based
- Client / Server model to support remote visualization
- Scriptable / Extendable
- Built on top of vtk (largely)
- *In situ* capabilities

vtk

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

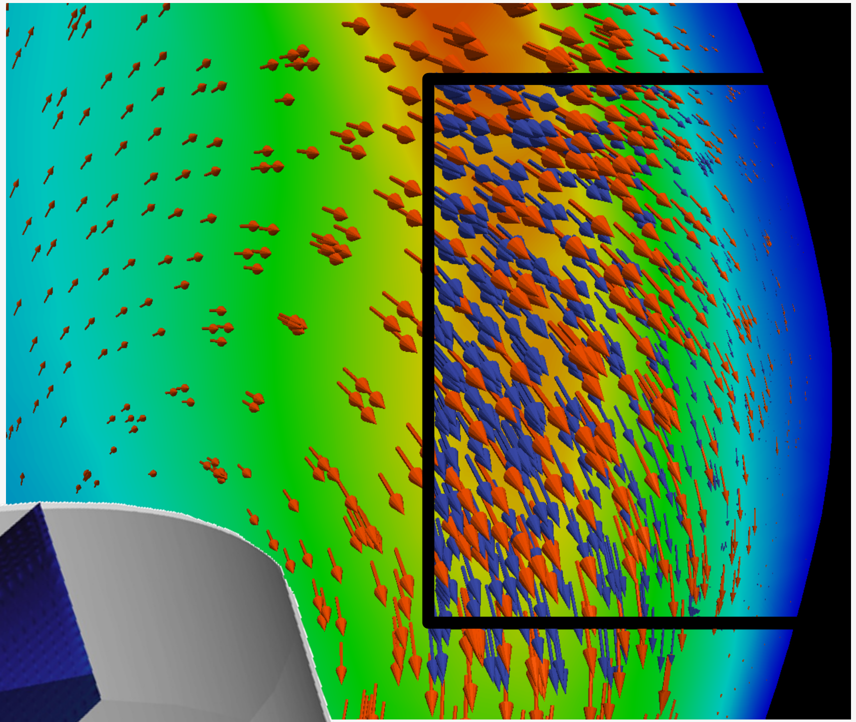
Data File Formats (ParaView & VisIt)

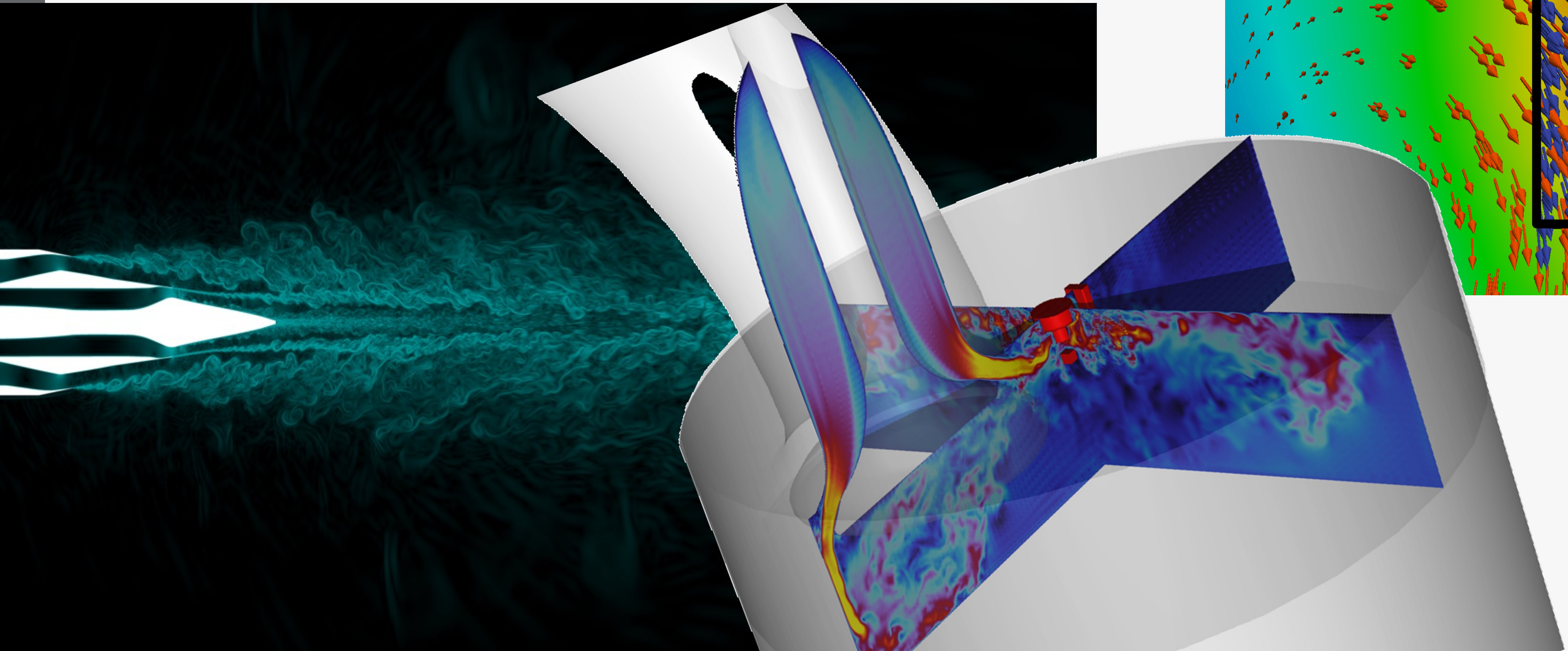
- VTK
- Parallel (partitioned) VTK
- VTK MultiBlock (MultiGroup, Hierarchical, Hierarchical Box)
- Legacy VTK
- Parallel (partitioned) legacy VTK
- EnSight files
- EnSight Master Server
- Exodus
- BYU
- XDMF
- PLOT2D
- PLOT3D
- SpyPlot CTH
- HDF5 raw image data
- DEM
- VRML
- PLY
- Polygonal Protein 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 Representations

Data Representations: Cutting Planes

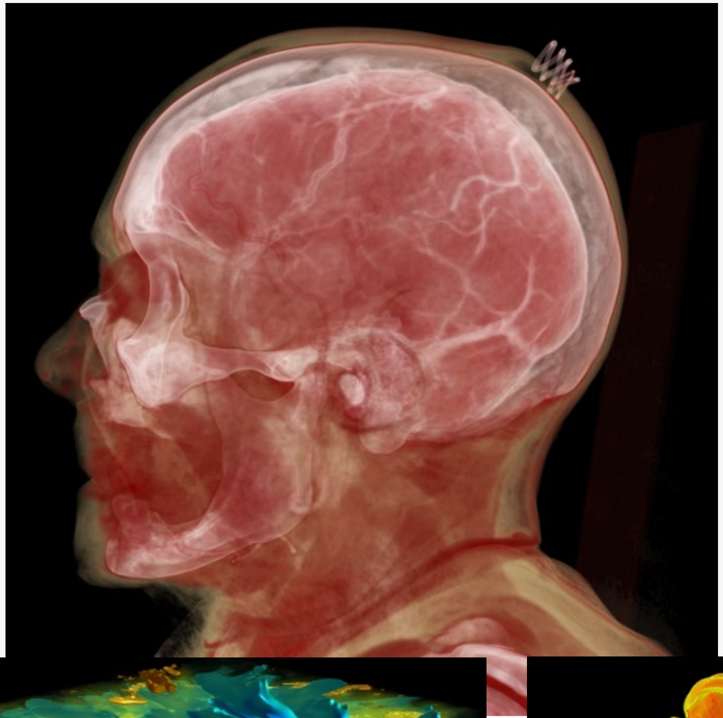
- Slice a plane through the data
  - Can apply additional visualization methods to resulting plane
- VisIt & ParaView & vtk good at this
- VMD has similar capabilities for some data formats

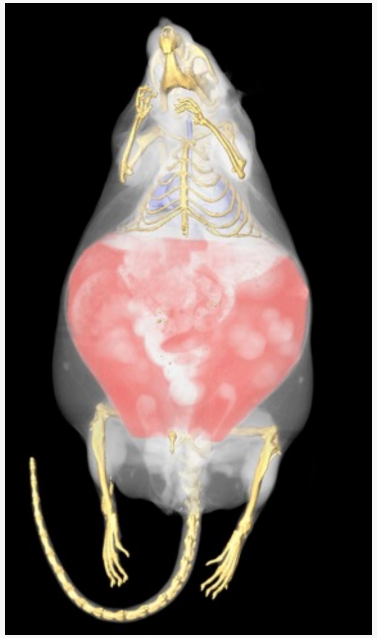


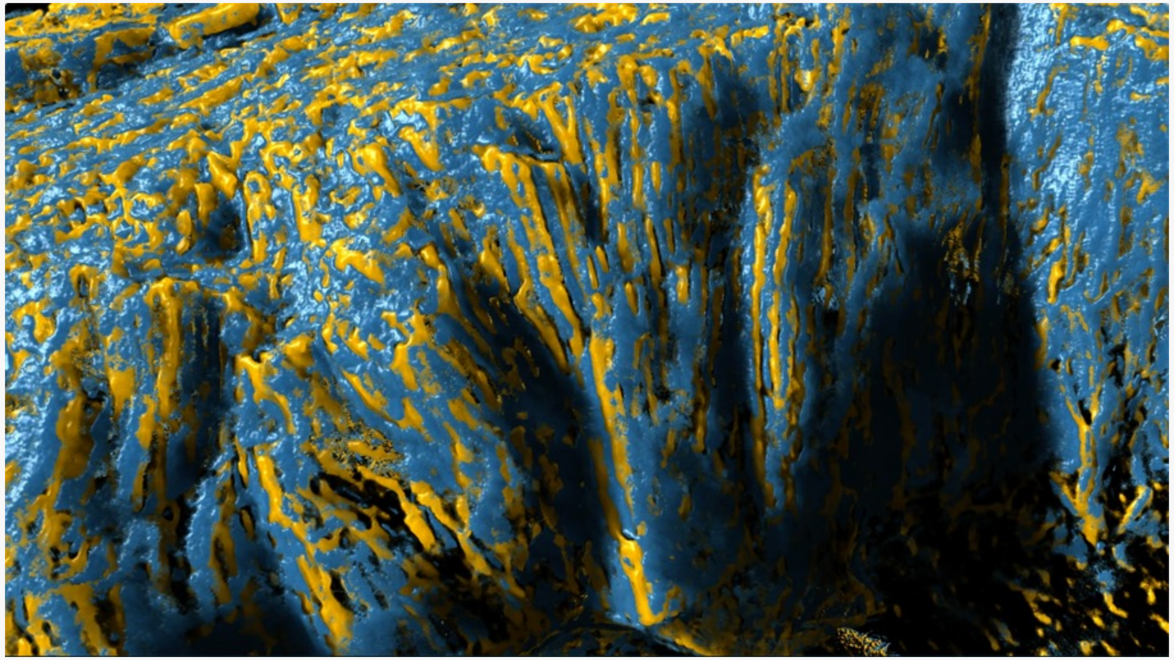


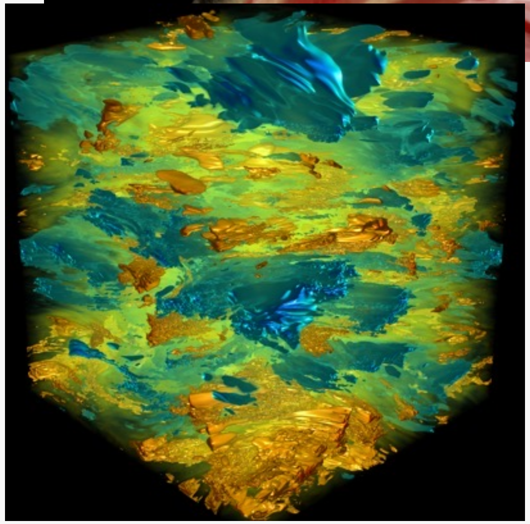
Argonne NATIONAL LABORATORY

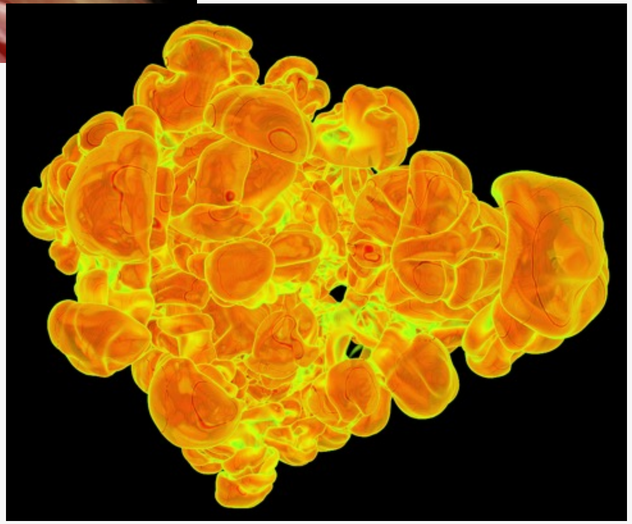
Data Representations: Volume Rendering

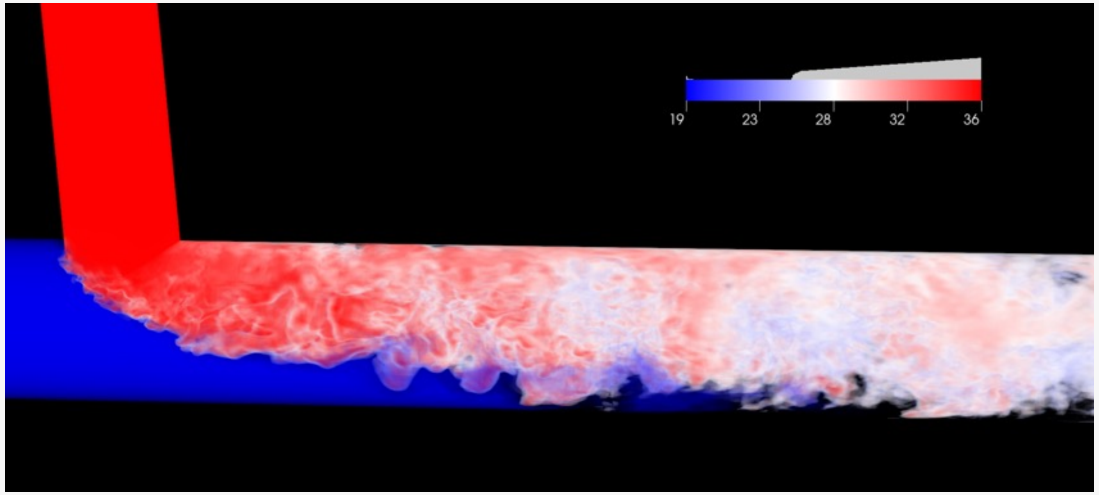












Data Representations: Contours (Isosurfaces)

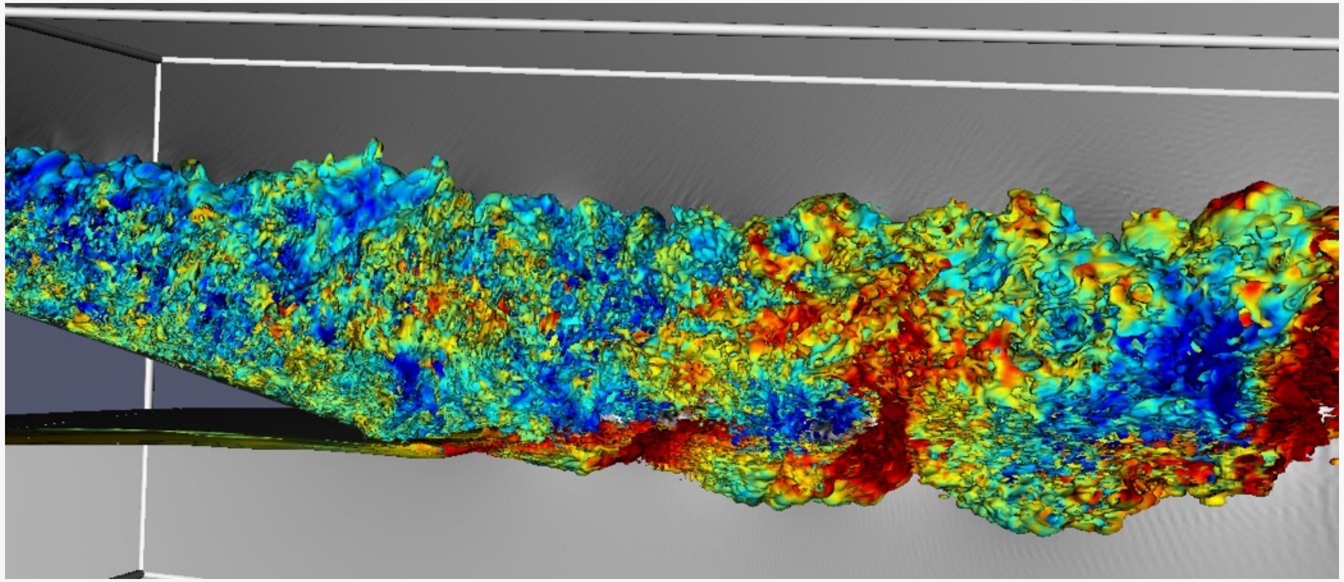
A Line (2D) or Surface (3D), representing a constant value

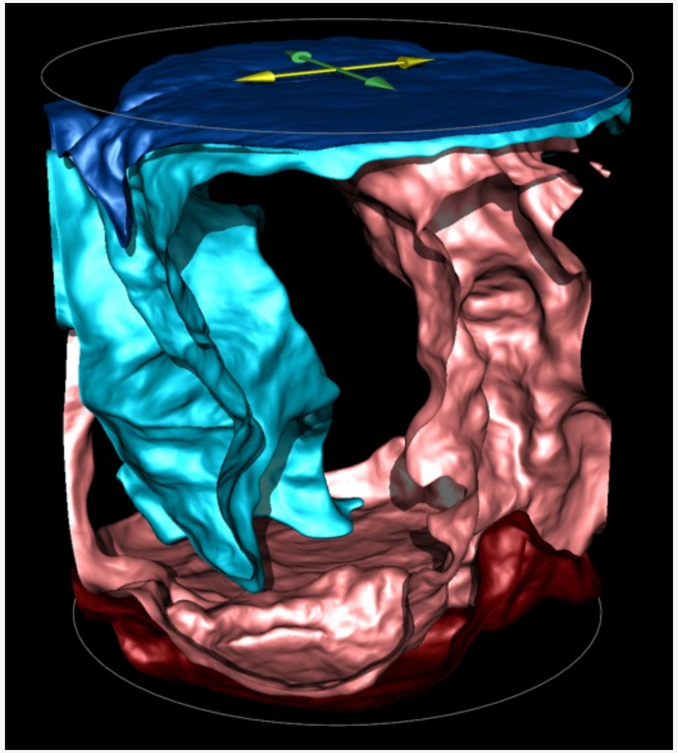
VisIt & ParaView:

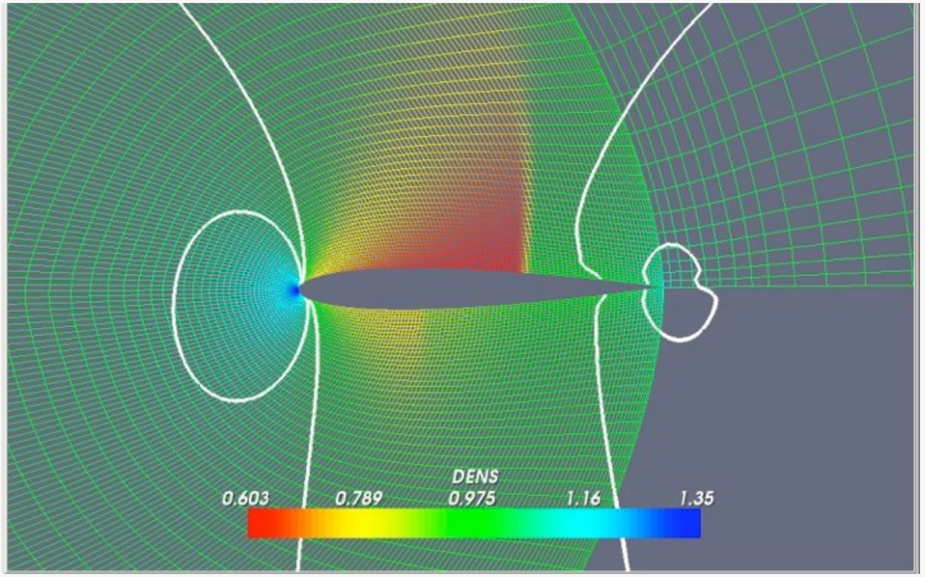
– good at this

vtk:

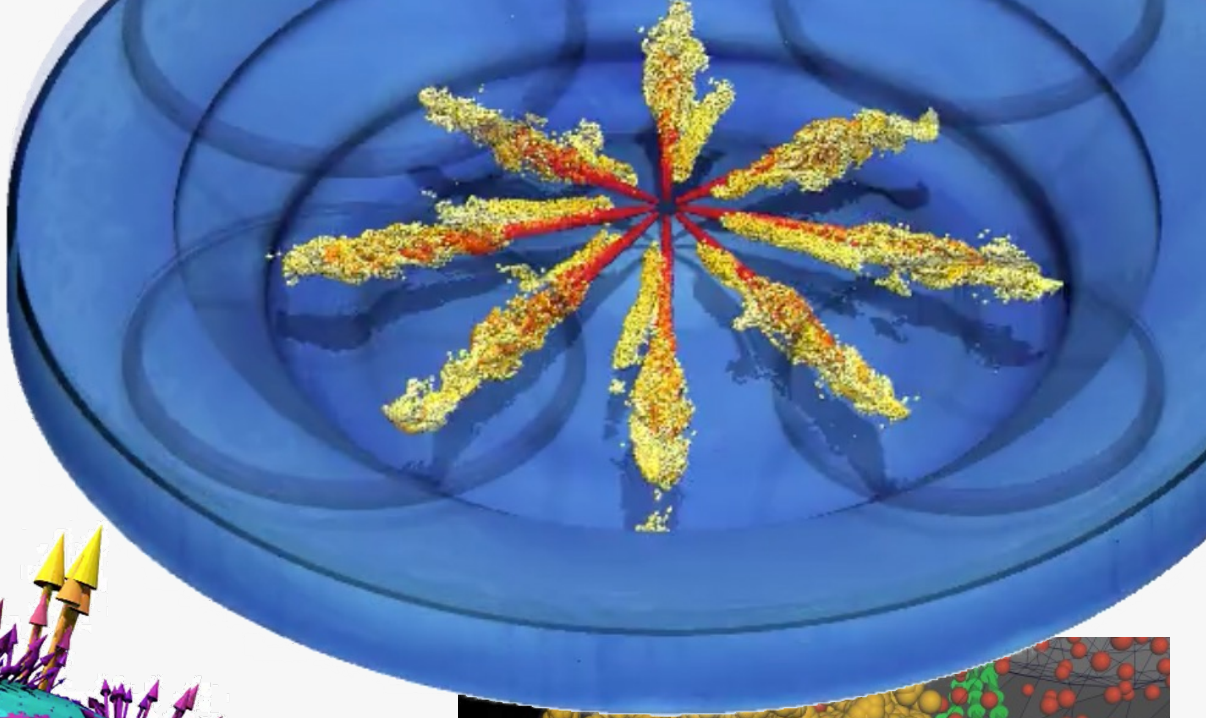
– same, but again requires more effort







Data Representations: Glyphs



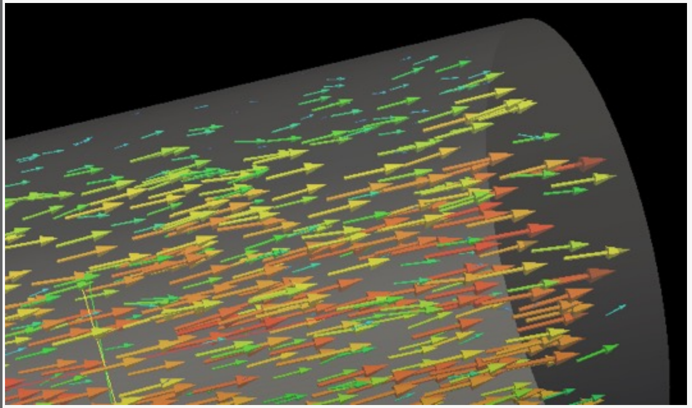
2D or 3D geometric object to represent point data

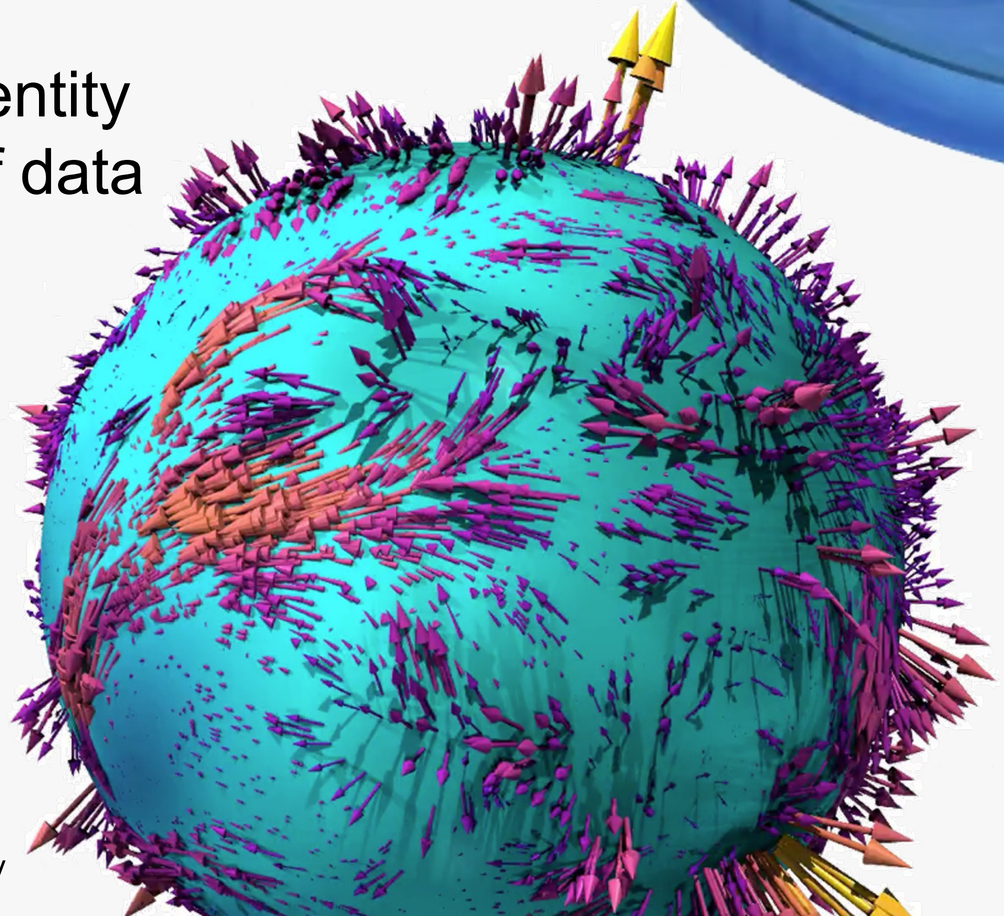
Location dictated by coordinate

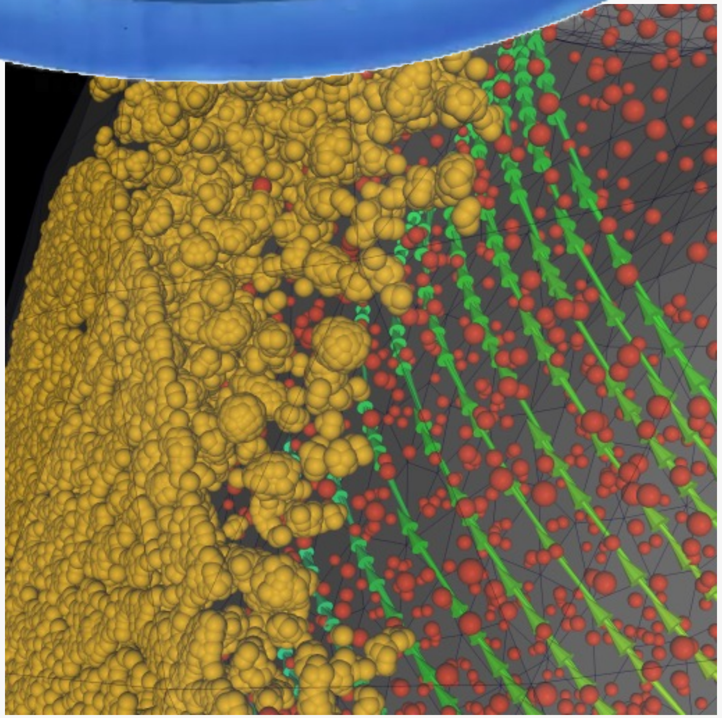
– 3D location on mesh
– 2D position in table/graph

Attributes of graphical entity dictated by attributes of data

– color, size, orientation





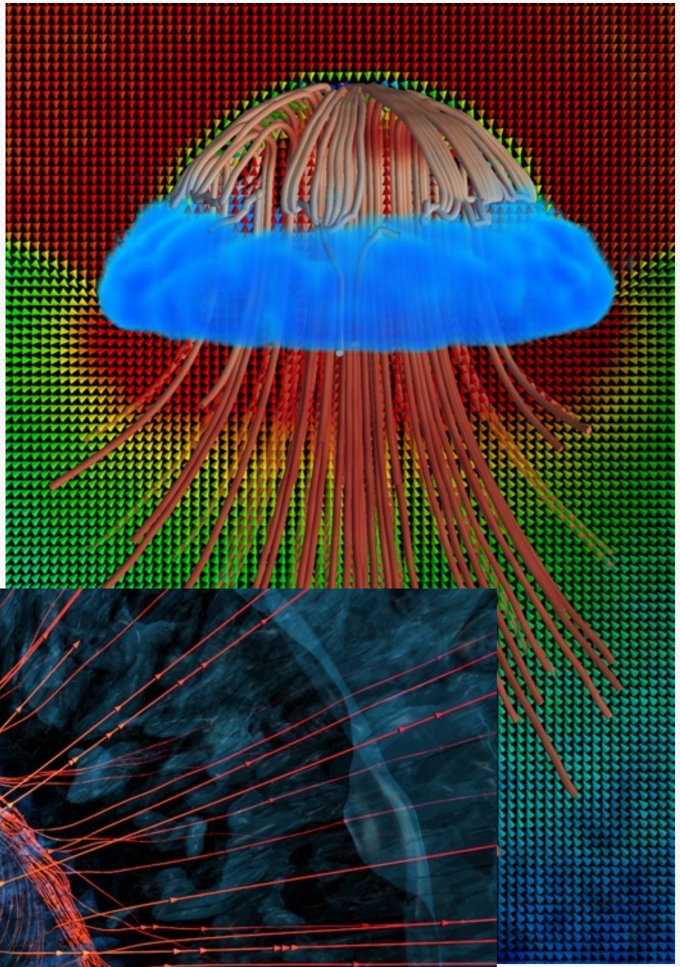


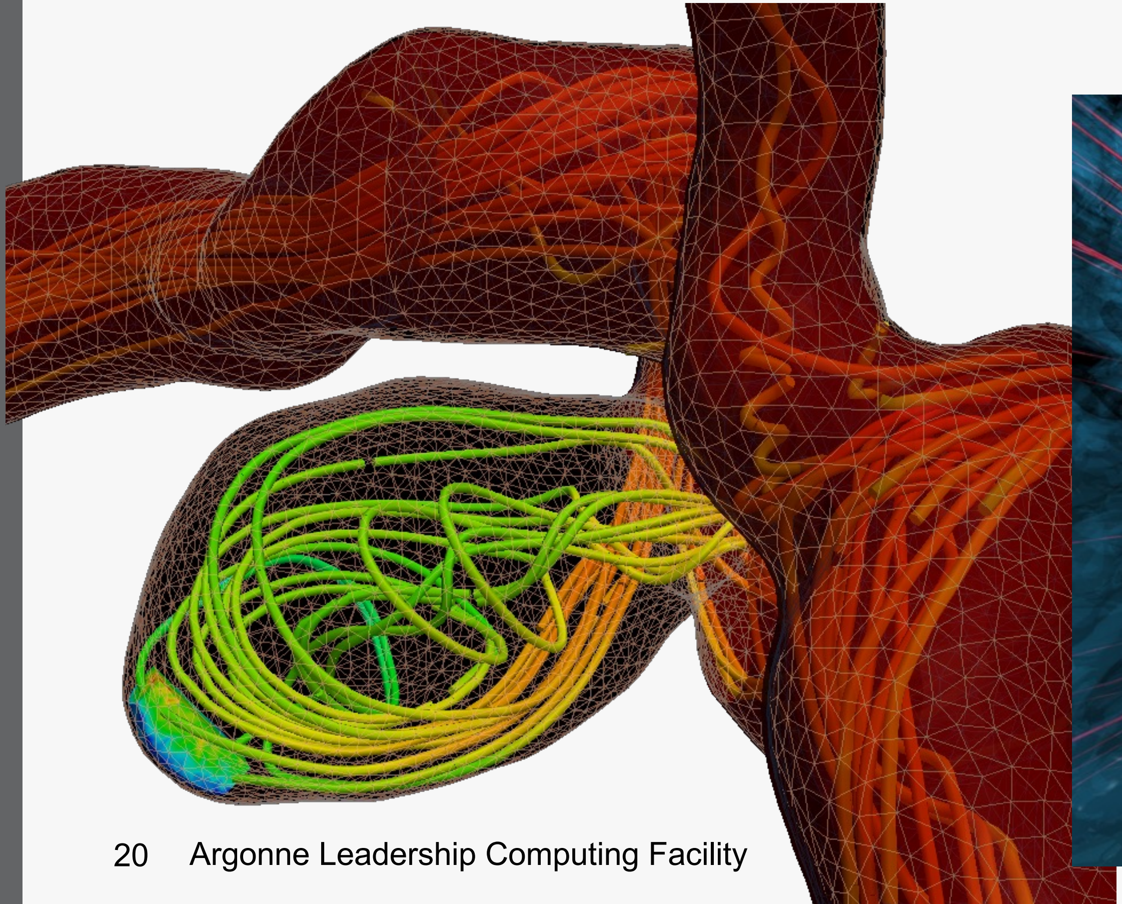
Data Representations: Streamlines

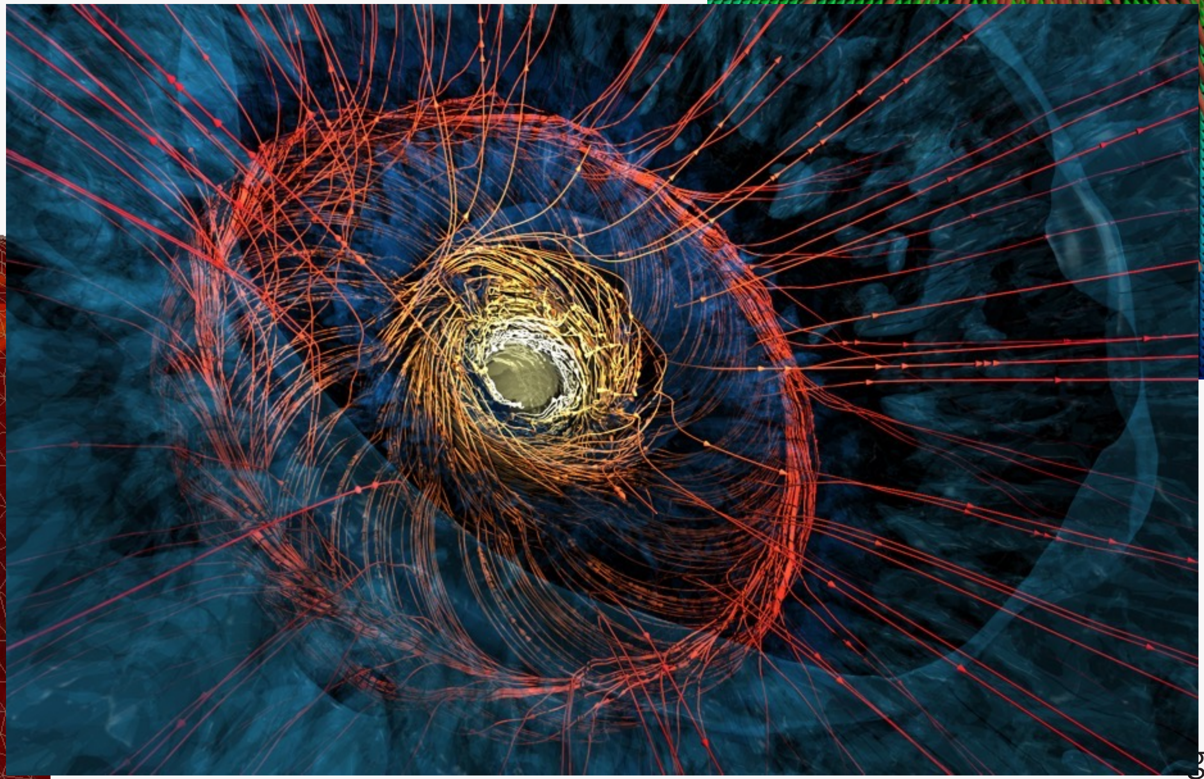
From vector field on a mesh (needs connectivity)

– Show the direction an element will travel in at any point in time.

VisIt & ParaView & vtk good at this





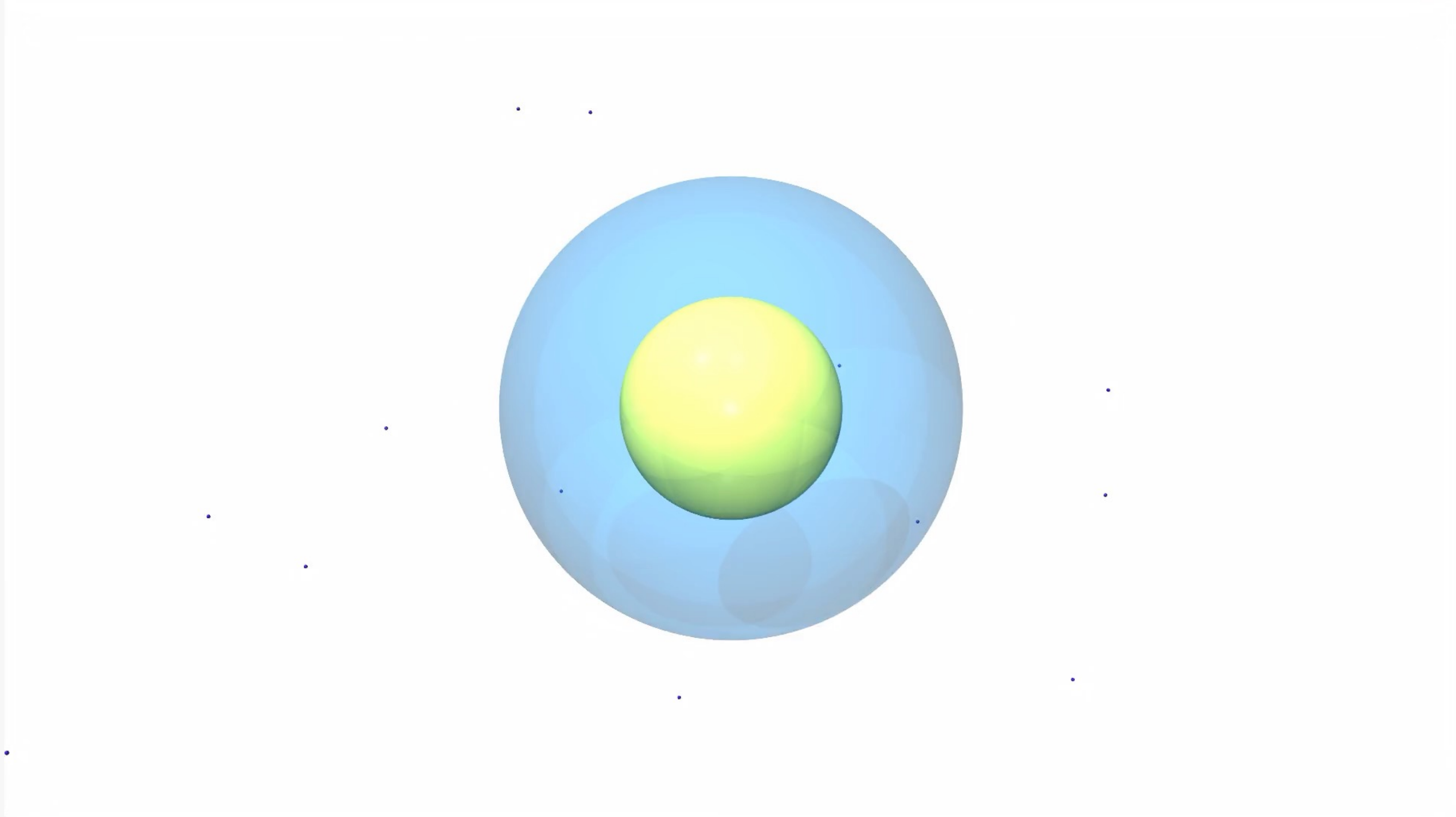


Data Representations: Pathlines

From vector field on a mesh (needs connectivity)

– Trace the path an element will travel over time.

VisIt & ParaView & vtk good at this



Data Representations: Pathlines

From vector field on a mesh (needs connectivity)

– Trace the path an element will travel over time.

VisIt & ParaView & vtk good at this

Molecular Dynamics Visualization

VMD:

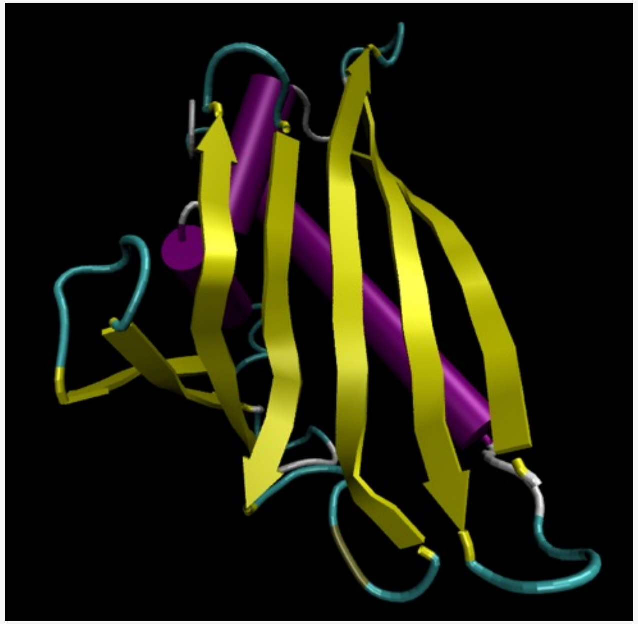
- Lots of domain-specific representations
- Many different file formats
- Animation
- Scriptable

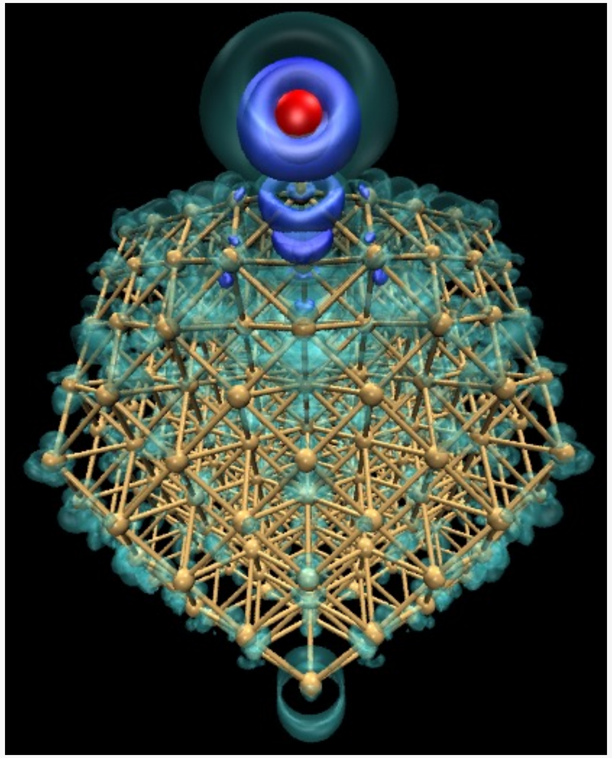
VisIt & ParaView:

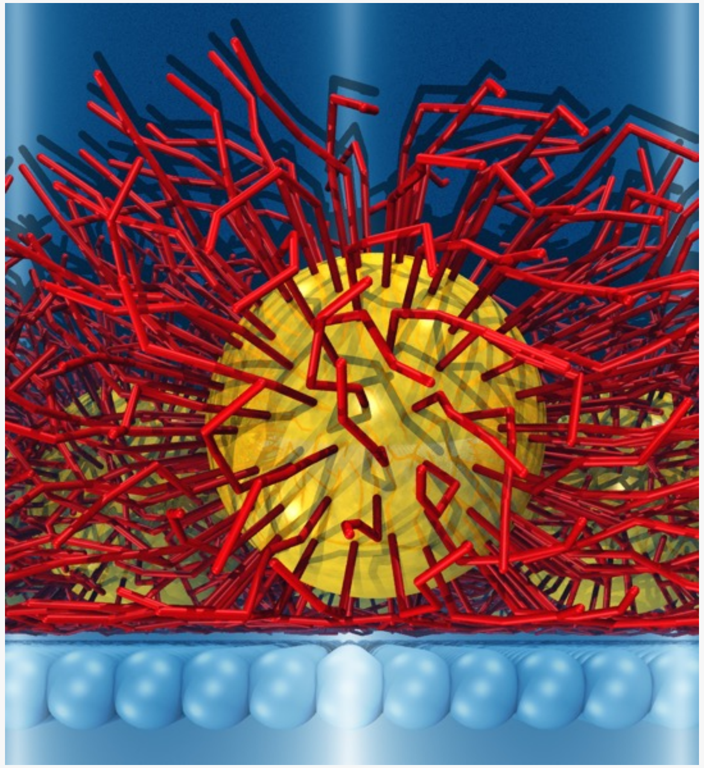
- Limited support for these types of representations, but improving

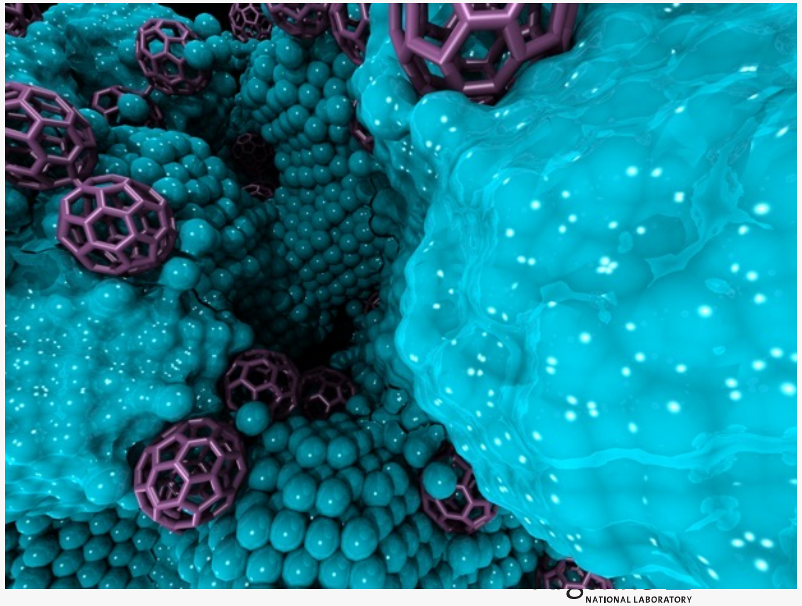
VTK:

- Anything's possible if you try hard enough



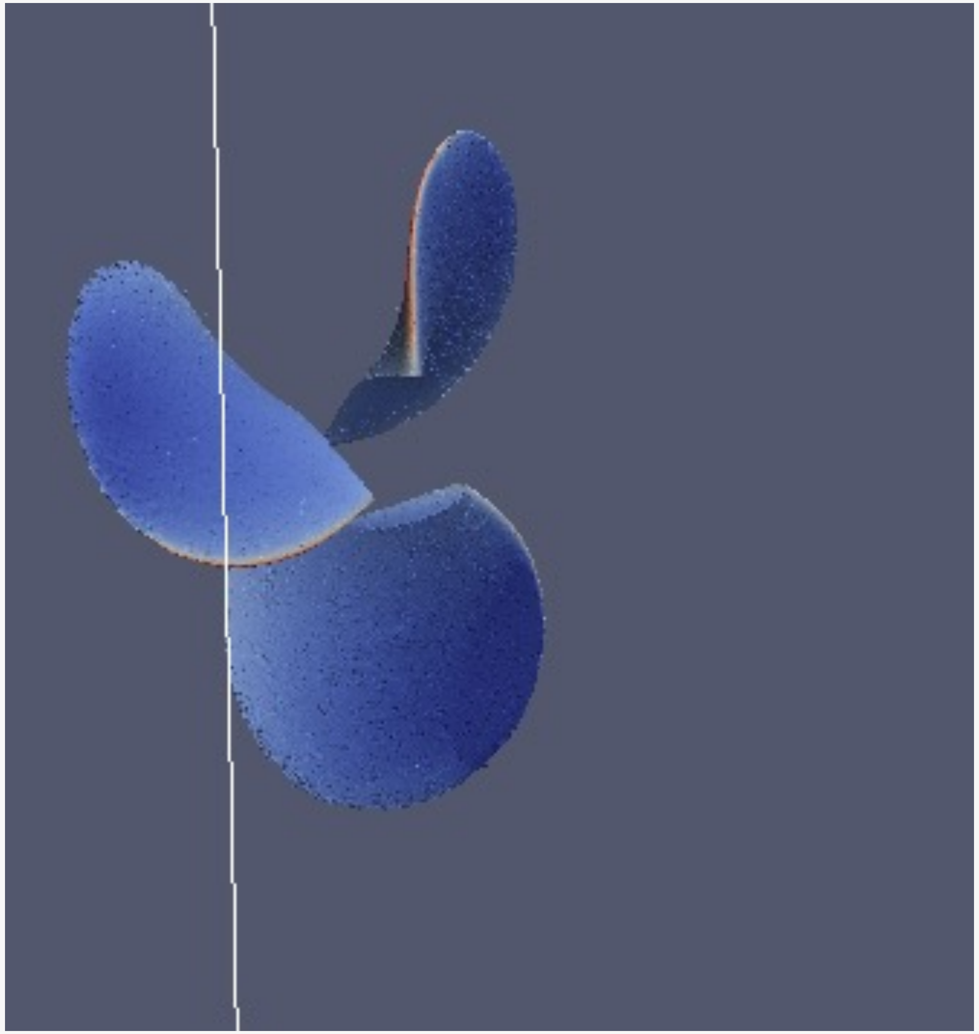




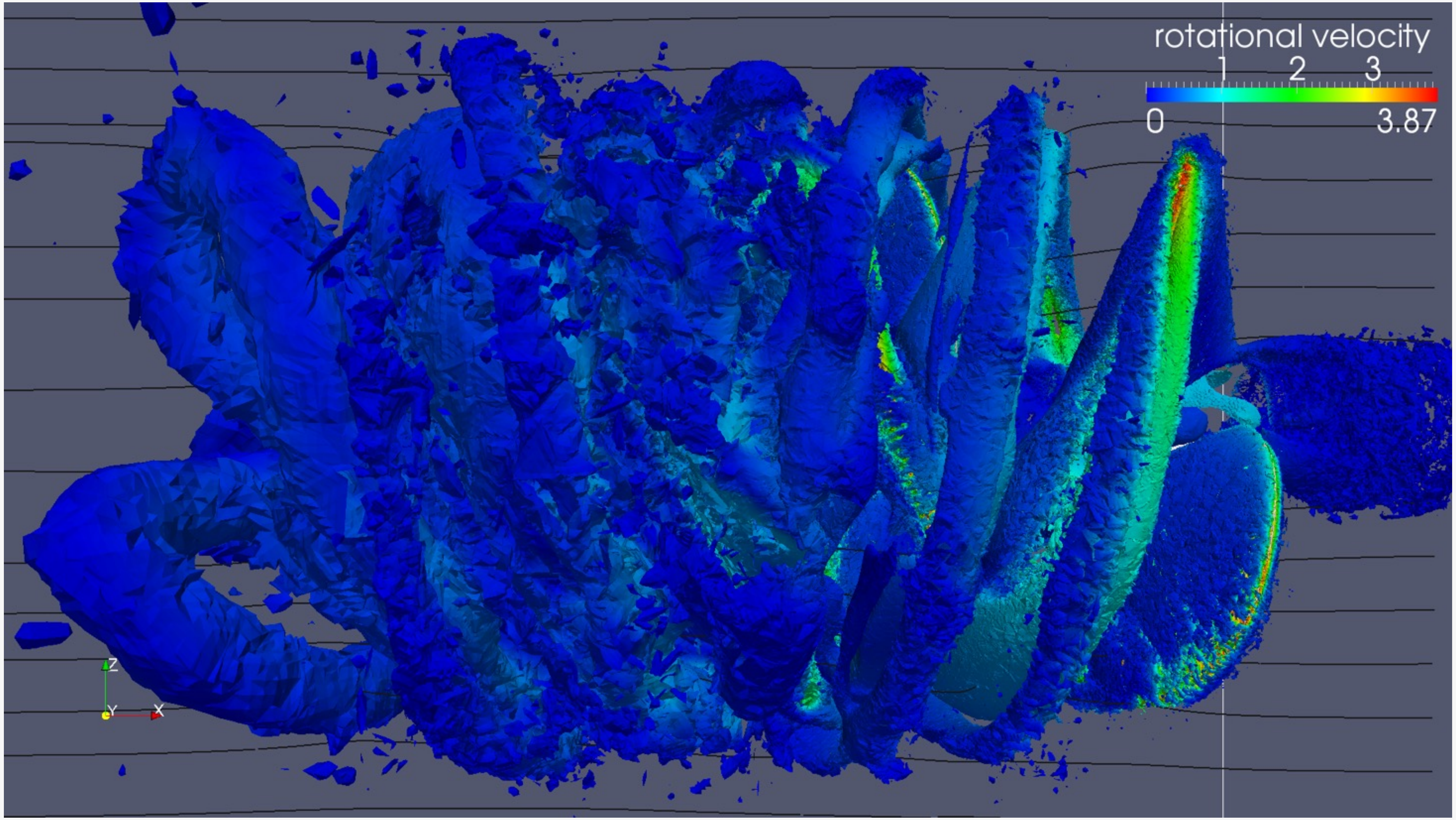


Visualization for Debugging

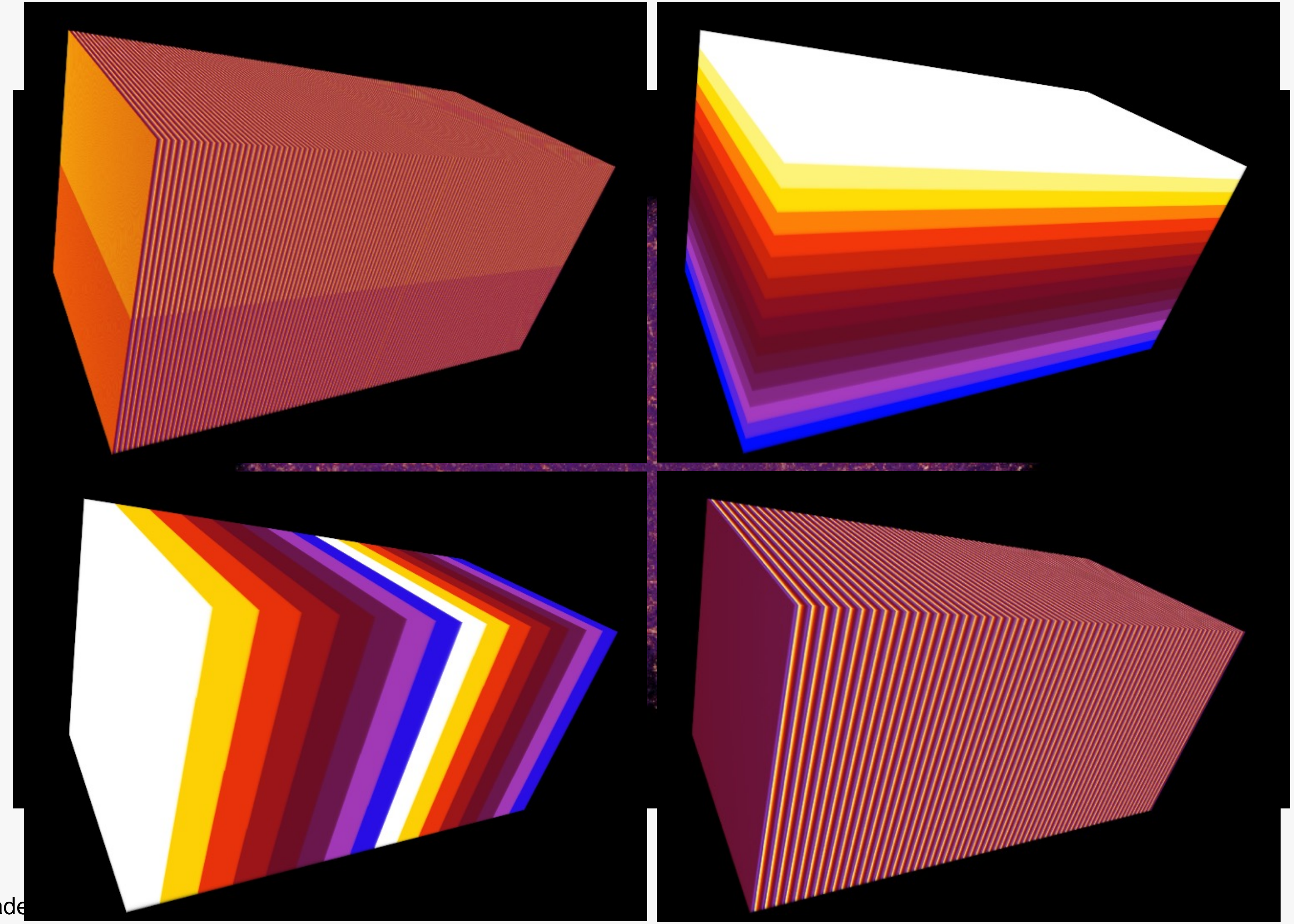
Visualization for Debugging



Visualization for Debugging



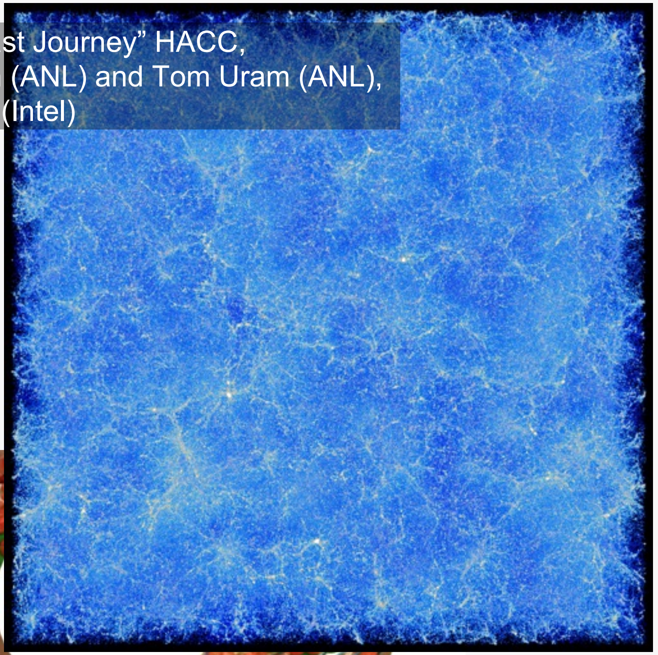
Visualization as Diagnostics: Color by Thread ID



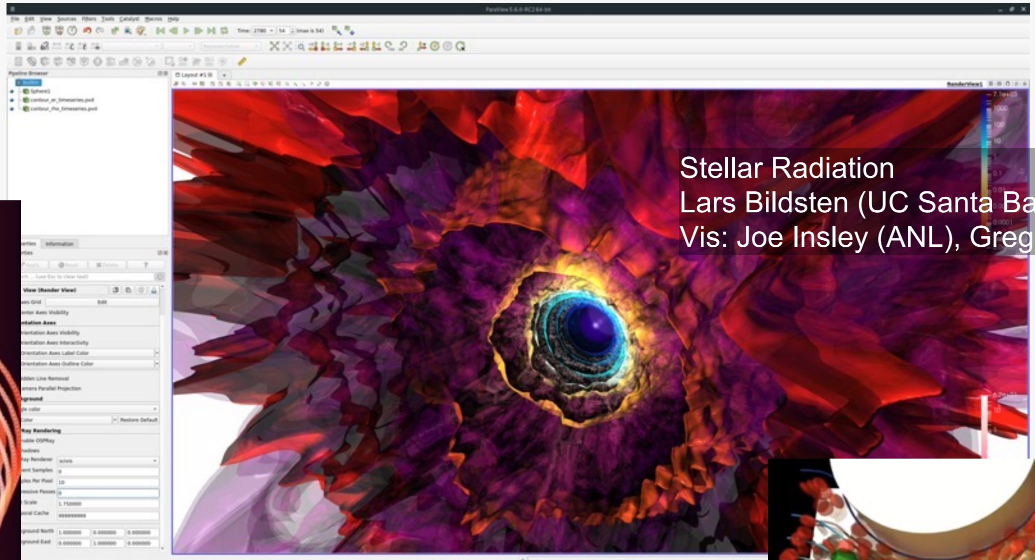
Argonne NATIONAL LABORATORY

Advanced Rendering

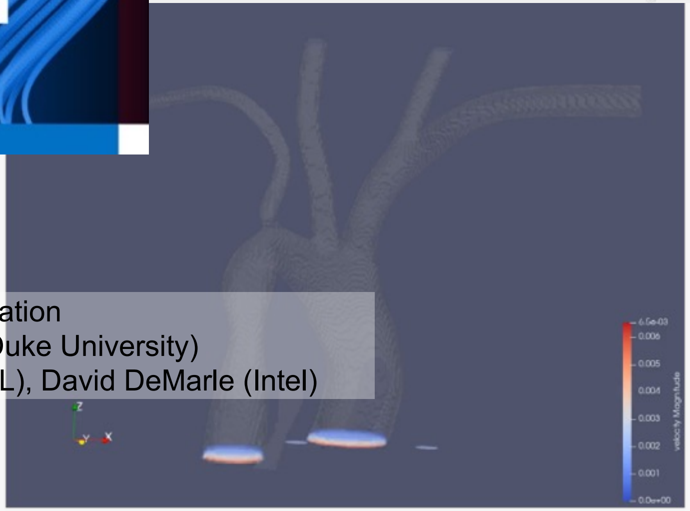
Intel® oneAPI Rendering Toolkit (“Render Kit”/"Render Framework") Open Source Software for Advanced Rendering and Visualization



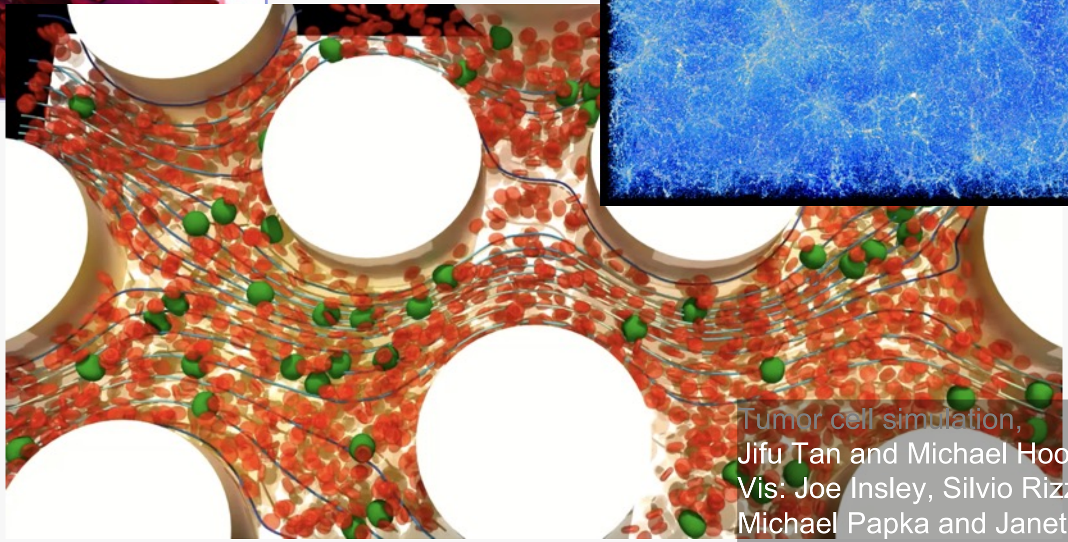
51 M particle “Last Journey” HACC,
Katrin Heitmann (ANL) and Tom Uram (ANL),
Vis: Aaron Knoll (Intel)



Stellar Radiation
Lars Bildsten (UC Santa Barbara),
Vis: Joe Insley (ANL), Greg Johnson (Intel)



Harvey Proxy Simulation
Amanda Randles (Duke University)
Vis: Silvio Rizzi (ANL), David DeMarle (Intel)



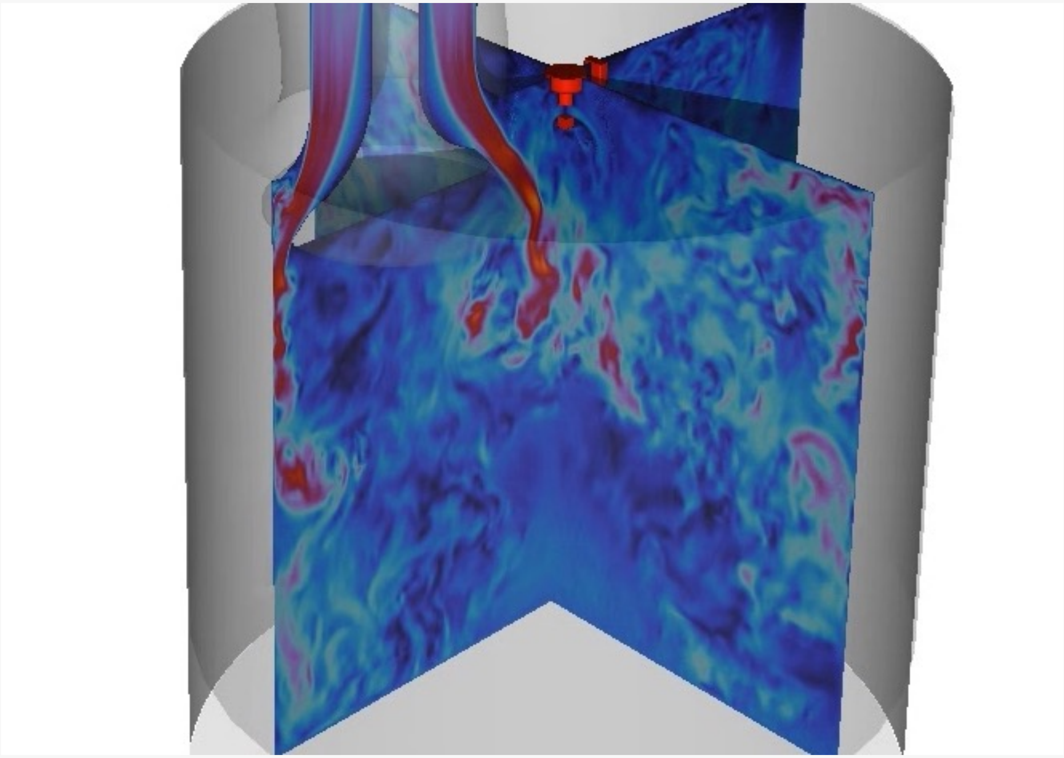
Tumor cell simulation,
Jifu Tan and Michael Hood, (NIU),
Vis: Joe Insley, Silvio Rizzi,
Michael Papka and Janet Knowles,
(ANL)

Slide courtesy of Intel Visualization Team

The Science

Internal Combustion Engine Simulation

TCC Engine Apparatus



Fluid Dynamics Simulation

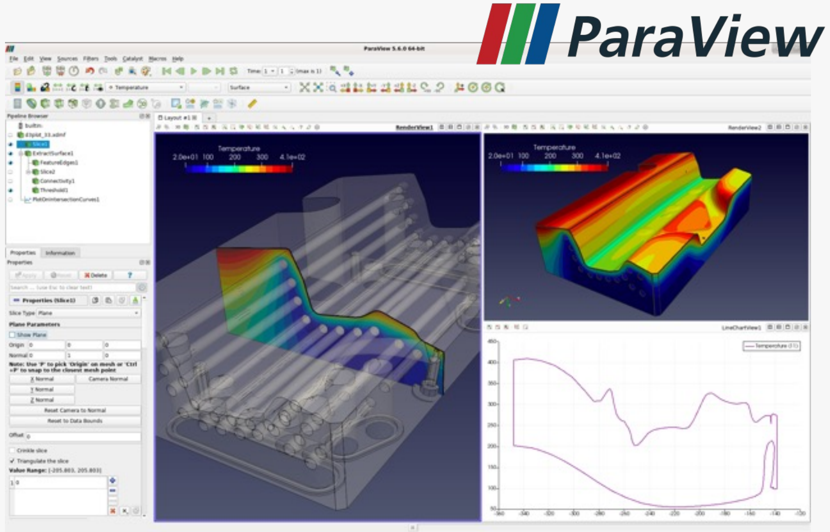
Goal

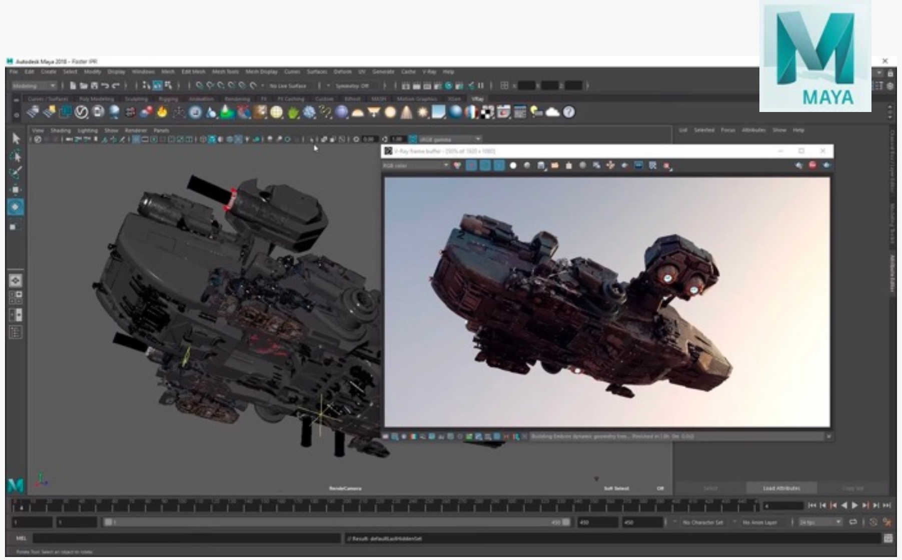
Provide context to tell the story/explain the science

Integrate production tools into the existing visualization pipeline

Tools used:

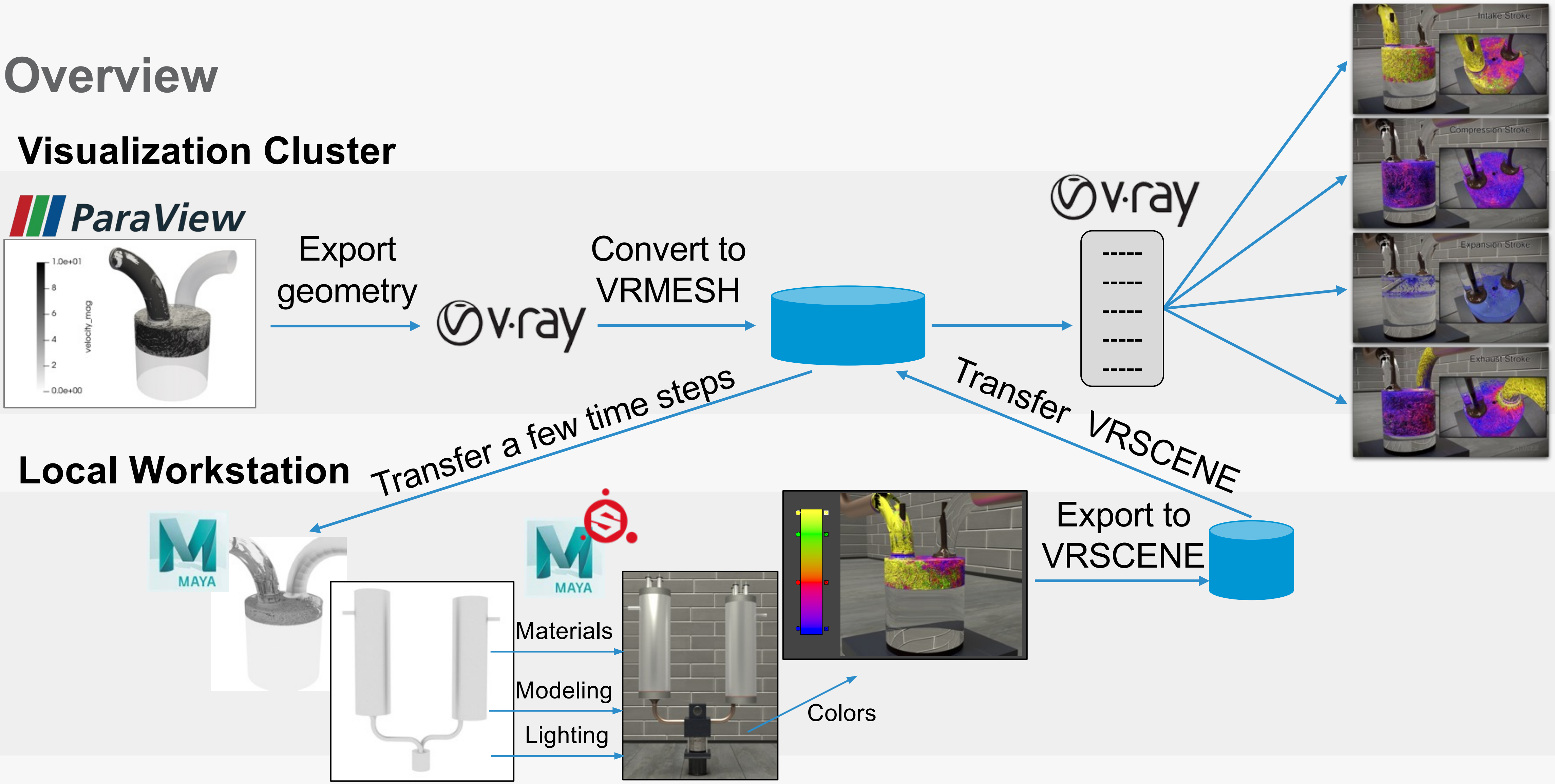
– ParaView
– Maya
– Substance Painter
– V-Ray
– Custom scripts and HPC Resources
– ffmpeg
– Premiere/After Effects





THE VISUALIZATION PIPELINE

Overview



Argonne NATIONAL LABORATORY

In Situ Visualization and Analysis

Five orders of magnitude between compute and I/O capacity on Titan Cray system at ORNL

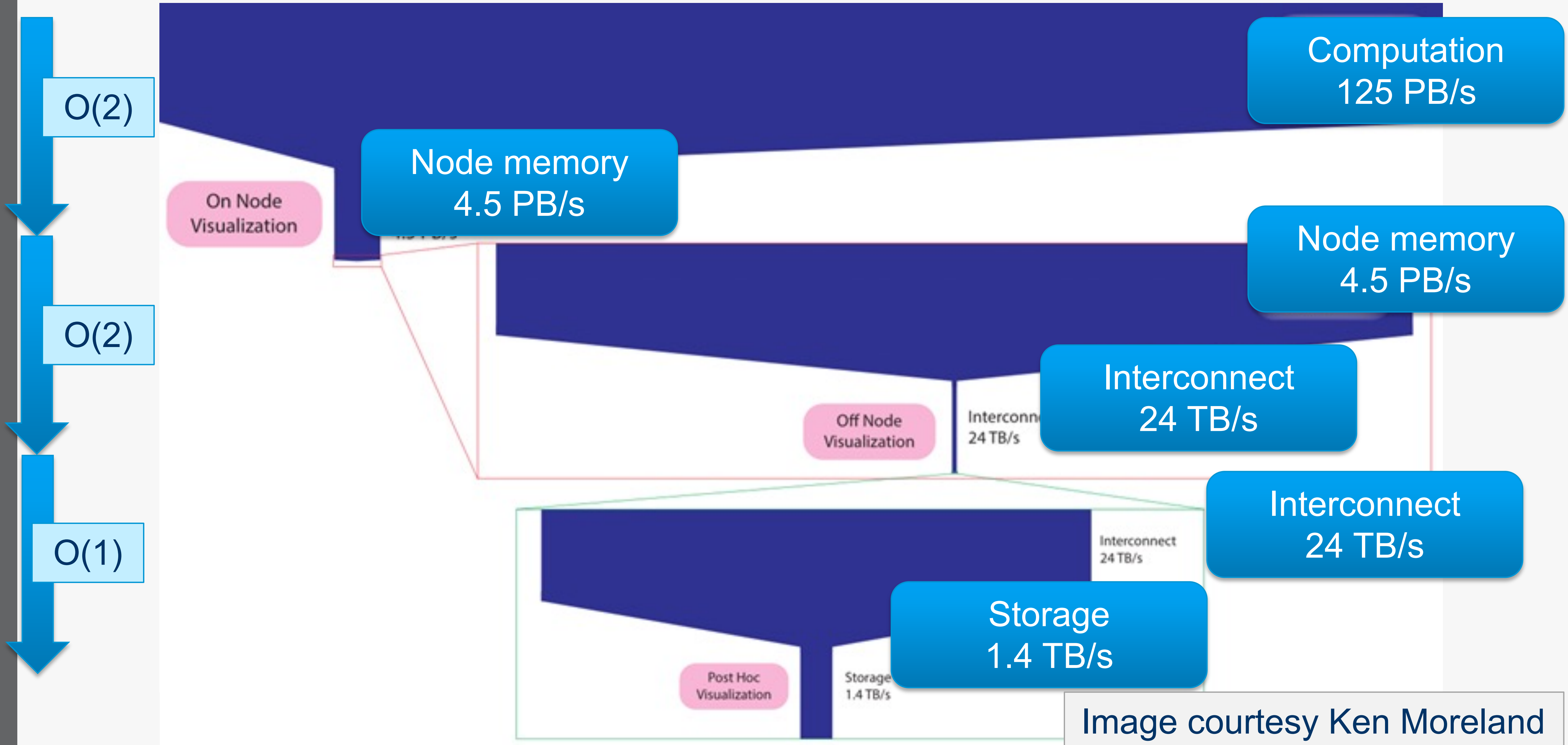


Image courtesy Ken Moreland

What are the problems?

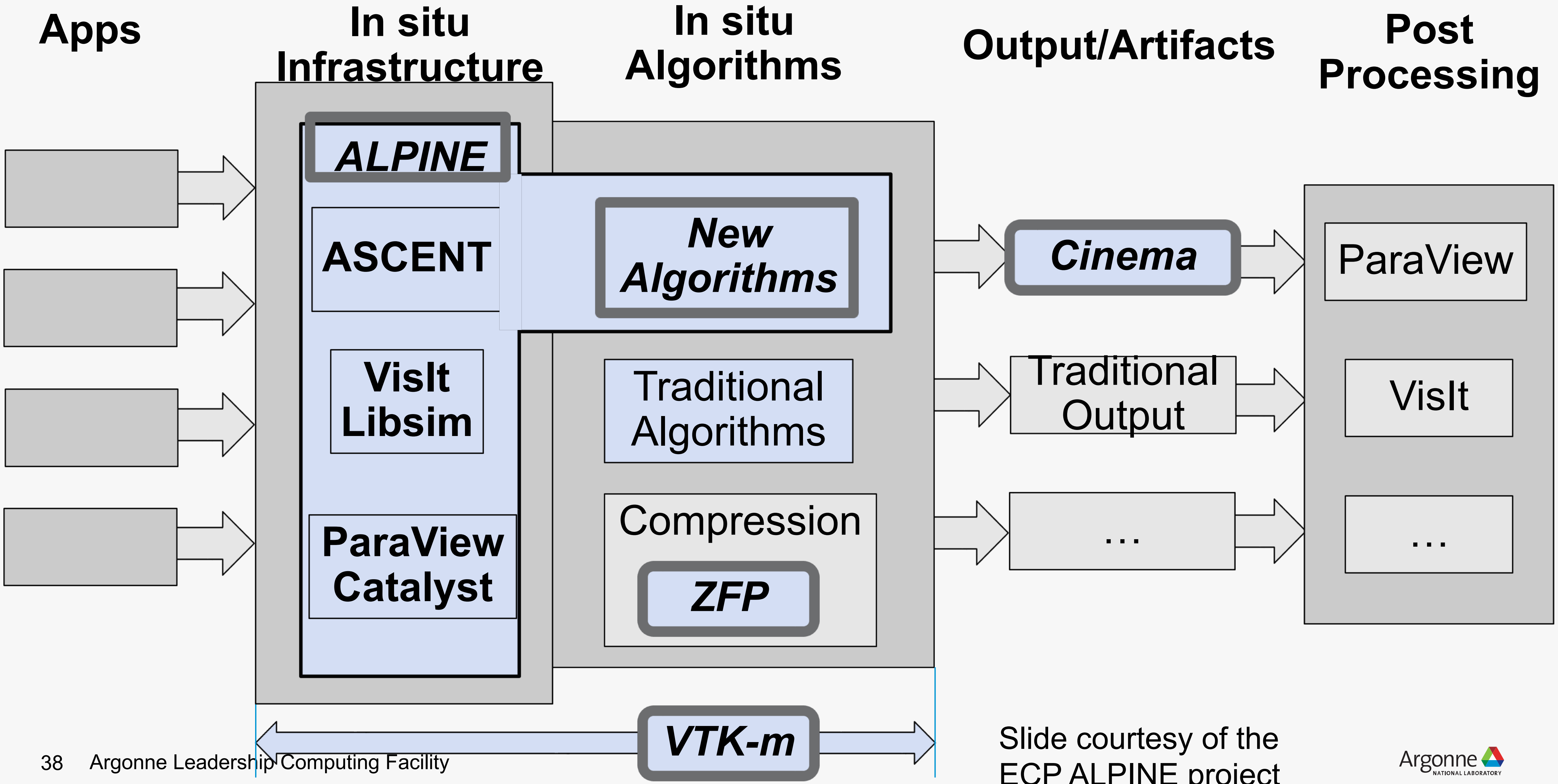
- Not enough I/O capacity on current HPC systems, and the trend is getting worse.
- If there's not enough I/O, you can't write data to storage, so you can't analyze it: lost science.
- Energy consumption: it costs a lot of power to write data to disk.
- Opportunity for doing better science (analysis) when have access to full spatiotemporal resolution data.

Slide courtesy the SENSEI team www.sensei-insitu.org

In Situ Frameworks and Infrastructures at ALCF

Name	Description	Contact person at ATPESC
ALPINE	In Situ algorithms and infrastructure for the Exascale Computing Project	Silvio Rizzi, Cyrus Harrison
ASCENT	A flyweight in situ visualization and analysis runtime for multi-physics HPC simulations	Cyrus Harrison
SENSEI	Write once run anywhere. Multiple backends. MxN in transit communication patterns	Silvio Rizzi, Joe Insley
ParaView/Catalyst	*In situ* use case library, with an adaptable application programming interface (API), that orchestrates the delicate alliance between simulation and analysis and/or visualization tasks	Dan Lipsa
Libsim	Originally developed to facilitate interactive connections from VisIt to running simulations	Cyrus Harrison
SmartSim	SmartSim is a software framework that facilitates the convergence of numerical simulations and AI workloads on heterogeneous architectures	Silvio Rizzi

Exascale Computing Project
Software Technology Data and Visualization



Slide courtesy of the ECP ALPINE project

Ascent

- Flyweight design, minimizes dependencies
- Data model based on Conduit from LLNL
- Vis and analysis algorithms implemented in VTK-m

```
//
// Run Ascent
//

Ascent ascent;
ascent.open();
ascent.publish(data);
ascent.execute(actions);
ascent.close();
```

Slide courtesy of the Ascent team

SENSEI: Write once run everywhere

- “Write once, run everywhere” design
- Data model based on VTK from Kitware
- Supports a variety of backends, including ParaView/Catalyst, VisIt/LibSim, ADIOS, Python
- MxN in transit capabilities

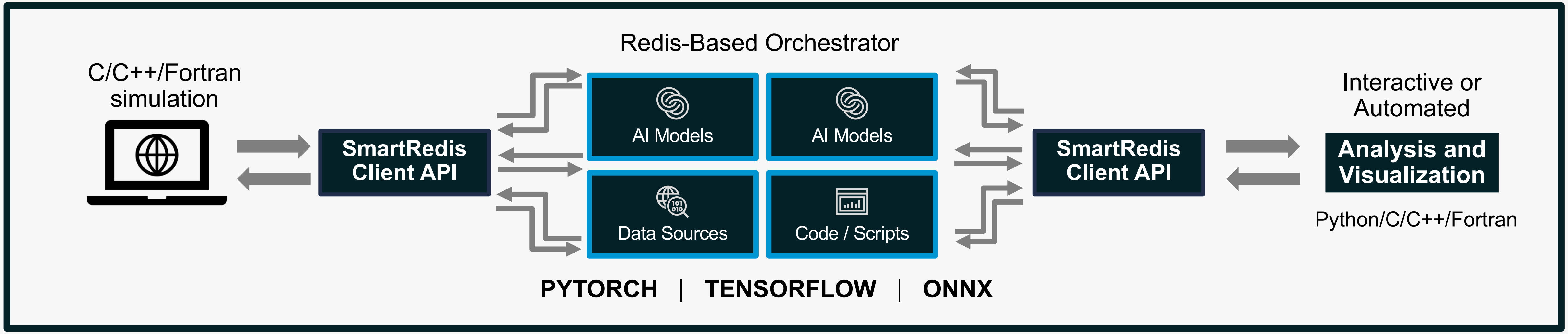
SmartSim Overview

The **SmartSim open-source library** enables scientists, engineers, and researchers to embrace a **"data-in-motion" philosophy** to accelerate the convergence of **AI/data science** techniques and **HPC simulations**

SmartSim enables **simulations** to be used as **engines** within a system, **producing data,** consumed by other services enable **new applications**

- Embed **machine learning** training and inference with **existing** in Fortran/C/C++ **simulations**
- **Communicate** data **between** C, C++, Fortran, and Python **applications**
- Analyze and visualize **data streamed** from **HPC applications** while they are **running**
- **Launch**, **configure**, and **coordinate** complex simulation, analysis, and visualization **workflows**

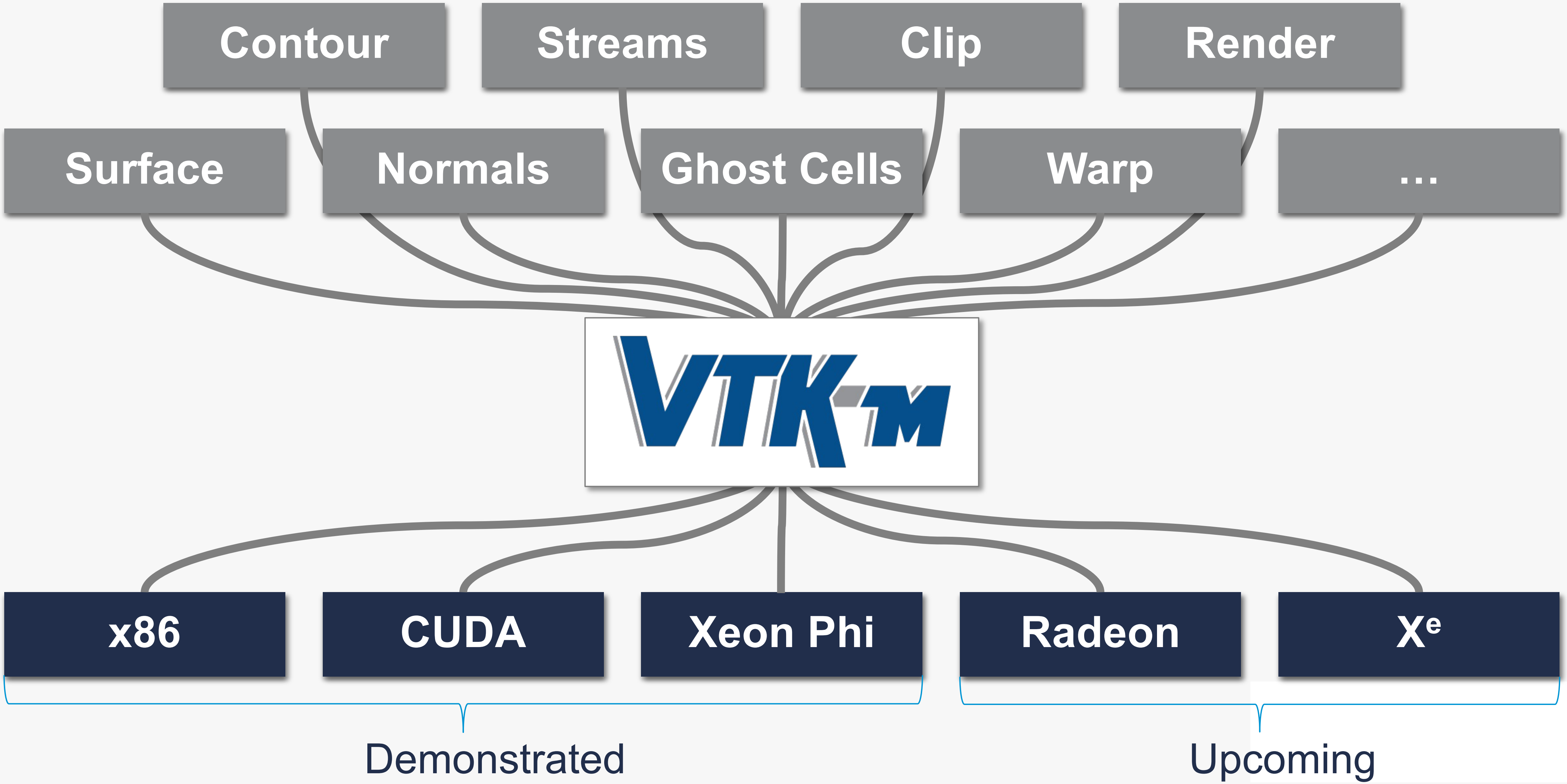
All of these can be done without touching the filesystem, i.e. **data-in-motion**



slide courtesy of the HPE SmartSim team

Infrastructures

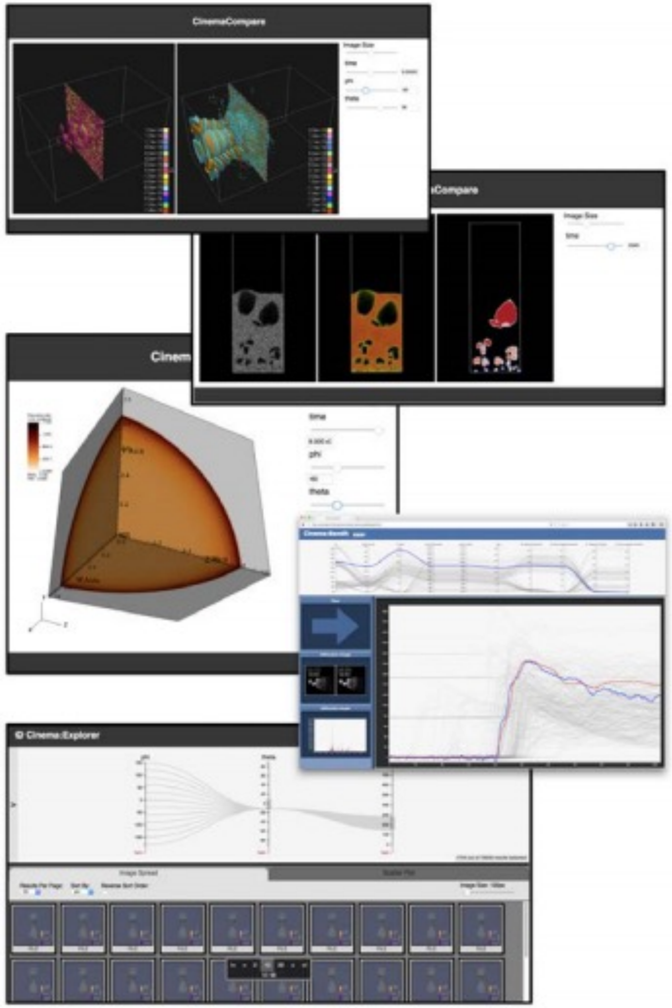
VTK-m's main thrust: a write-once-run-everywhere framework



Slide courtesy of the ECP VTK-m project

What is Cinema?

- **Cinema** is part of an integrated workflow, providing a method of extracting, saving, analyzing or modifying and viewing complex data artifacts from large scale simulations.
  - If you're having difficulty exploring the complex results from your simulation, Cinema can help.
- **The Cinema 'Ecosystem'** is an integrated set of writers, viewers, and algorithms that allow scientists to export, analyze/modify and view Cinema databases.
  - This ecosystem is embodied in widely used tools (**ParaView, VisIt, Ascent**) and the database specification.



Slide courtesy of Los Alamos National Lab and the CINEMA project

In Situ Computational Fluid Dynamics

Team: Paul Fisher et al.

NekRS

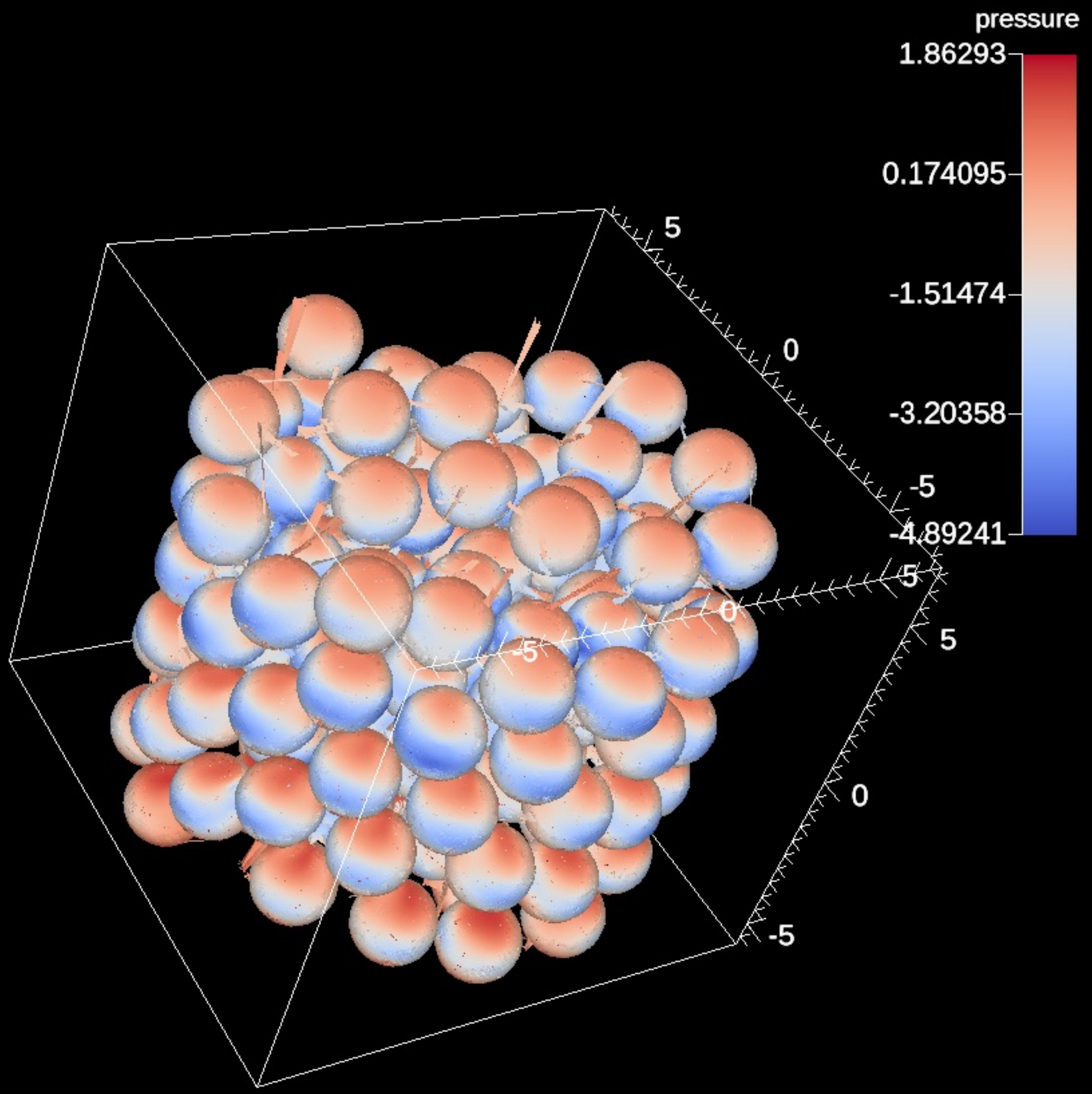
COPYRIGHT (c) 2019-2023 UCHICAGO ARGONNE, LLC

- CFD code
- Simulates turbulent incompressible or low Mach-number flows with heat transfer and species transport.
- Supports heterogeneous platforms
- Legacy code: Nek5000

As the resolution of the simulation increases, scientists turn to performing analysis in situ, doing their analysis while data is still resident in memory. Such capabilities enable study of turbulence statistics at these extremely high resolutions.

NekRS + Ascent

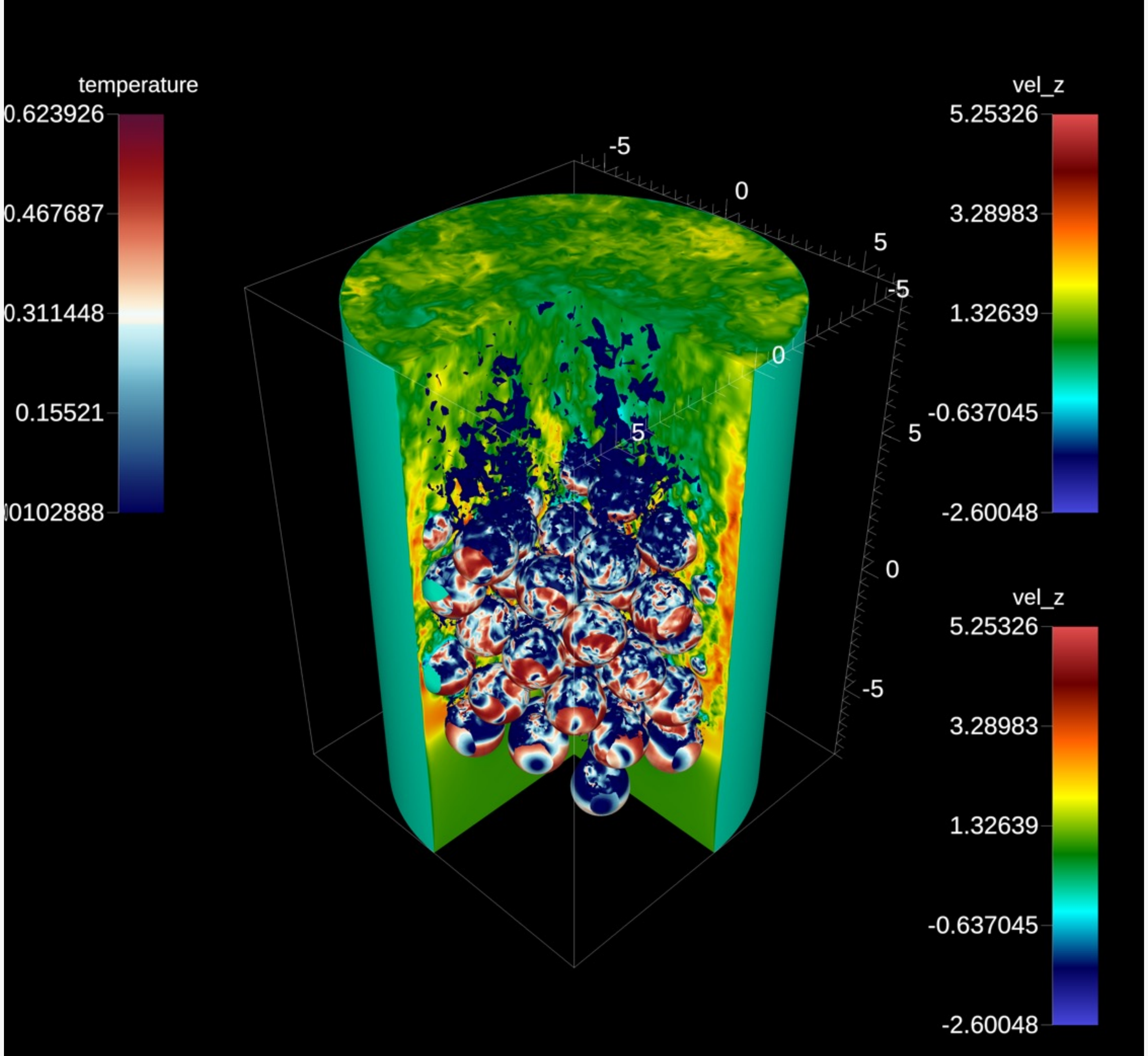
- Data is passed by reference (zero-copy)
  - CPU -> CPU or GPU -> GPU
- Ascent is disabled by default
  - Binary is bit by bit identical with non-instrumented code
- Full functionality of Ascent at your disposal
- Closely working with NekRS team, who are testing the instrumentation and providing feedback



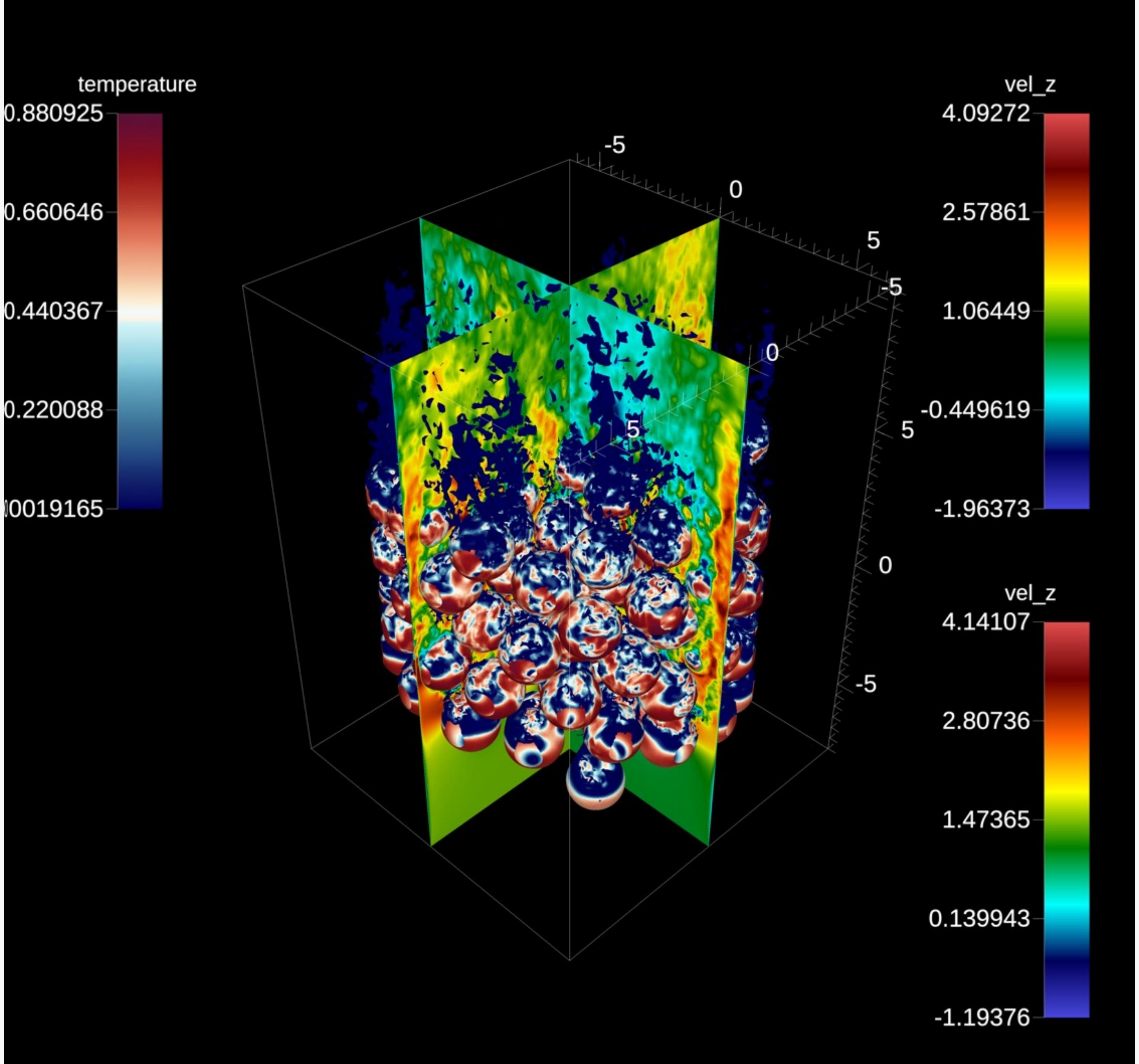

```
 ------------------------------------------------------------------------------+
70.103.01    Driver Version: 470.103.01    CUDA Version: 11.4                  |
---------------------------+-----------------------+---------------------------+
       Persistence-M| Bus-Id        Disp.A | Volatile Uncorr. ECC |
erf  Pwr:Usage/Cap|         Memory-Usage | GPU-Util  Compute M. |
                   |                      |               MIG M. |
===================+======================+======================|
A100-SXM...  On   | 00000000:07:00.0 Off |                    0 |
 P0   318W / 400W |  13934MiB / 40536MiB |     93%      Default |
                   |                      |             Disabled |
-------------------+----------------------+----------------------+
A100-SXM...  On   | 00000000:46:00.0 Off |                    0 |
 P0   289W / 400W |  13928MiB / 40536MiB |     95%      Default |
                   |                      |             Disabled |
-------------------+----------------------+----------------------+
A100-SXM...  On   | 00000000:85:00.0 Off |                    0 |
 P0   319W / 400W |  13952MiB / 40536MiB |     90%      Default |
                   |                      |             Disabled |
-------------------+----------------------+----------------------+
A100-SXM...  On   | 00000000:C7:00.0 Off |                    0 |
 P0   343W / 400W |  13952MiB / 40536MiB |     90%      Default |
                   |                      |             Disabled |
-------------------+----------------------+----------------------+

-----------------------------------------------------------------+
                                                                 |
CI        PID   Type   Process name                  GPU Memory |
ID                                                   Usage      |
=================================================================|
N/A     12689      C   ...-ascent-polaris/bin/nekrs    13931MiB |
N/A     12690      C   ...-ascent-polaris/bin/nekrs    13925MiB |
N/A     12691      C   ...-ascent-polaris/bin/nekrs    13949MiB |
N/A     12692      C   ...-ascent-polaris/bin/nekrs    13949MiB |
-----------------------------------------------------------------+
```

Polaris - 40 ranks



Yu-Hsiang Lan, Misun Min, Paul Fisher



Yu-Hsiang Lan, Misun Min, Paul Fisher

Next steps…

Bi-directional steering

Visualization Help

support@alcf.anl.gov

Publication Images & Covers

Animations

- SC Visualization Showcase [Best Vis Finalist 2014-2020]
- APS Division of Fluid Dynamics Gallery of Fluid Motion
- SC Gordon Bell Submissions
- Press Releases

In Situ Vis and Analysis

Argonne NATIONAL LABORATORY

Additional information

ALPINE: https://alpine.dsscale.org/
Ascent: https://github.com/Alpine-DAV/ascent
SENSEI: https://sensei-insitu.org/
SmartSim: https://developer.hpe.com/platform/smartsim/home/
ParaView/Catalyst: https://www.paraview.org/in-situ/
Libsim: https://www.visitusers.org/index.php?title=VisIt-tutorial-in-situ
VTK-m: https://m.vtk.org/
Cinema: https://cinemascience.github.io/
OSPRay: https://github.com/ospray/ospray

QUESTIONS?

Joe Insley
insley@anl.gov

Silvio Rizzi
srizzi@anl.gov

Janet Knowles
jknowles@anl.gov

Victor Mateevitsi
vmateevitsi@anl.gov