Visualizing your Data

Joseph Insley
Lead, Visualization & Data Analysis
Argonne Leadership Computing Facility

Silvio Rizzi
Computer Scientist
Argonne Leadership Computing Facility

Janet Knowles
Principal Software Engineering Specialist
Argonne Leadership Computing Facility

Victor Mateevitsi
Computer Scientist
Argonne Leadership Computing Facility

ALCF Visualization Group

Silvio Rizzi

Joe Insley

Janet Knowles

Victor Mateevitsi

Here’s the plan…

— **Visualization resources, tools and formats**

— **Visualization Success Stories**

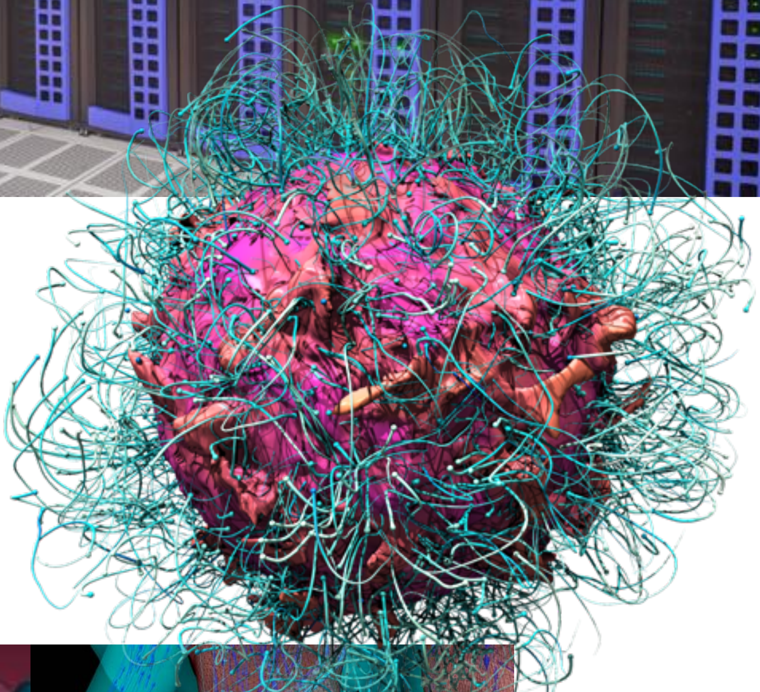
- Blood Flow
- Cosmology
- Combustion

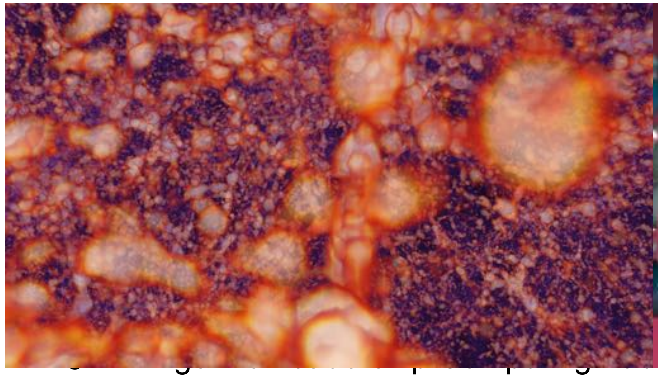
— **In Situ Visualization and Analysis**

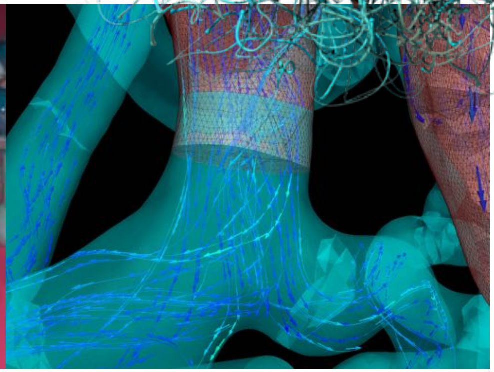
Visualization Resources, Tools and Data Formats

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
- Aggregate GPU memory of ~3TB
- Cray CS System
- 216 port FDR IB switch with uplinks to our QDR infrastructure
- Mounts the Theta, Eagle, and Grand file systems







All Sorts of Tools

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

■ Available on Cooley

* Available on Theta

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)

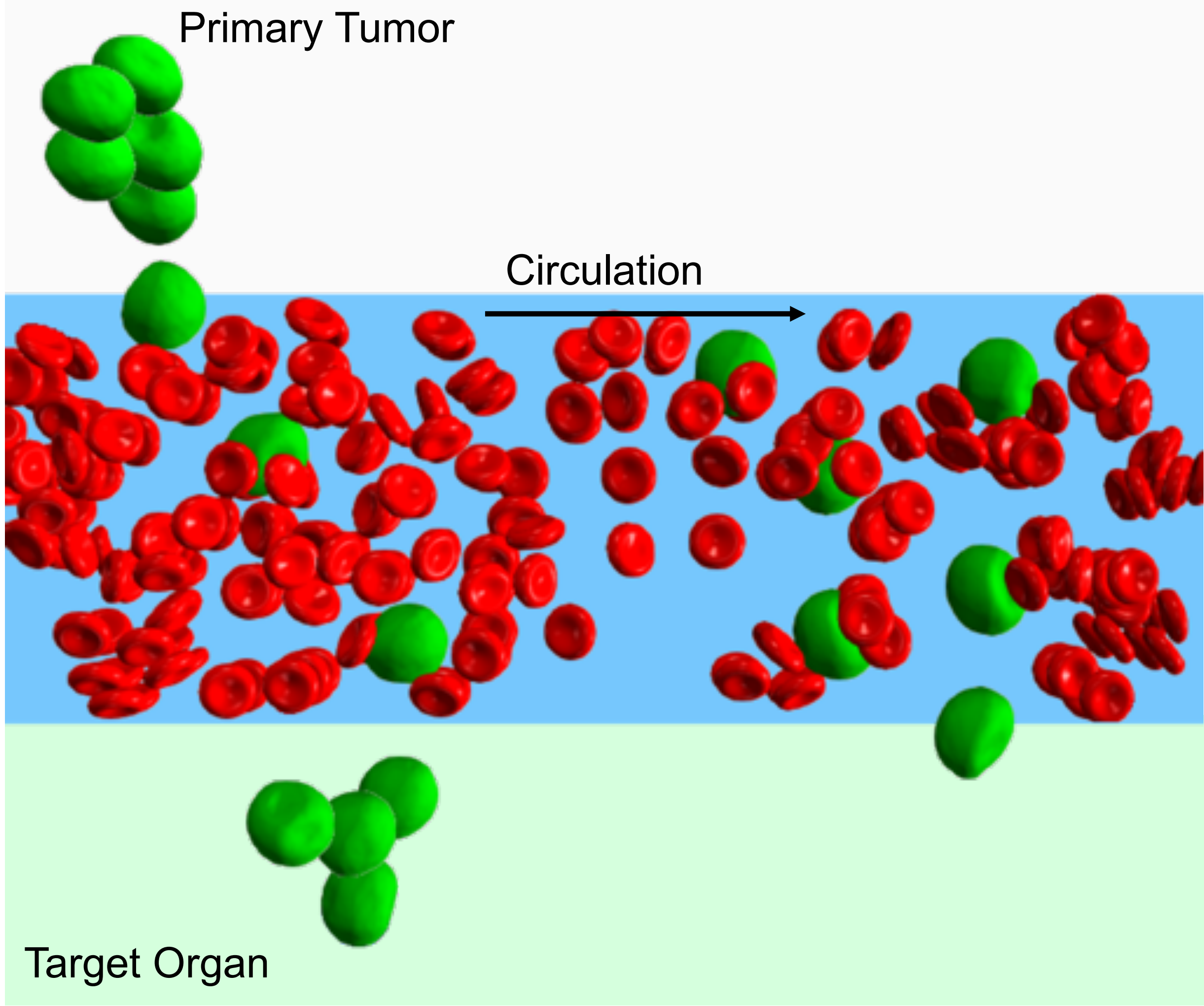
Data File Formats (ParaView & VisIt)

- 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

Success Stories: Blood Flow

Motivation

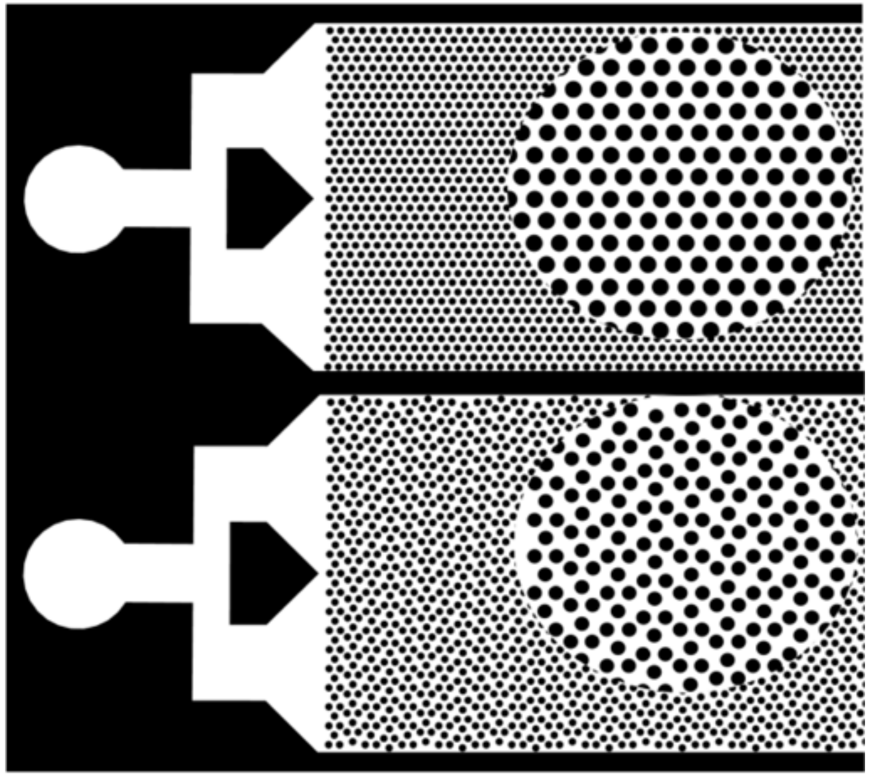
- Circulating tumor cells (CTCs)
  - Can lead to metastasis
  - Metastasis of cancer accounts for 90% of cancer deaths
- Isolation of CTCs
  - Enabling earlier diagnosis
  - Earlier treatment and better outcomes



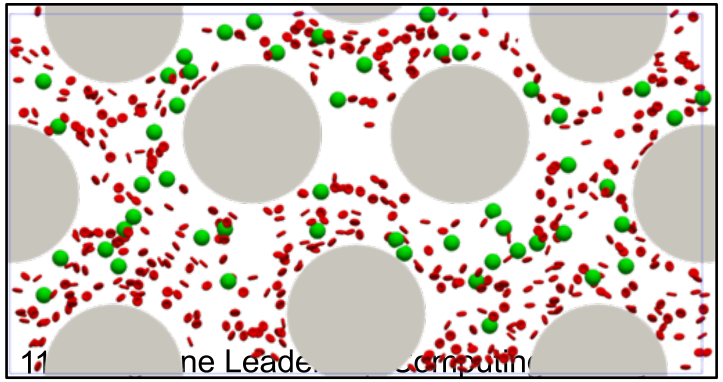
CTCs in the Device

- CTC physical properties
  - Receptor density on cell surfaces
  - Deformability
  - Size
- Device design and operation parameters
  - Pressure gradient
  - Micropost size and spacing
  - Ligand coating density on microposts
  - Red blood cell (RBC) volume fraction

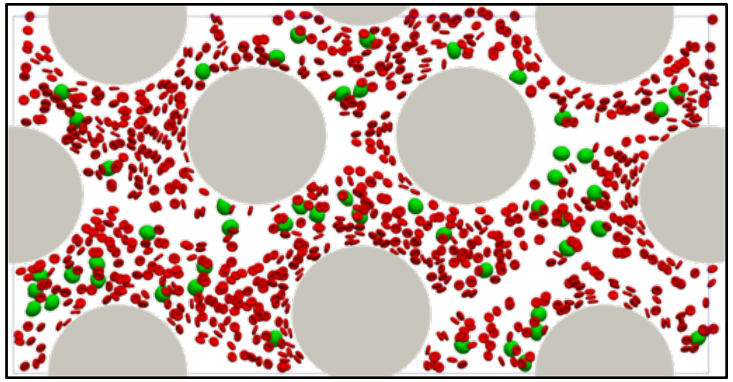
Regular and shifted position of microposts



5% RBC volume fraction

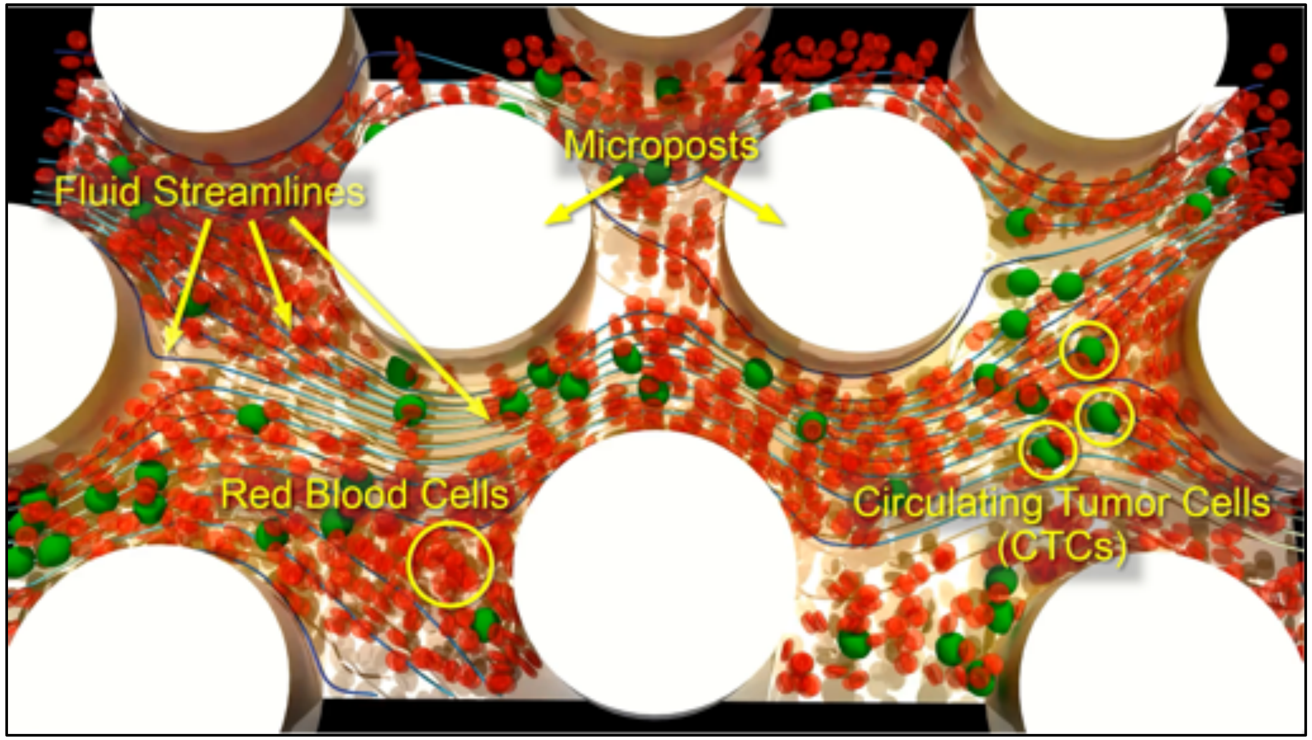


10% RBC volume fraction



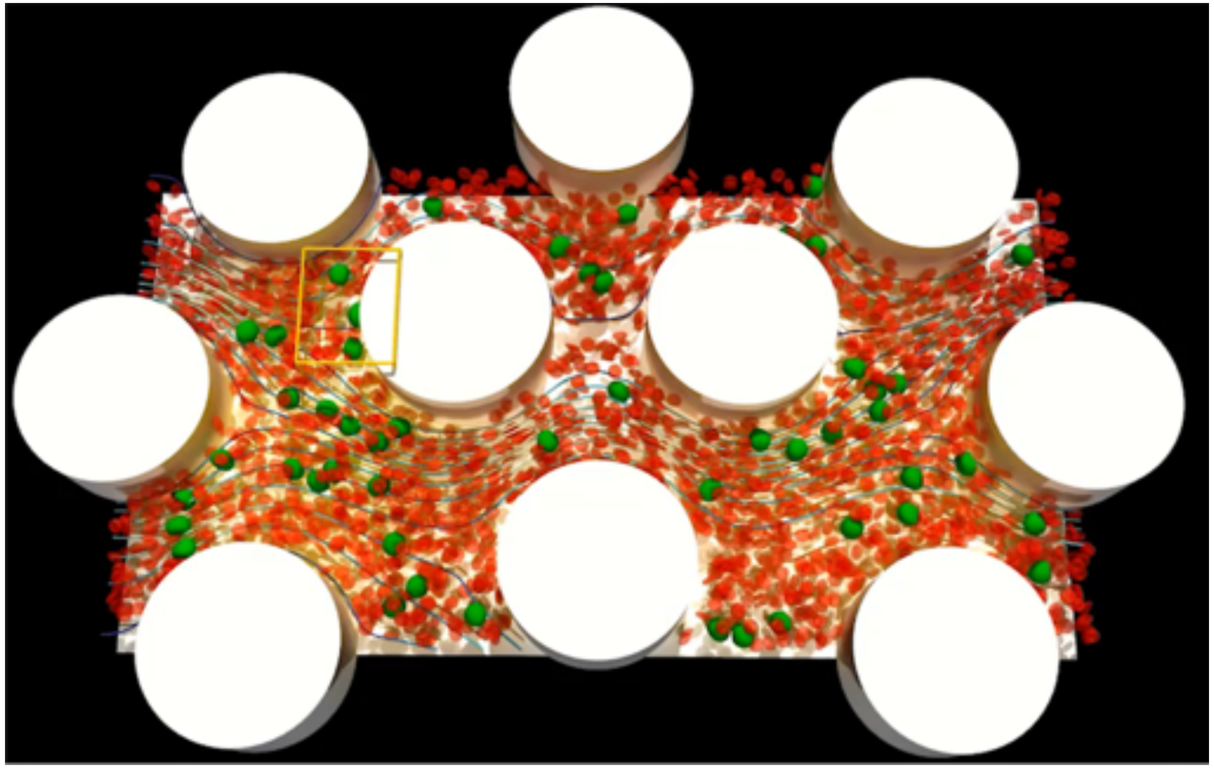
Modeling Overview

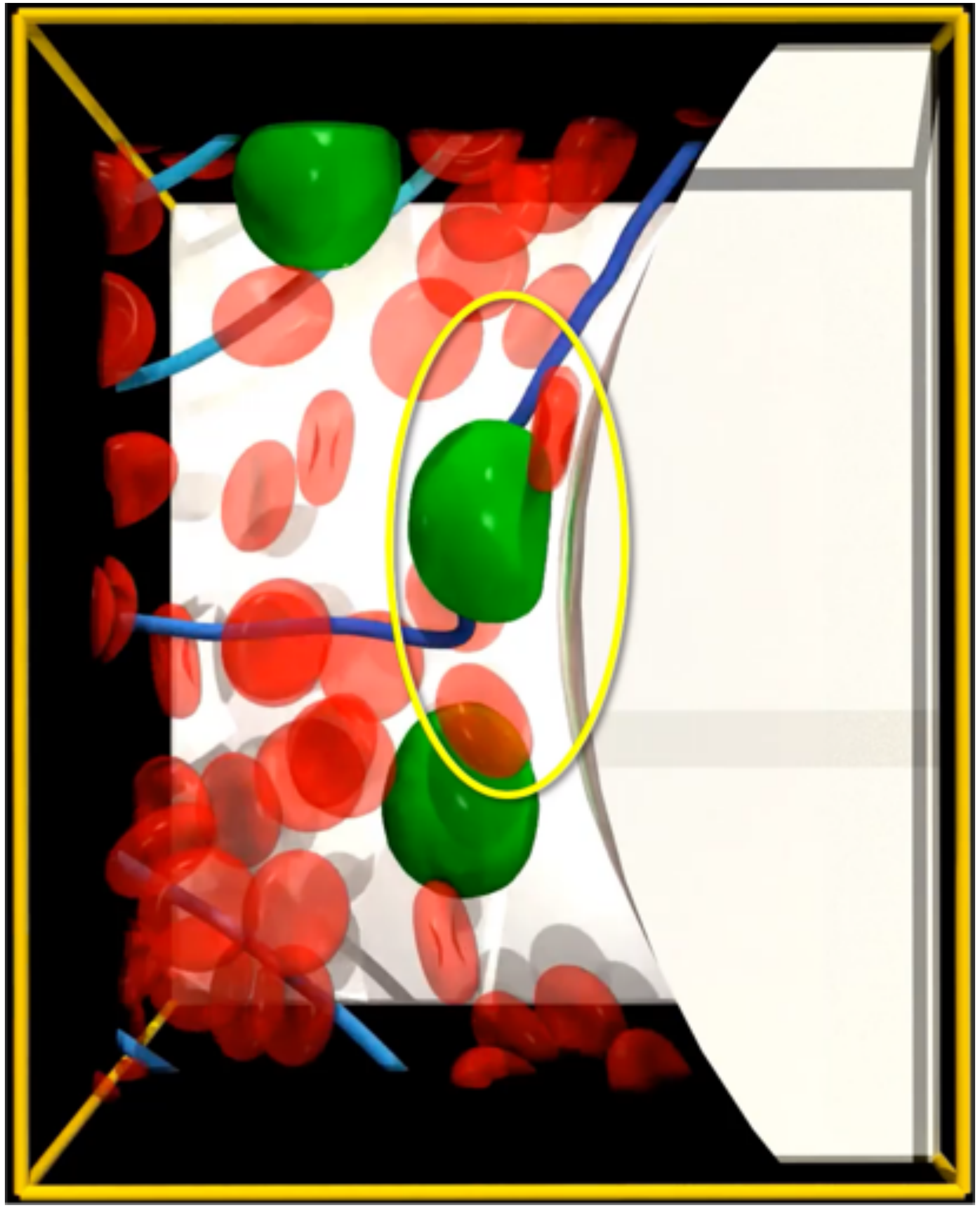
- Fluid solver: Lattice Boltzmann Method (LBM)
  - Highly efficient in parallel processors
  - Palabos software library
- Cell membrane: coarse-grained molecular dynamics
  - Particle based method
  - LAMMPS software library
- Fluid-Solid Interaction: immersed boundary method
  - In-house interface code



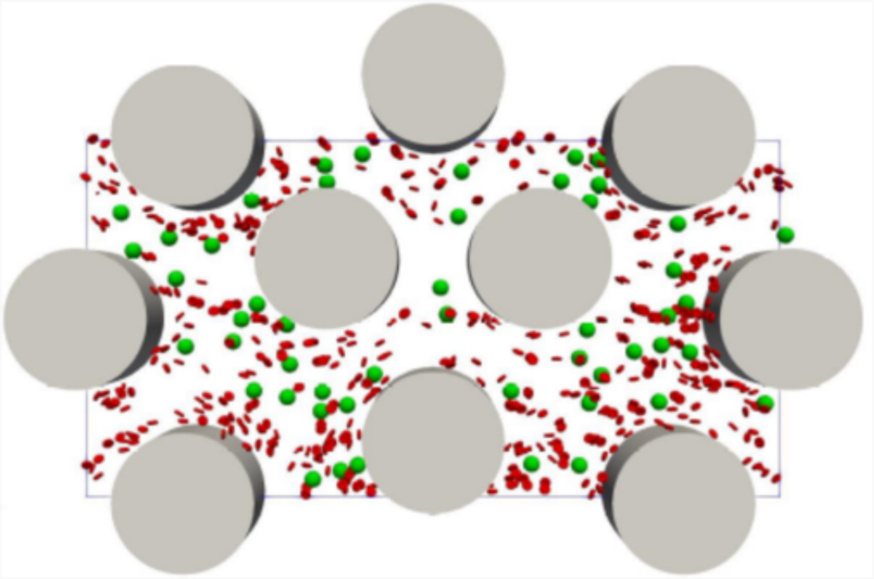
Adhesion locations

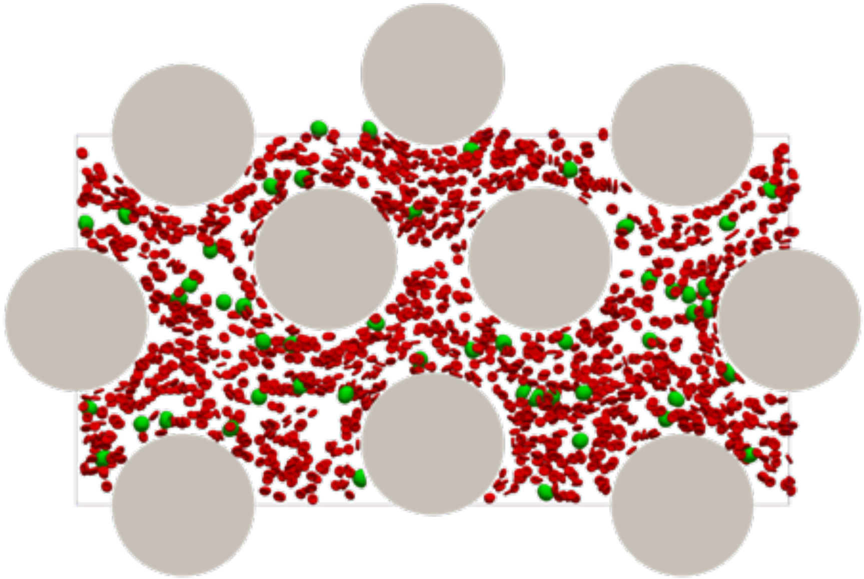
- Impact of
  - CTC size
  - Ligand coating density

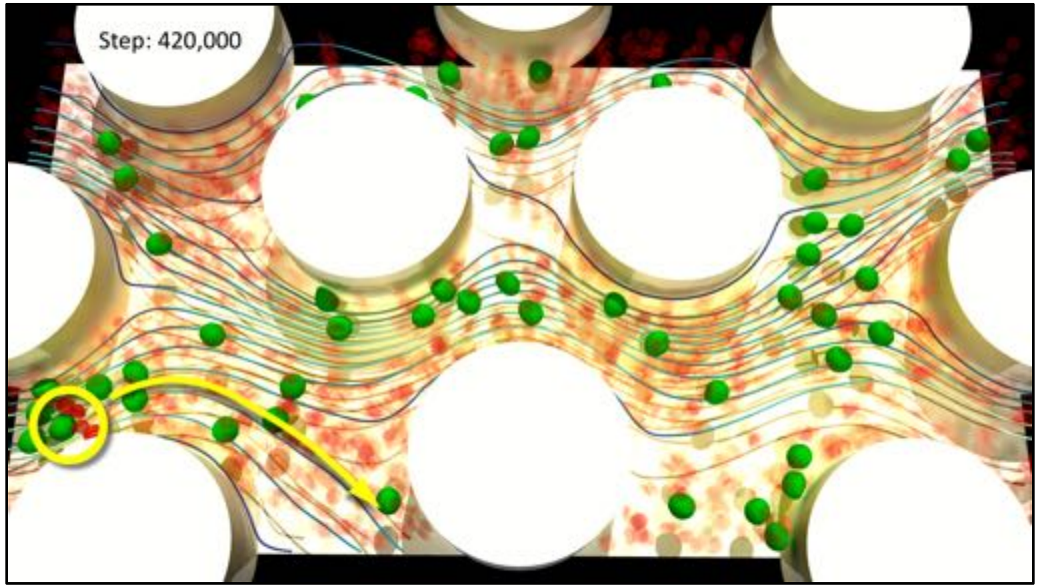


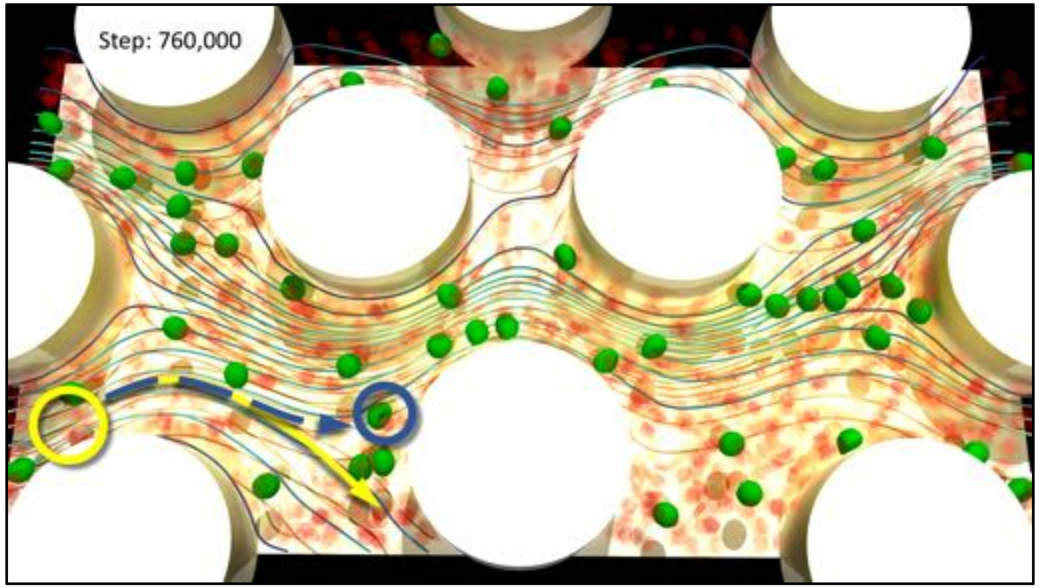


Transport Trajectories



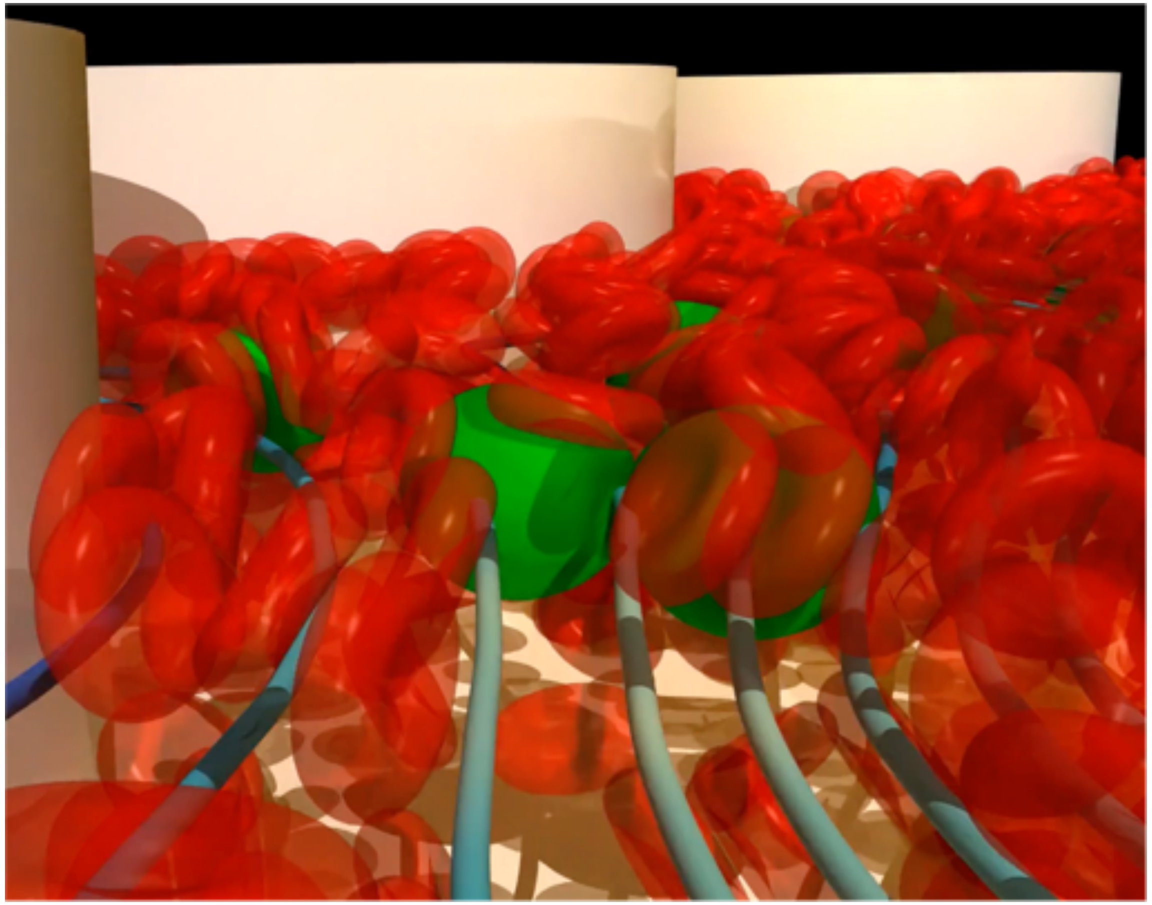






Animation

- Simulation
  - Computed on Mira at ALCF
  - 512 processors
  - ~20,000 core hours
  - ~2 Million time steps
- Visualization
  - ParaView
    - Surface rendering of cells
    - Streamline generation of fluid
    - Temporal interpolation of fluid
  - OSPRay



Visualization of Flow of Circulating Tumor Cells Mixed with Blood Cell Suspensions in Microfluidics

Science:

Jifu Tan

Northern Illinois University, Department of Mechanical Engineering

Michael Hood

Northern Illinois University, Department of Mechanical Engineering

Visualization:

Joseph A. Insley

Northern Illinois University, School of Art and Design

Argonne National Laboratory

Michael E. Papka

Northern Illinois University, Department of Computer Science

Argonne National Laboratory

Silvio Rizzi

Argonne National Laboratory

Janet Knowles

Argonne National Laboratory

Success Stories: Cosmology

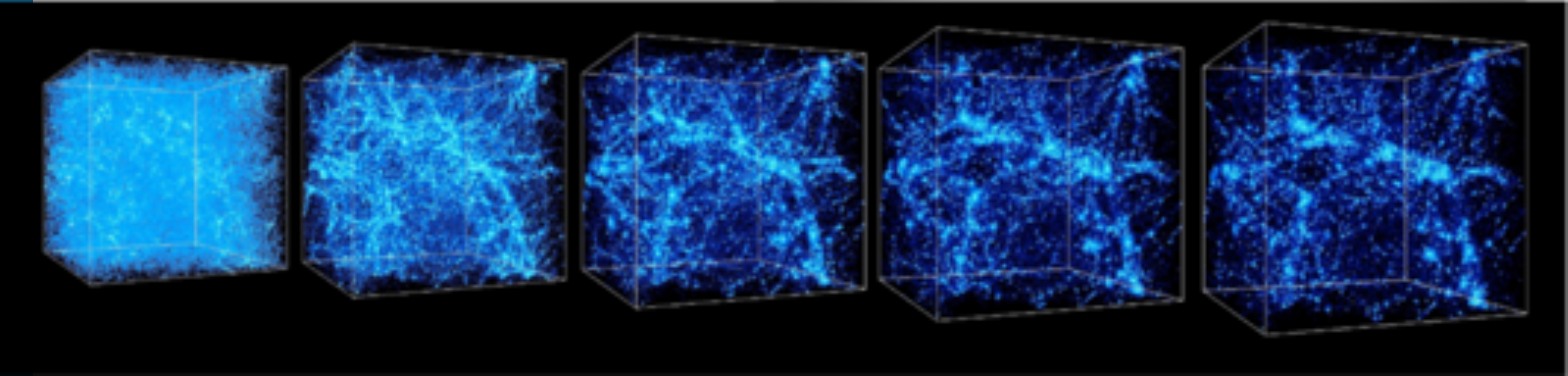
The Science

Slide courtesy the HACC Team

Cosmological N-body: Gravity in an Expanding Universe

- Gravity dominates on the largest scales
- Solve Vlasov-Poisson equations
- Use N-body methods with mass tracer particles
  - —6D phase space of positions + velocities
- Large-scale structures condense out of initially smooth distribution of matter

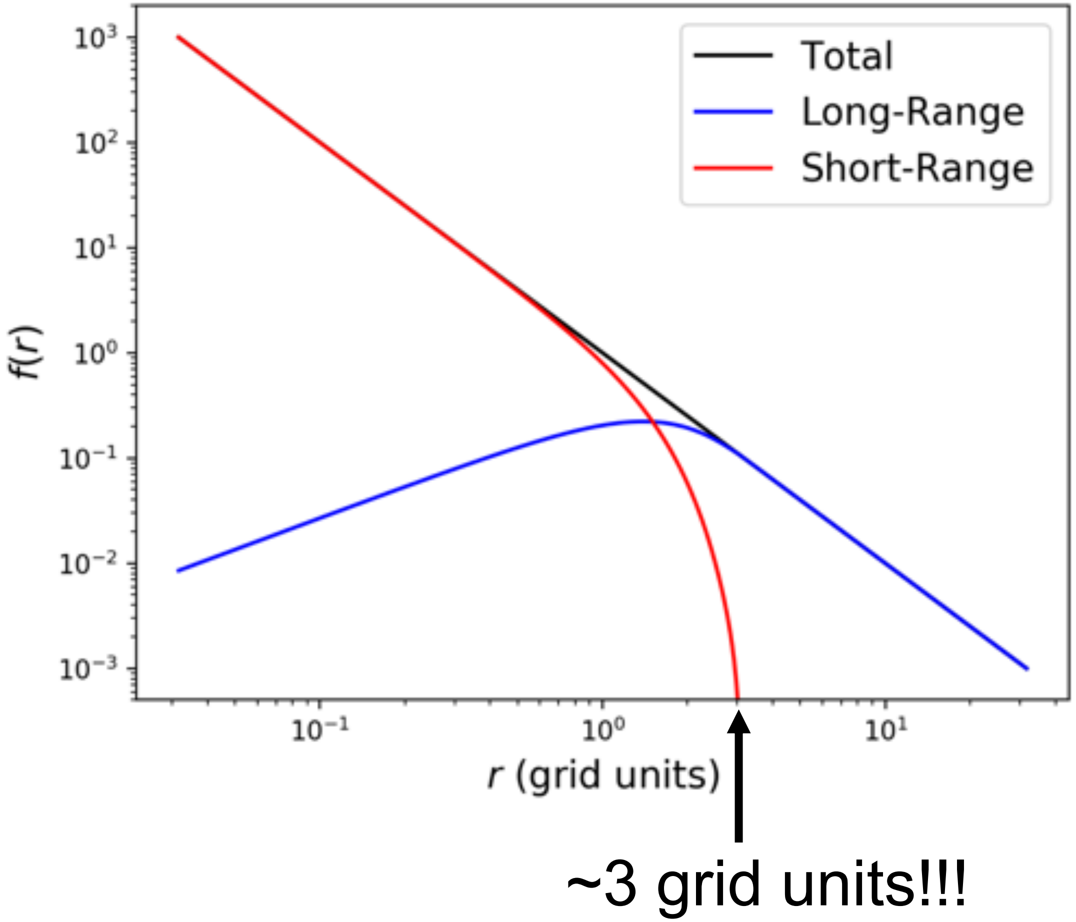
$$
\begin{aligned}
\frac{\partial f_i}{\partial t} + \dot{\mathbf{x}}\frac{\partial f_i}{\partial \mathbf{x}} - \nabla\phi\frac{\partial f_i}{\partial \mathbf{p}} &= 0, \qquad \mathbf{p} = a^2\dot{\mathbf{x}}, \\
\nabla^2\phi &= 4\pi G a^2(\rho(\mathbf{x},t) - \langle\rho_{\mathrm{dm}}(t)\rangle) = 4\pi G a^2 \Omega_{\mathrm{dm}}\delta_{\mathrm{dm}}\rho_{\mathrm{cr}}, \\
\delta_{\mathrm{dm}}(\mathbf{x},t) &= (\rho_{\mathrm{dm}} - \langle\rho_{\mathrm{dm}}\rangle)/\langle\rho_{\mathrm{dm}}\rangle, \\
\rho_{\mathrm{dm}}(\mathbf{x},t) &= a^{-3}\sum_i m_i \int d^3\mathbf{p}\, f_i(\mathbf{x},\dot{\mathbf{x}},t).
\end{aligned}
$$



HACC: Gravity Force Splitting

Slide courtesy the HACC Team

- HACC = Hardware/Hybrid Accelerated Cosmology Code
  - Gravity is infinite and unshielded
  - 1 kpc force resolution in 1 Gpc box, 10^6 dynamic range
- Operator splitting
  - Kicks: forces used to updated velocities, positions fixed
    - Long-range: Particle-Mesh, deposit onto grid, FFT-based Poisson solver, ~10^4 resolution from ~10k^3 grids, requires double precision
    - Short-range: Particle-Particle interactions, FLOPS intense, maximize architecture, ~10^2 resolution, single precision sufficient
  - Stream: velocities used to update positions, velocities fixed
  - Symplectic integration
- HACC Spectral Force Handover Technology
  - Use low-order Cloud-in-Cell (CIC) deposit
  - Spectral shaping reduces noise and emulates smoother deposit
  - Extremely compact, ~3 grid units, limit particle comparisons



The Last Journey
An Extreme-Scale Simulation on the Mira Supercomputer

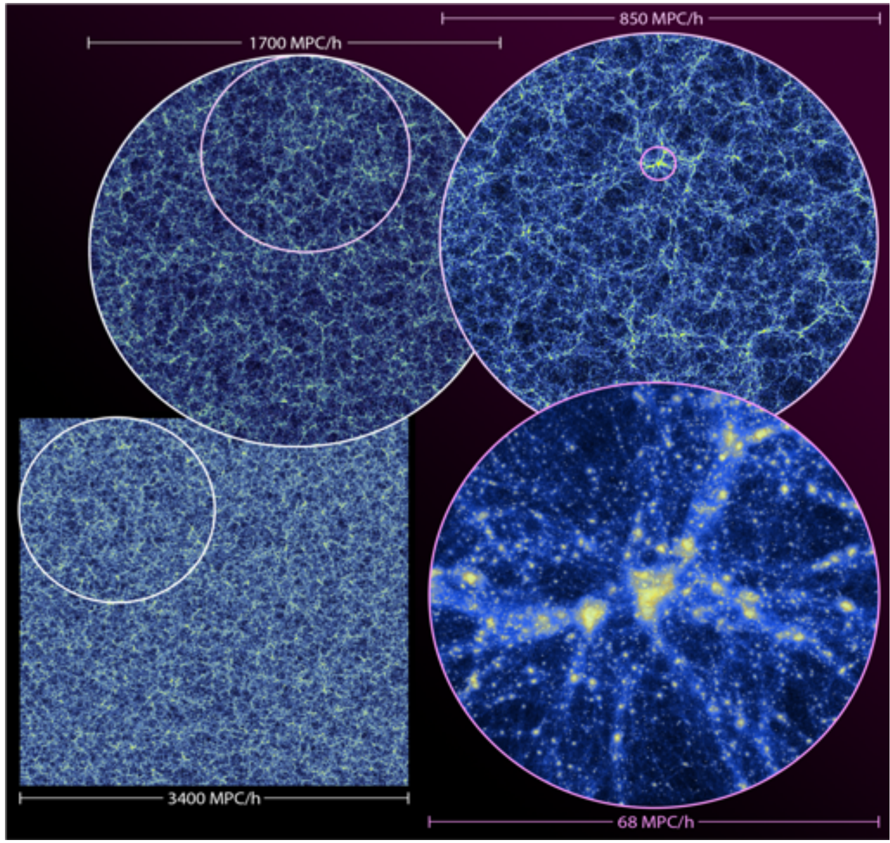
Slide courtesy the HACC Team

The Science

- Observational tools: measurements of the growth and distribution of structures on the largest scales, including how galaxies group into clusters and how those clusters are distributed.

- Determine how the content of the universe is divided between normal (baryonic) matter, dark matter, and dark energy, and also to try to understand more about the physical properties of dark matter and dark energy.

- Current and near future astronomical sky surveys can measure positions of millions or billions of galaxies. Massive simulations of cosmological structure formation have become indispensable tools for interpreting the measurements



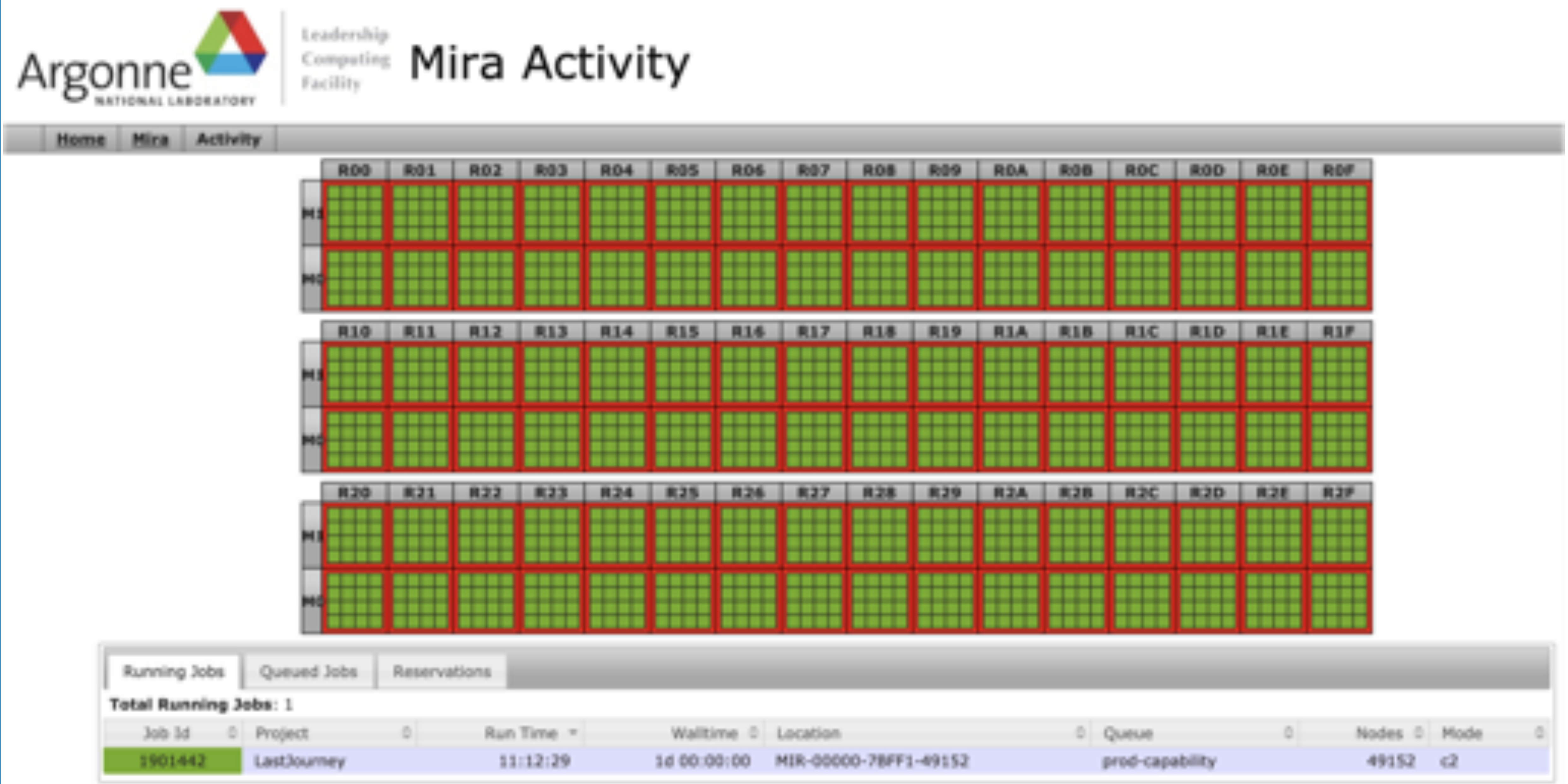
Zoomed views of a thin density slice of the simulation volume, ending with the largest cluster in the simulation. (Heitmann et al 2021)

The Last Journey
An Extreme-Scale Simulation on the Mira Supercomputer

Slide courtesy the HACC Team

The Impact

- The team used almost 400 million core-hours over the last six months of operation on the Mira supercomputer
- The Hardware/Hybrid Accelerated Cosmology Code (HACC) was used to simulate the movements of 1.24 trillion particles tracing the mass distribution in the universe over cosmic time.
- HACC's CosmoTools infrastructure was used to perform the largest data analysis and reduction operations in-situ, dramatically reducing the amount of data storage required while still producing a rich set of outputs that will be used to support a wide variety of cosmological measurements.
- Outputs from the Last Journey simulation will be used to help with planning and analysis of current and upcoming sky surveys including the Dark Energy Spectroscopic Instrument (DESI), the Legacy Survey of Space and Time (LSST), the Stage 4 Cosmic Microwave Background experiment (CMB-S4), and the SPHEREx space telescope.

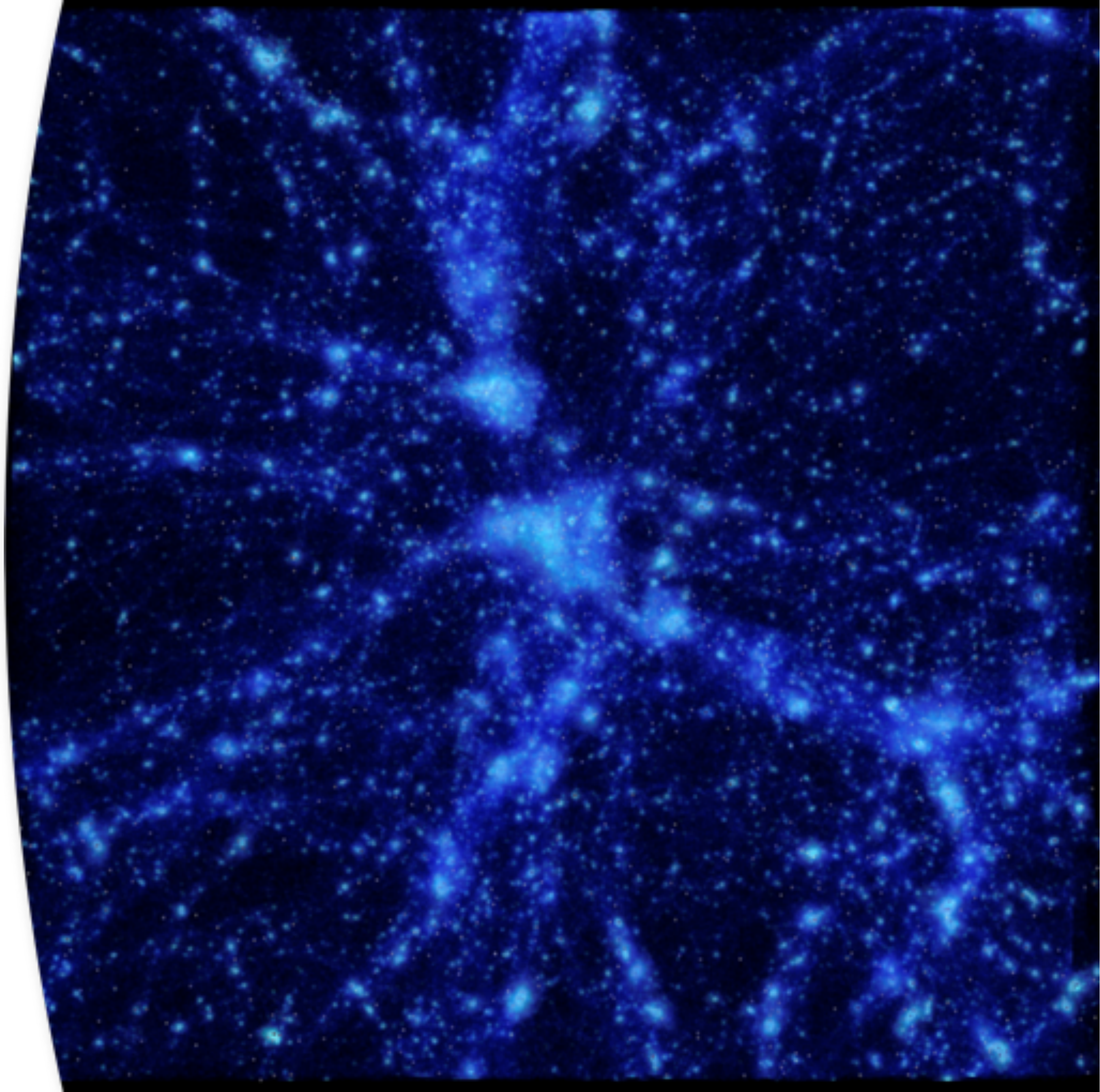

Argonne NATIONAL LABORATORY
Leadership Computing Facility
Mira Activity
Home
Mira
Activity
R00 R01 R02 R03 R04 R05 R06 R07 R08 R09 R0A R0B R0C R0D R0E R0F
R10 R11 R12 R13 R14 R15 R16 R17 R18 R19 R1A R1B R1C R1D R1E R1F
R20 R21 R22 R23 R24 R25 R26 R27 R28 R29 R2A R2B R2C R2D R2E R2F
M1
M0
Running Jobs
Queued Jobs
Reservations
Total Running Jobs: 1
Job Id
Project
Run Time
Walltime
Location
Queue
Nodes
Mode
1901442
LastJourney
11:12:29
1d 00:00:00
MIR-00000-7BFF1-49152
prod-capability
49152
c2

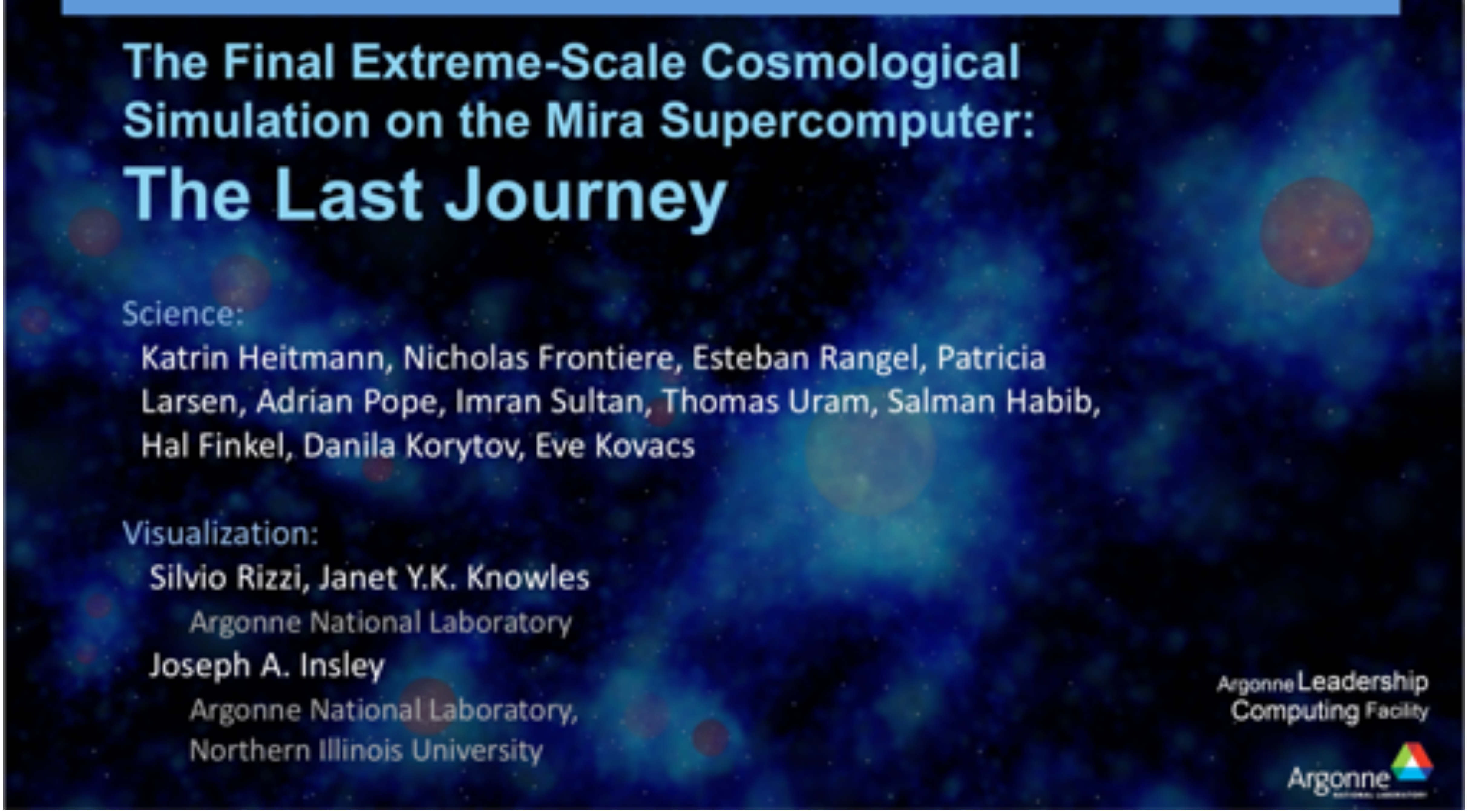
Visualizing data from the Last Journey simulation

- Convert data from simulation to regular grids representing density for volume rendering with our vl3 framework

- Convert raw particles and halo particles to VTK for visualization with ParaView

- Combine point rendering and sphere glyphs in ParaView to highlight halos and filament structure



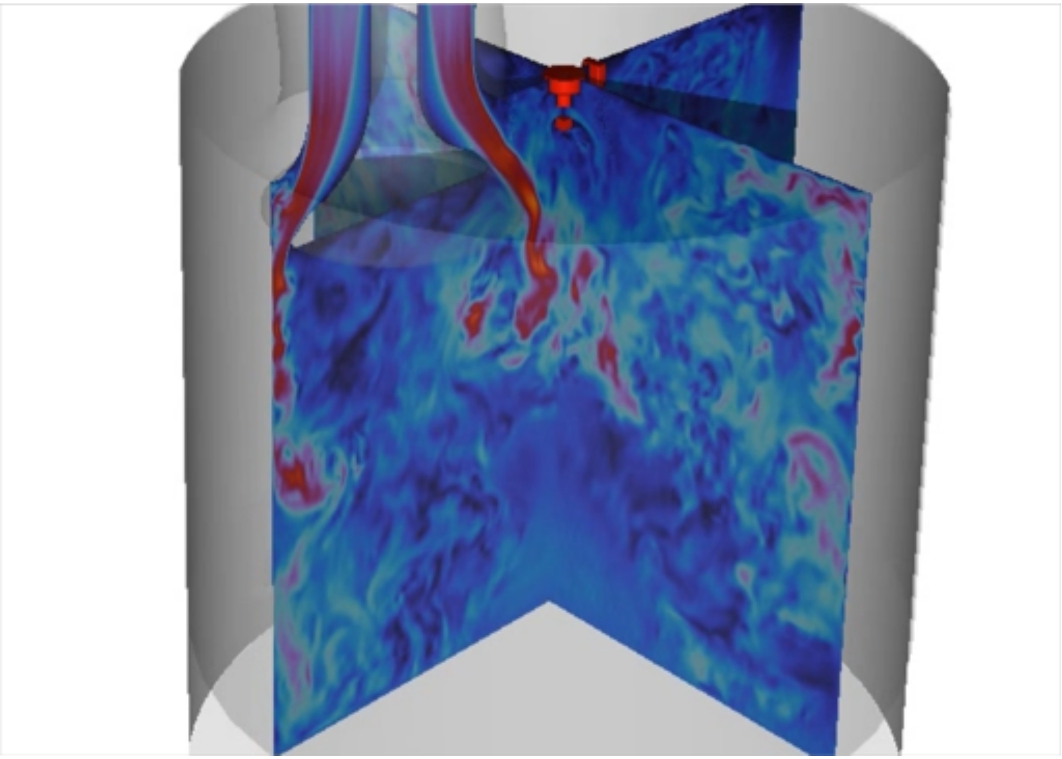

The Final Extreme-Scale Cosmological
Simulation on the Mira Supercomputer:
The Last Journey
Science:
Katrin Heitmann, Nicholas Frontiere, Esteban Rangel, Patricia Larsen, Adrian Pope, Imran Sultan, Thomas Uram, Salman Habib, Hal Finkel, Danila Korytov, Eve Kovacs
Visualization:
Silvio Rizzi, Janet Y.K. Knowles
Argonne National Laboratory
Joseph A. Insley
Argonne National Laboratory,
Northern Illinois University
Argonne Leadership Computing Facility
Argonne

Success Stories: Combustion

The Science

Internal Combustion Engine Simulation

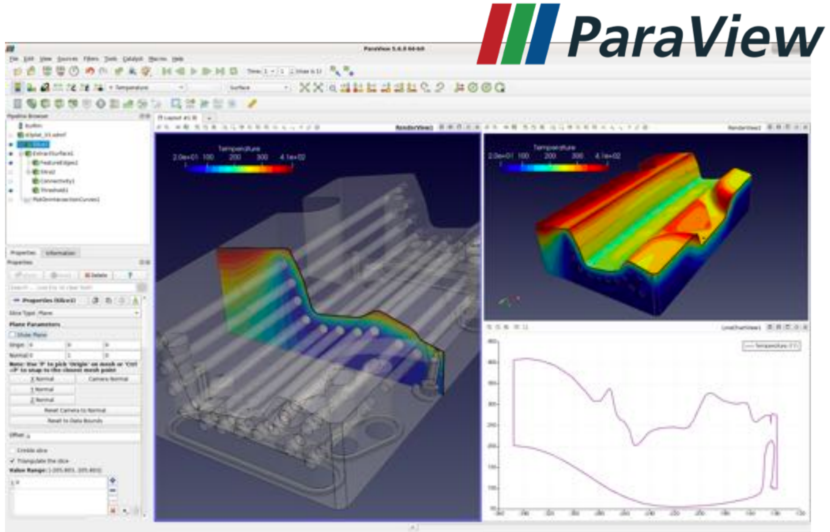
TCC Engine Apparatus

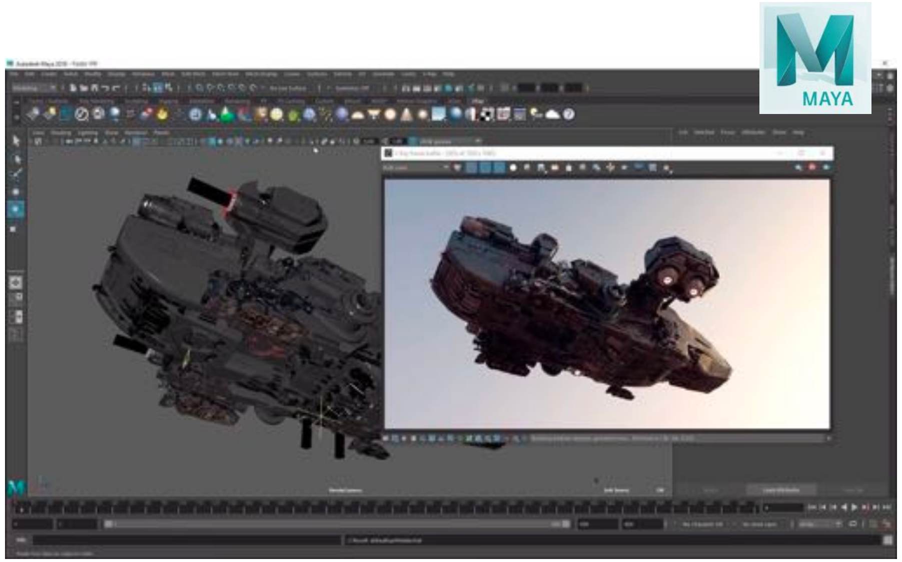


Fluid Dynamics Simulation

Goal

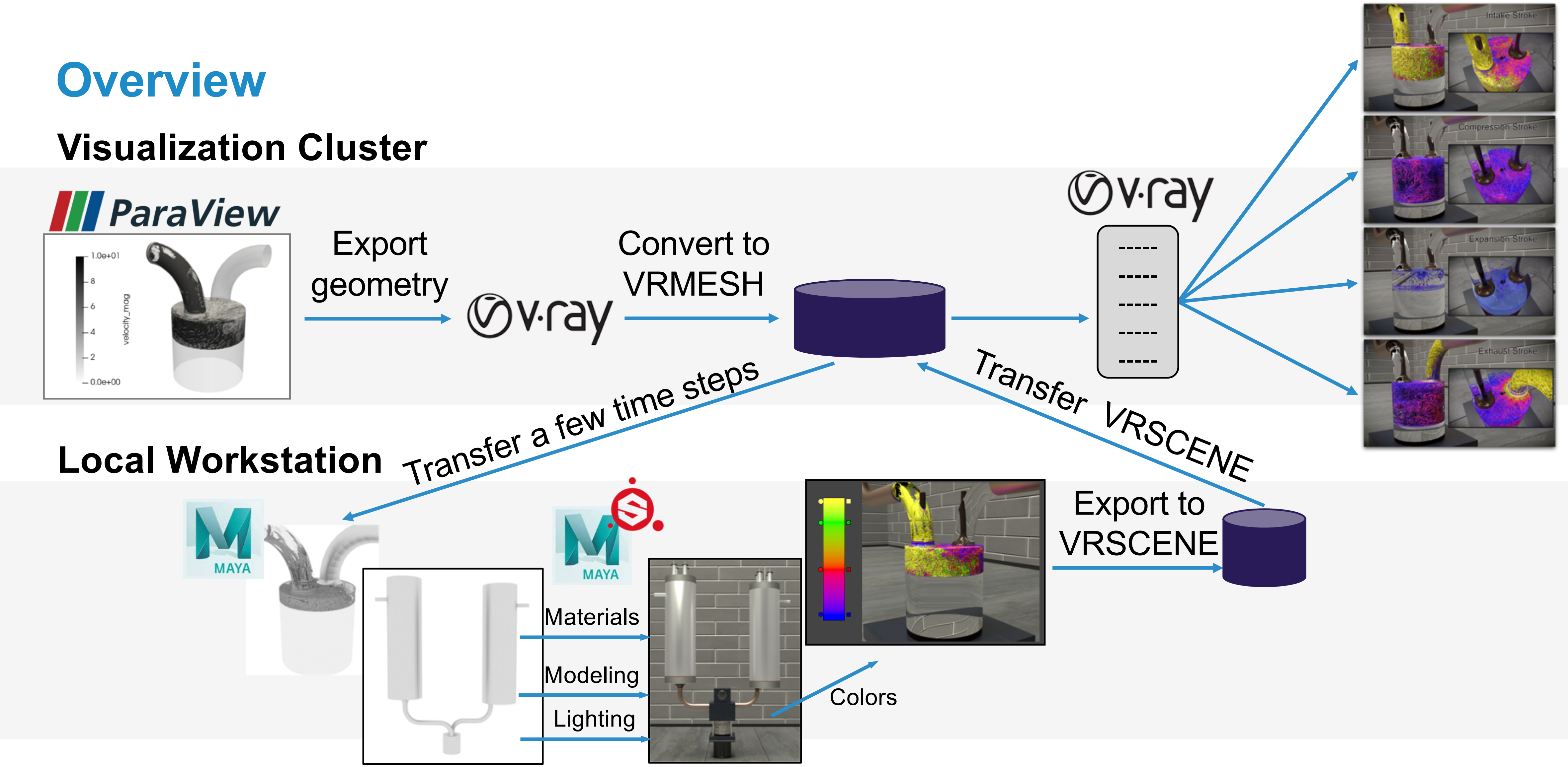
- Provide context to tell the story/explain the science
- Integrate production tools into the existing visualization pipeline
- Tools used:
  - —ParaView
  - —Maya
  - —Substance Painter
  - —V-Ray
  - —Custom scripts and HPC Resources
  - —ffmpeg
  - —Premiere/After Effects

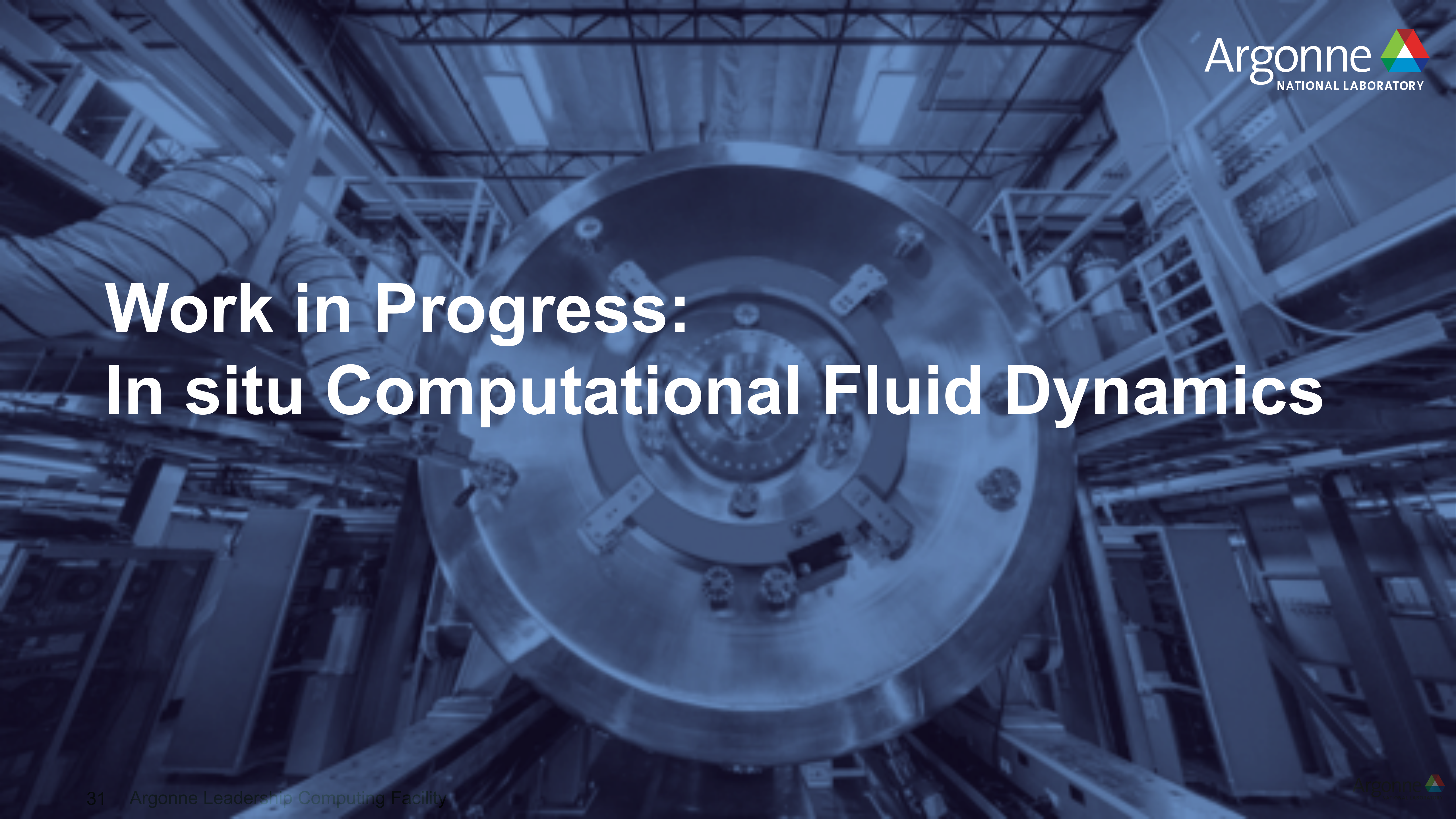




THE VISUALIZATION PIPELINE

Overview





Work in Progress: In situ Computational Fluid Dynamics

In Situ Visualization and Analysis: Motivation

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

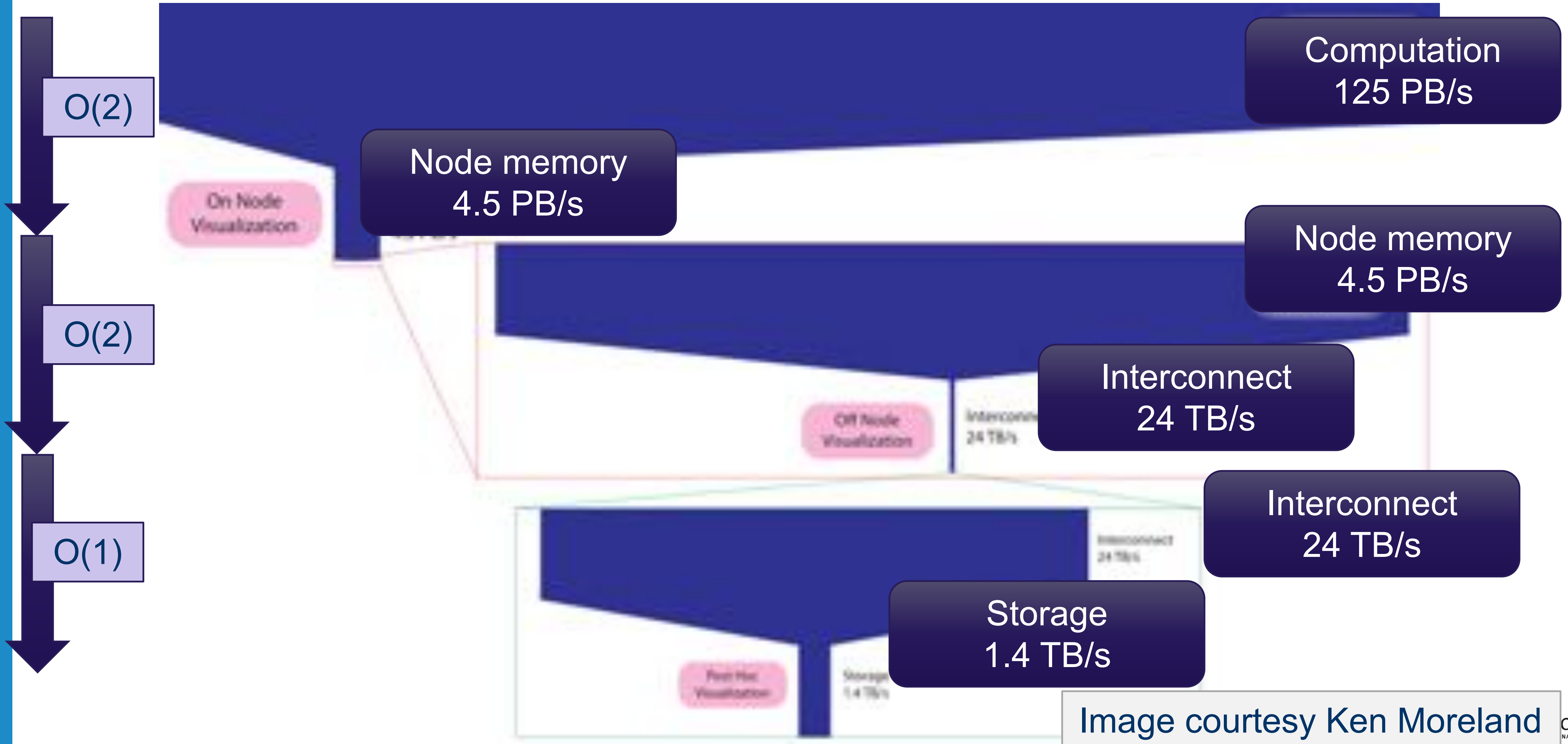
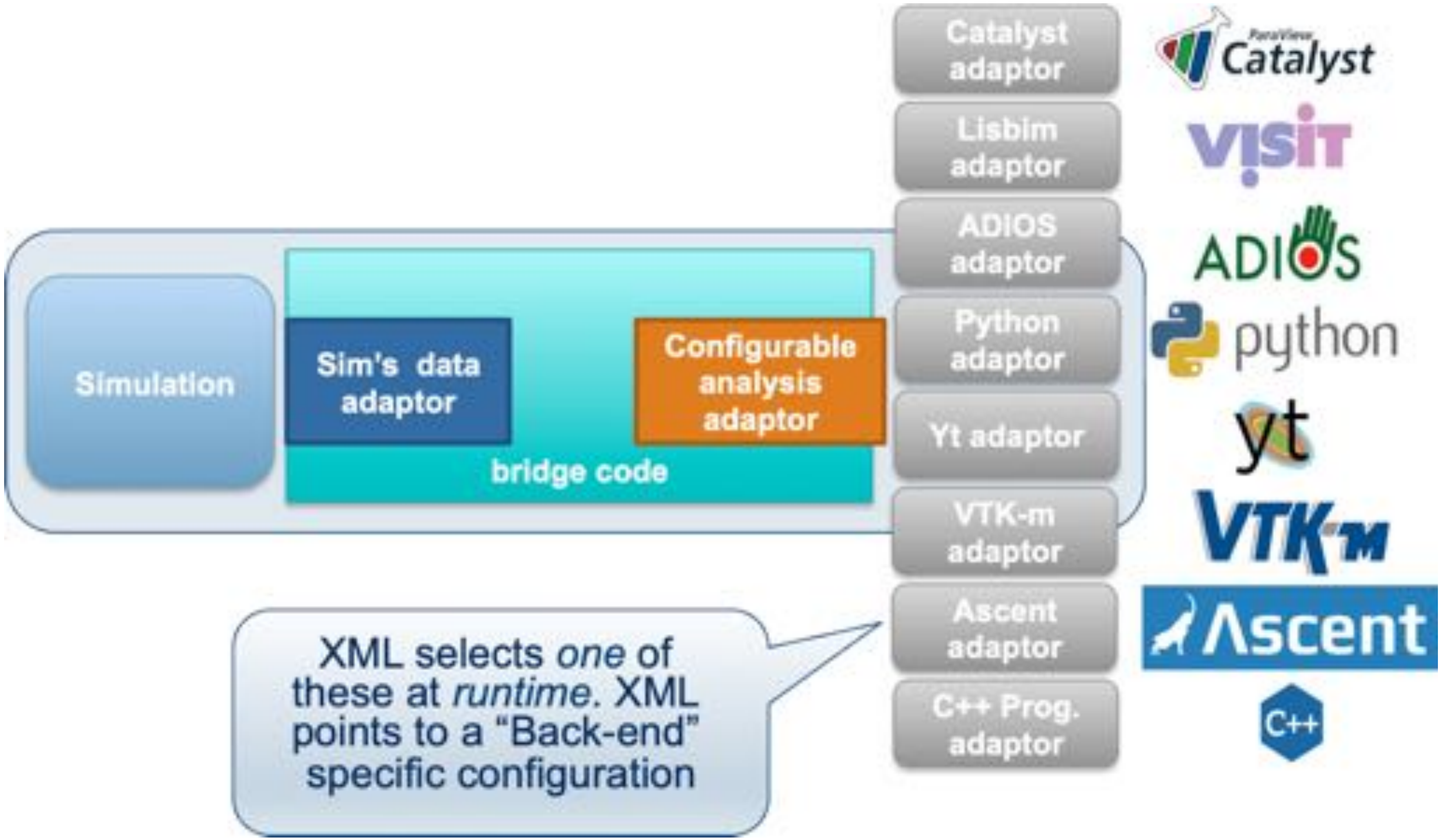


Image courtesy Ken Moreland

Background

SENSEI: Scalable Environments for Scientific Explorations In Situ

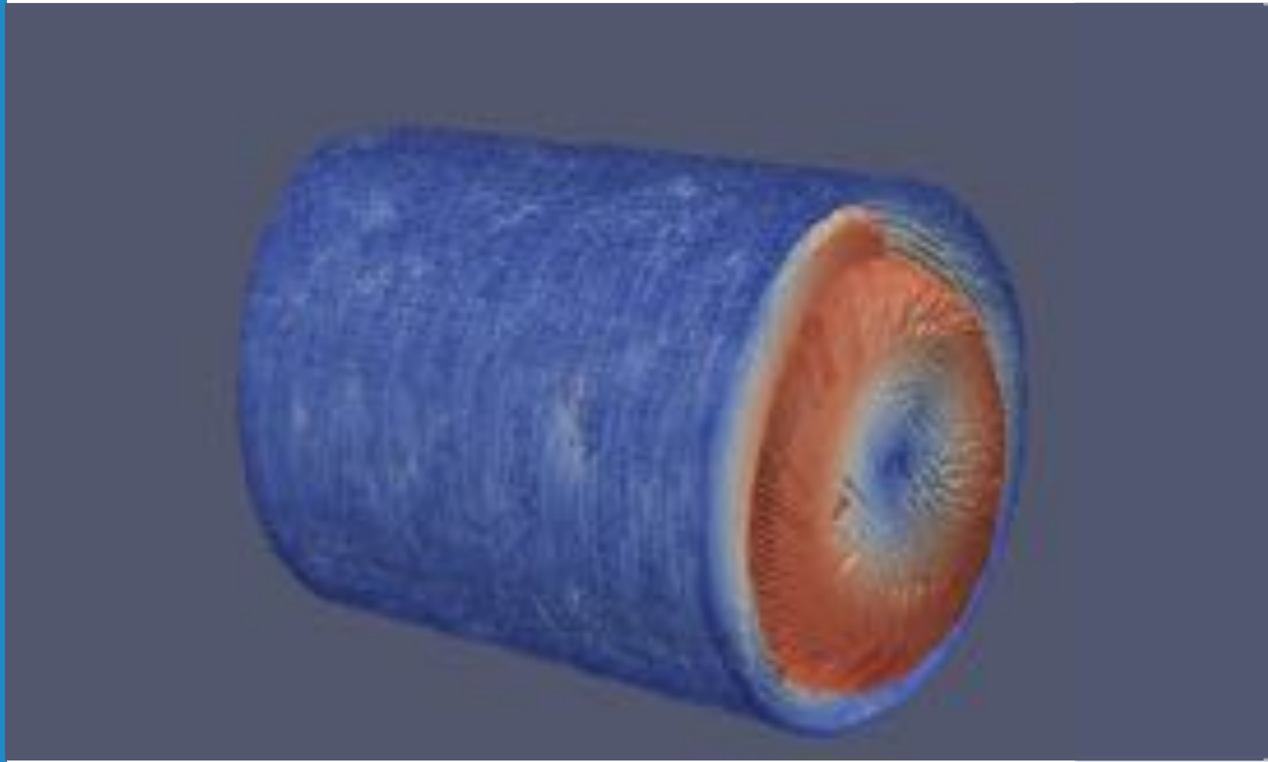


Loring et al., EGPGV 2020.

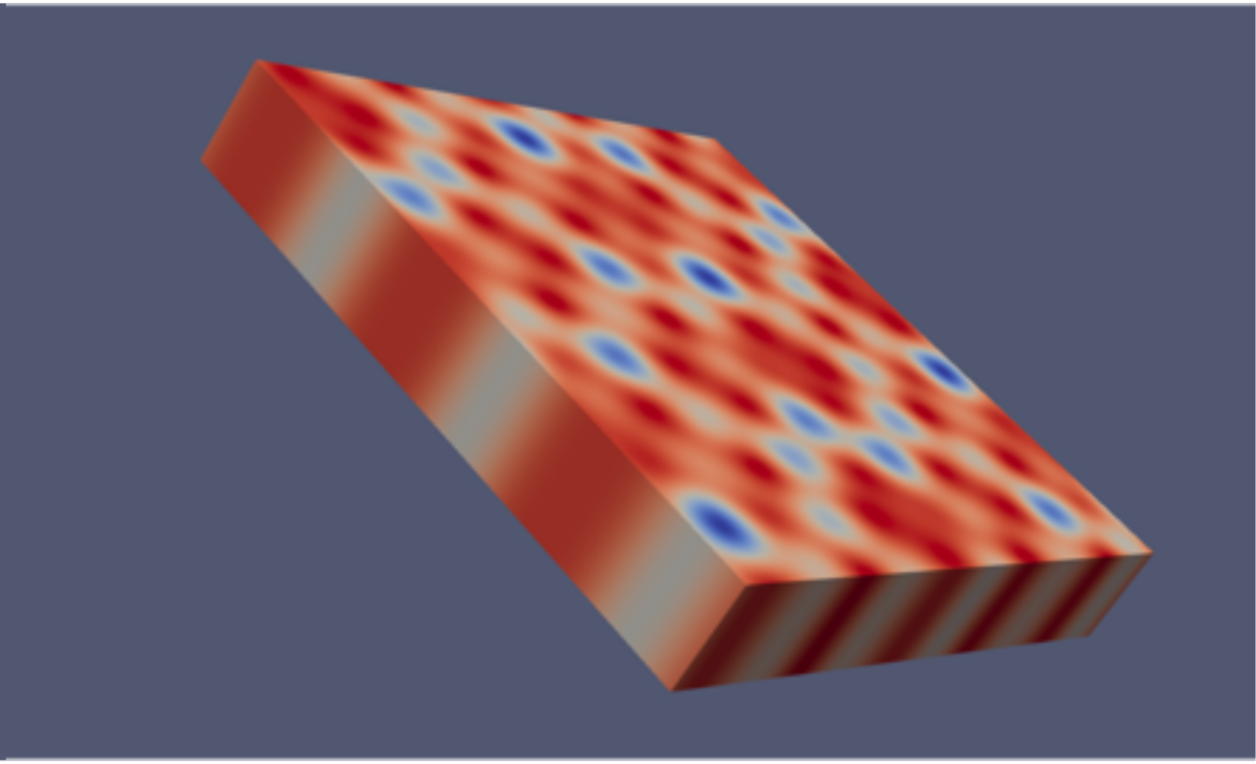
The Science

Nek5000/NekRS are computational fluid dynamics codes that simulate laminar, transitional, and turbulent incompressible or low Mach-number flows with heat transfer and species transport, which can handle general 2D and 3D domains.

We demonstrated successful full-machine runs (4096 Theta nodes / 262,144 ranks) and generated, at scale, *in situ* renderings, without any data I/O. Additionally, we successfully used ADIOS2 for *in transit* visualization at smaller scales using heterogeneous pipelines. Currently hardening the implementation as well as planning larger runs. The *in transit* runs used 512 ranks on Theta to run the simulation and 96 ranks on Cooley to render it.



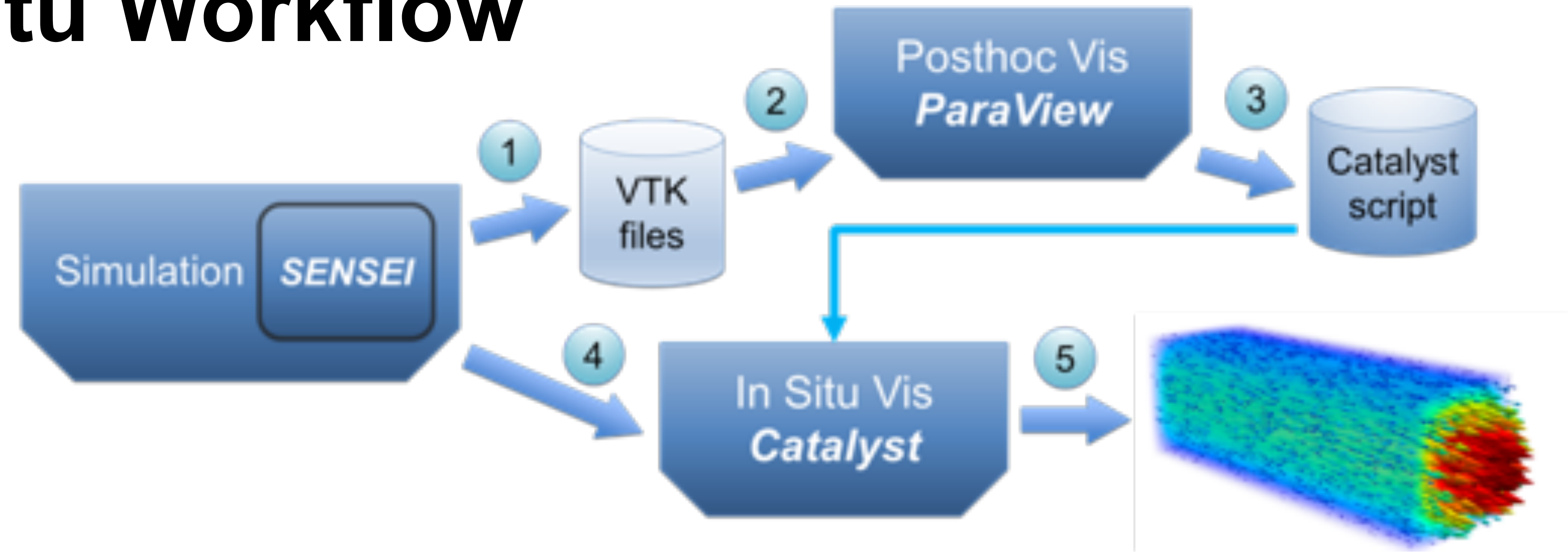
Swirling, turbulent air during the compression stroke of an internal combustion engine. Data: Saumil Patel



NekRS with SENSEI instrumentation


Argonne
Leadership
Computing
Facility
Theta Activity
Home
Theta
2021 Aug 5 7:14:54 PM
Options
Running
Starting
Queued
Reservations
Total Running Jobs: 3
Job Id
Project
Nodes
Start Time
Run Time
Walltime
Queue
Mode

Blood Flow In Situ Workflow

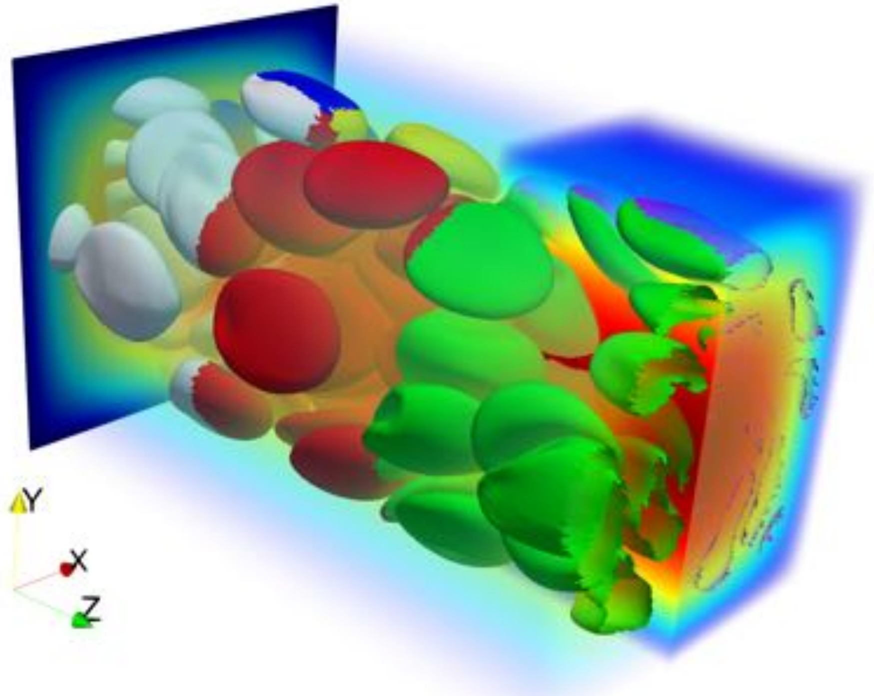


Workflow to create and in situ visualization pipeline[1]

1. Run a small simulation with SENSEI and save a representative data set (PostHocIO).
2. Visualize with ParaView and define desired parameters.
3. Save the state as a Catalyst script
4. Run the simulation with Catalyst enabled and using the Catalyst script.
5. In situ visualization

[1] A. Bucaro et al. "Instrumenting Multiphysics Blood Flow Simulation Codes for In Situ Visualization and Analysis," *2021 IEEE 11th Symposium on Large Data Analysis and Visualization (LDAV),* pp. 88-89.

Parallelization Results:



An *in situ* visualization of a blood flow simulation generated using PostHocIO showing how fluid and cells are divided among six processors. Only one 3D subdomain of the fluid velocity is shown in the front for clear visualization.

Additional Resources

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
- *InSitu* Vis and Analysis

Argonne NATIONAL LABORATORY

QUESTIONS?

Joe Insley
insley@anl.gov

Silvio Rizzi
srizzi@anl.gov

Janet Knowles
jknowles@anl.gov

Victor Mateevitsi
vmateevitsi@anl.gov