Visualization 2020:

Can you read the bottom line

Charles 'Chuck' Hansen
SCI Institute
University of Utah

Visualization 2020:

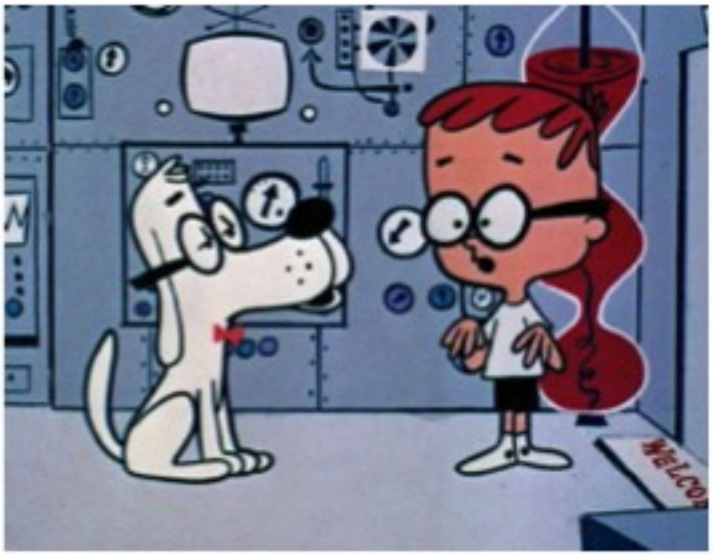
Can you read the bottom line

Charles 'Chuck' Hansen

S
A L
I S H
A N O R
A 2 0 2 0
V I S U A L
I Z A T I O N

www.eyechartmaker.com

Thirty years ago …



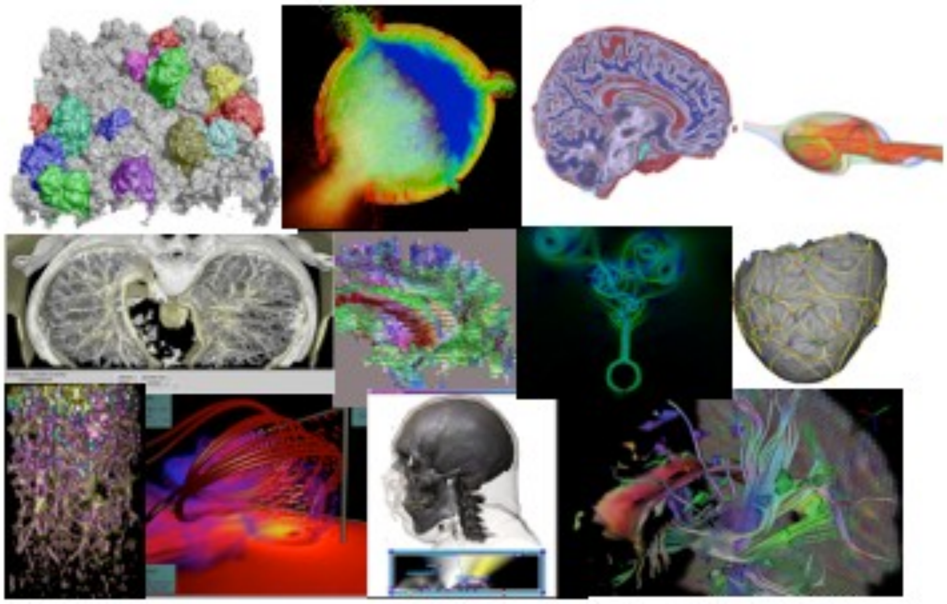
Credit: Peter Lee, Microsoft Research

Thirty years ago …

Visualization was frustratingly slow

Today …

Visualization spans all disciplines



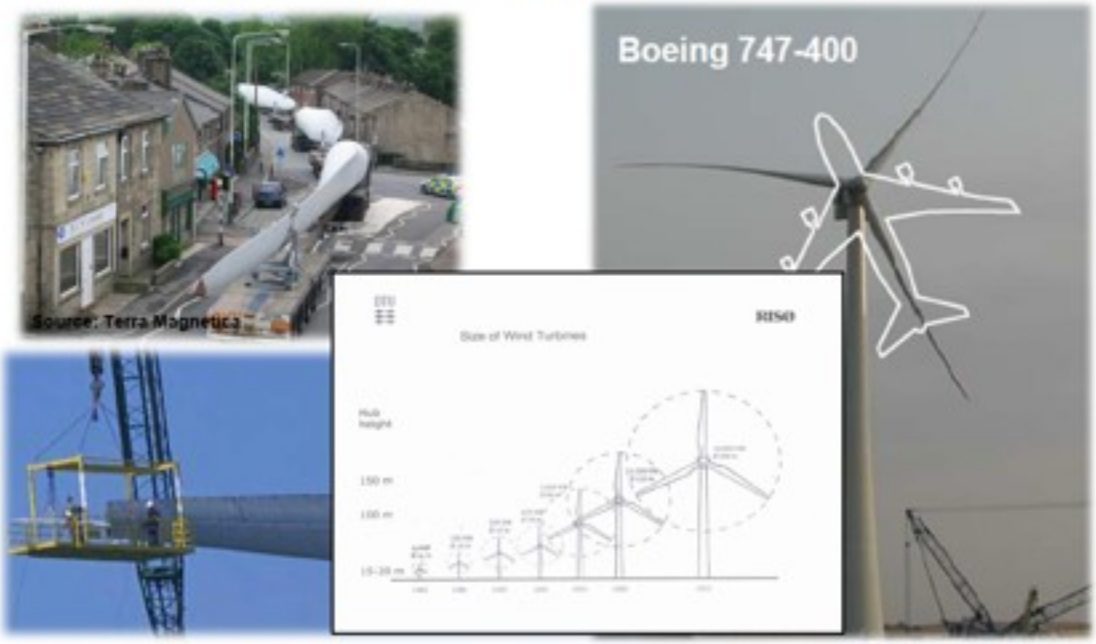
Applications Drive Visualization

Power is Critical for Exascale

Source: AMEC

Credit: William P. Mahoney III

Setting the Context - Scale



NCAR UCAR | NCARScience air • planet • people 2

Credit: William P. Mahoney III

Overarching Wind Energy Science Challenges

Boundary layer meteorology (0 to 200 m above ground) is not well understood nor is this layer well measured

The wind energy industry greatly under appreciates the complexity of the airflow in this layer

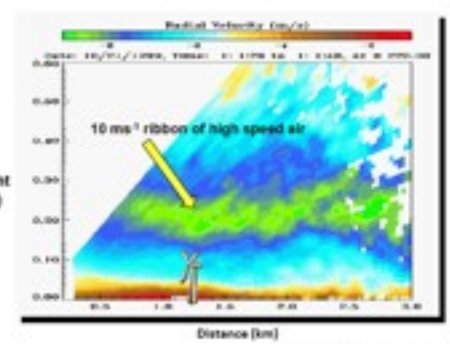
The wind industry has historically assumed less turbulence and more wind with height above the ground

Image source: Wind Measure International

NCAR UCAR | NCARScience air • planet •

Low-Level Jets of High Wind (U.S. Midwest)

Lidar (laser radar) measured wind velocity toward lidar

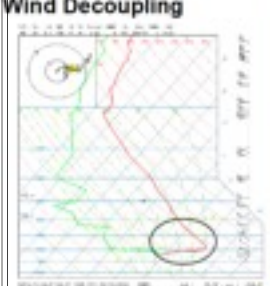


Courtesy, Robert Banta, NOAA

Low-level jet streams can damage wind generators

NCAR UCAR | NCARScience air • planet • people

Predicting Inversions – Wind Decoupling



Nocturnal inversion – Denver 19 September 2010
15 degree C difference over ~1500 ft

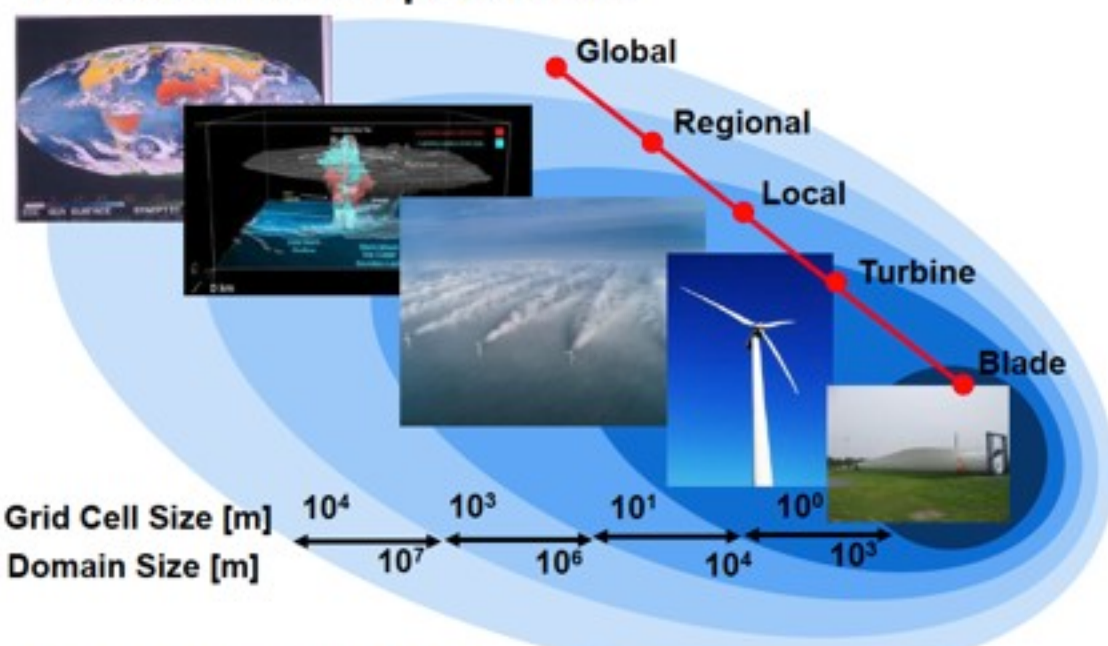
NCAR UCAR | NCARScience air • planet • people

Credit: William P. Mahoney III

Mother Nature Is Not Always Kind

Credit: William P. Mahoney III

Research in Complex Flows



Adopted from Mike Robinson (DOE/NREL)

Credit: William P. Mahoney III

Weather & Solar Energy Related Industry Issues

Financing	• Wind energy resource estimates at wind farm sites are over-estimated on average
Maintenance Costs	• Wind turbines are failing faster than predicted (up to 40% earlier)
Variability is Costly	• Wind & solar power variability complicate power integration and load balancing across the grid – requires reserves
Prediction Errors Costly	• Wind energy prediction has typical errors of 10-15% (flat terrain) to 15-25% (complex terrain)
Extreme Weather Impacts	• Wind turbines are not designed to handle extreme weather conditions (shear, ice, snow, high wind, etc.). More representative weather datasets are needed for turbine design

Credit: William P. Mahoney III

Weather & Solar Energy Related Industry Issues

Financing

Maintenance Costs

Variability is Costly

Prediction Errors Costly

Extreme Weather Impacts

WIRED GEAR SCIENCE ENTERTAINMENT BUSINESS SECURITY DESIGN OPINION MAGAZINE

SCIENCE

Charles C. Mann | Tuesday March 25

Renewables Aren't Enough. Clean Coal Is the Future

Coal supplies over 40 percent of global electricity needs, and that percentage is going up. The only real question is how to minimize the damage. Dan Winters Proof that good things don't always come in nice packages can be found by taking the fast train from Beijing to

Credit: William P. Mahoney III

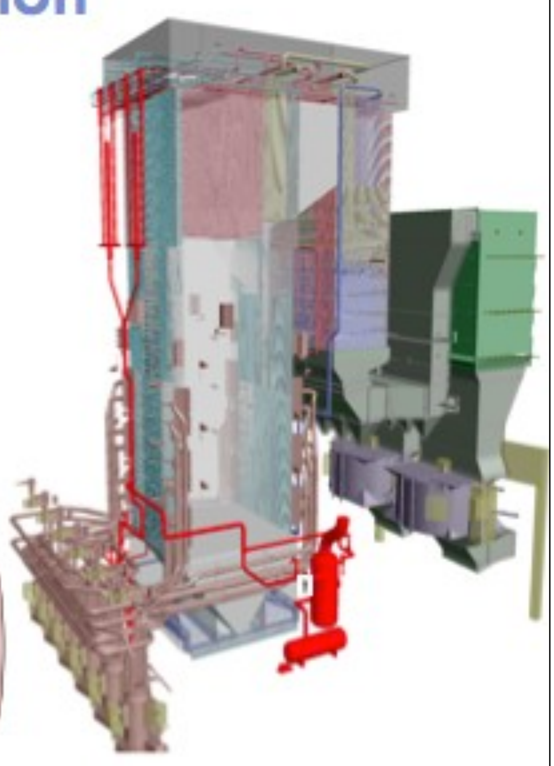
CCMSC

CARBON CAPTURE MULTIDISCIPLINARY SIMULATION CENTER

ALSTOM — Institute for Clean and Secure Energy, The University of Utah — SCI www.sci.utah.edu — Berkeley University of California — BYU Brigham Young University — NNSA National Nuclear Security Administration

Overarching Application

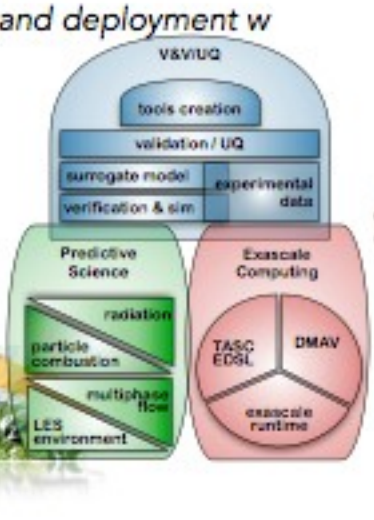
- high efficiency advanced ultra-supercritical (AUSC) oxy-coal tangentially-fired power boiler
  - *extreme computing*
  - *predictive science w hybrid validation/UQ*
    - expensive function evaluation
    - expensive data
  - *rapid design and deployment w Alstom*



CARBON CAPTURE MULTIDISCIPLINARY SIMULATION CENTER

ALSTOM

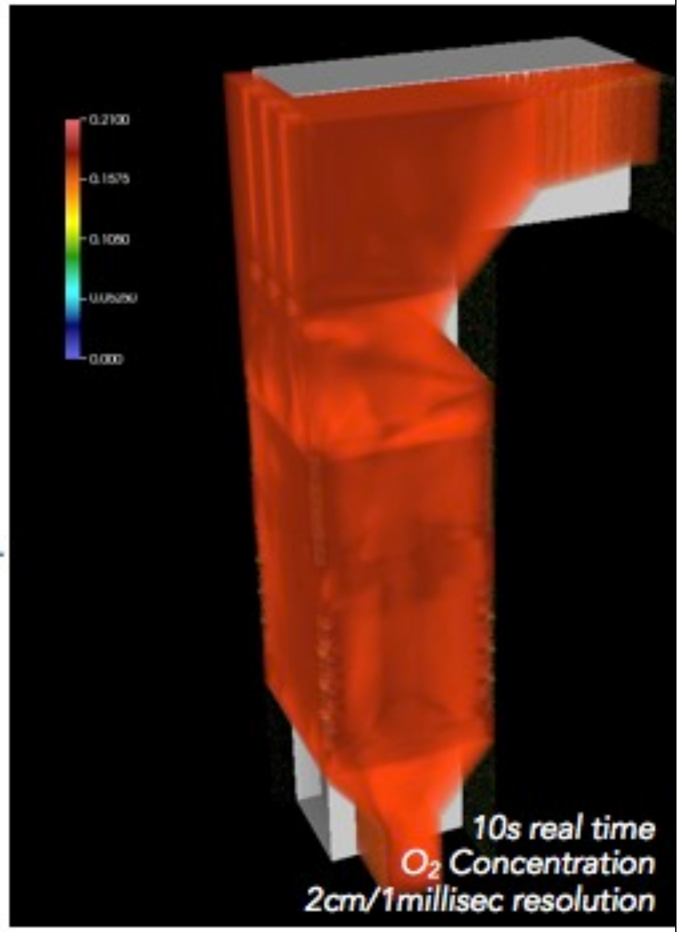
- *global reach: present in 100 countries*
- *2011/12 sales: $26.5 billion*
- *93,000 employees*



Why exascale?

- dynamics
  - pulverized-coal fuel
  - mixing & reaction
  - deposition
- multi-phase radiation
  - participating media
  - absorbing/emitting/scattering
- resolved physics
  - 1mm spatial resolution
  - 30 seconds of real time
  - 1microsecond temporal resolution
- validation
  - quantify degree of consistency w mmts.
  - reduce risk for rapid deployment
  - expensive function evaluation → exascale

CARBON CAPTURE
MULTIDISCIPLINARY
SIMULATION CENTER

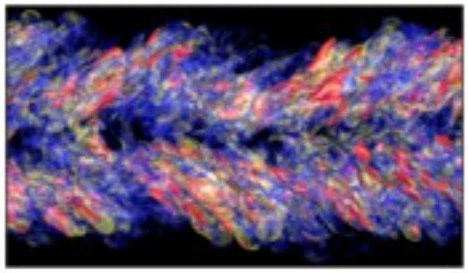


2020 Visualization

Challenges of Petascale DNS: Mountains of Data

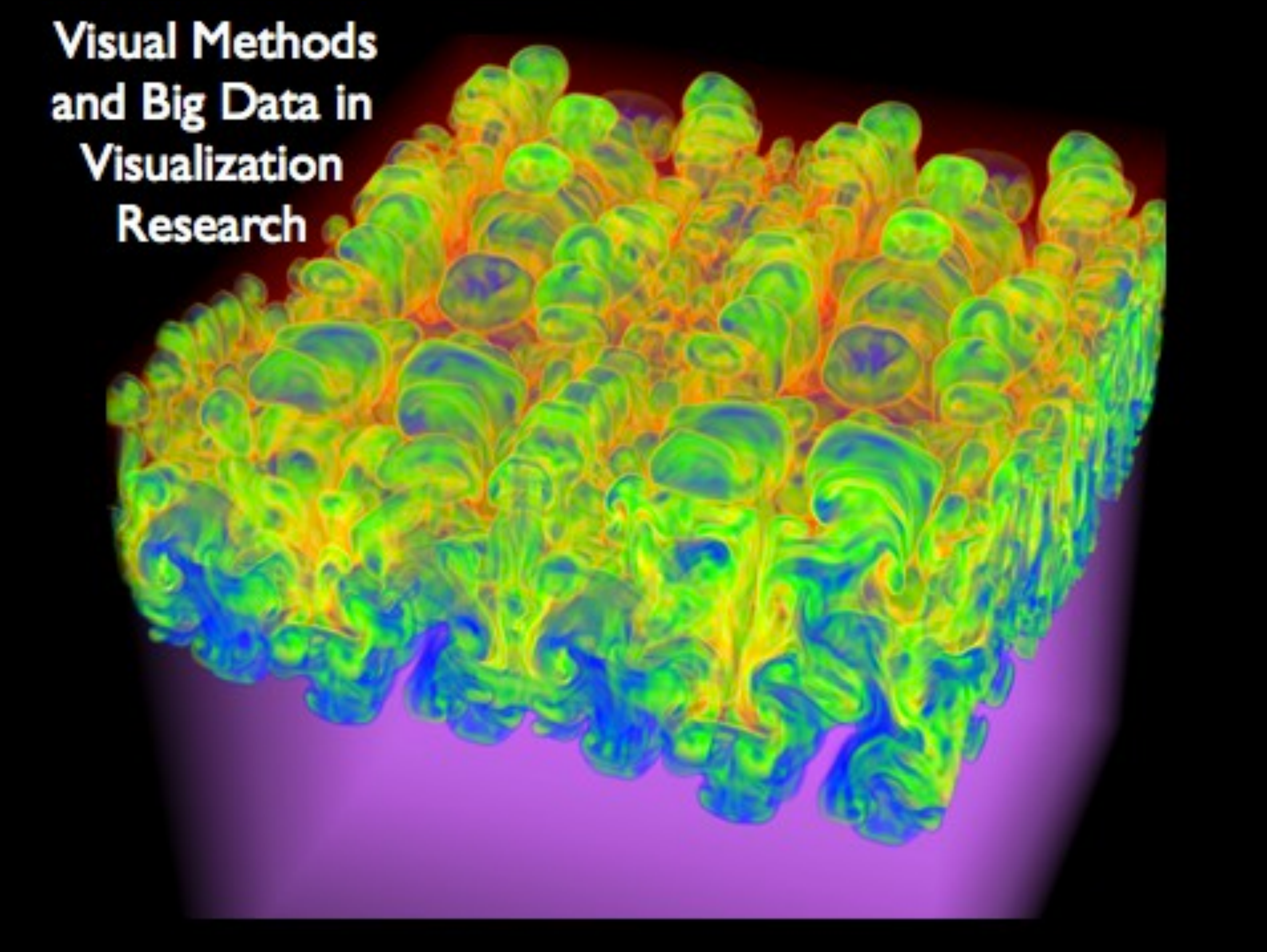
- Data size:
  - –O(3/4 PB) raw field data and 7TB of particle data on Jaguar CrayXT5, I/O 20 GB/s ADIOS
- Data complexity:
  - –Data is multi-variate (~50 species)
  - –Turbulence chaotic phenomena:
    - –Wide range of scales
    - –High intermittency, higher moments matter!
    - –Time-varying
    - – Organized coherent motions
    - –Non-locality important for spatial and temporal correlation of scalars and vectors

•HPSS storage facility at NERSC



Exascale Implies Big Data

Dan Ariely
Center for Advanced Hindsight
Duke University



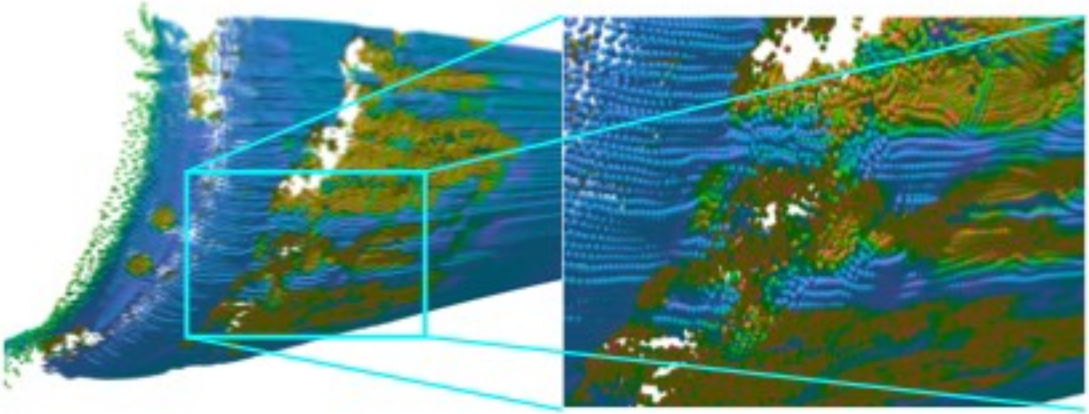
Perceptual Cues for Shading

Jim Blinn:

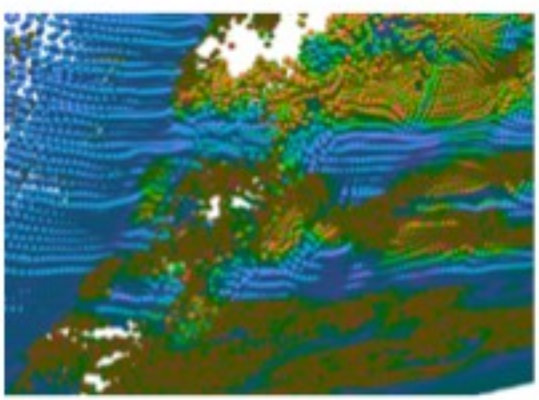
"Lighting models... there's something that always bothered me about lighting models. Bui Tuong Phong is[was] a great guy and he did wonderful work ... The thing is, this has no physical basis whatsoever ... I'd like to see cosine power retired and better approximations being done."

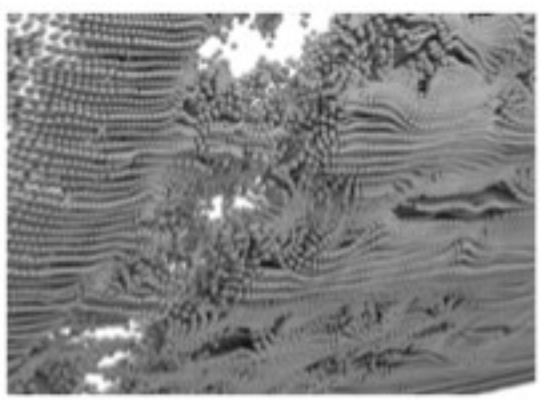
- SIGGRAPH 98 Keynote Speech

Indirect Shading of Particles: Ray tracing

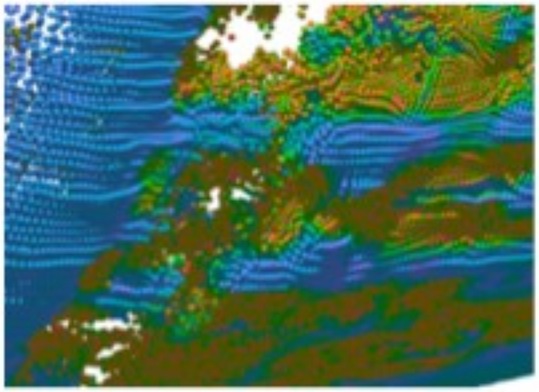


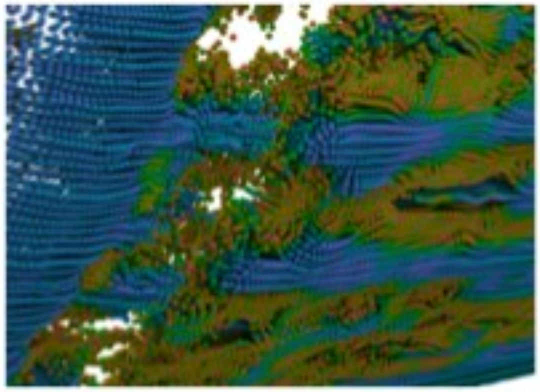
Indirect Shading of Particles: Ray tracing



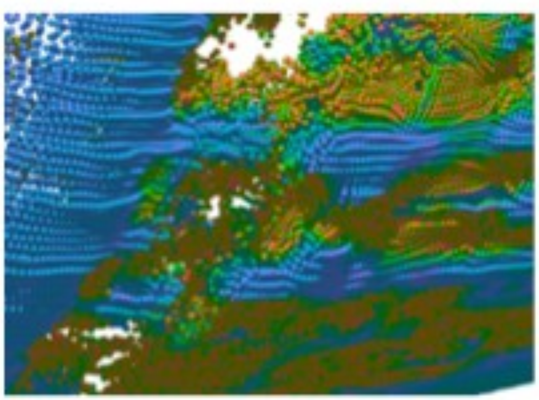


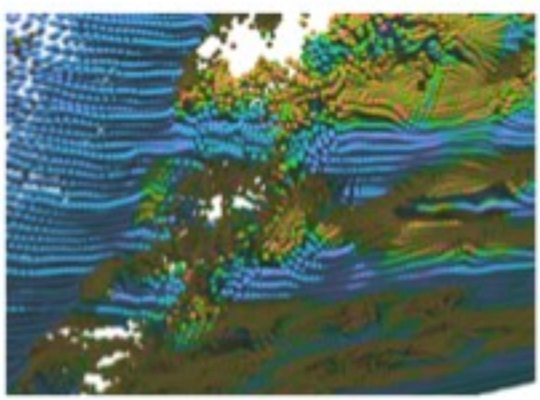
Indirect Shading of Particles: Ray tracing

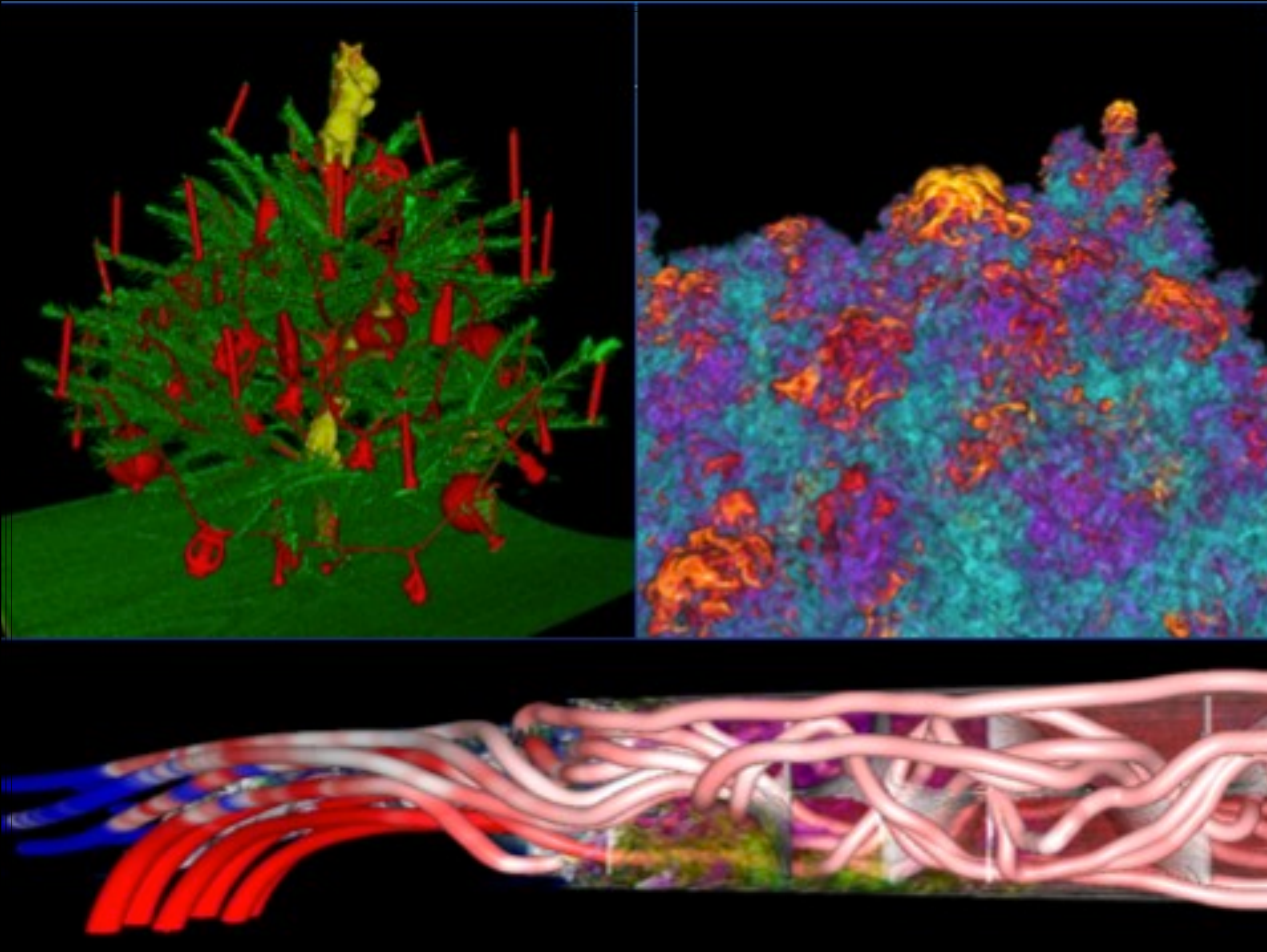


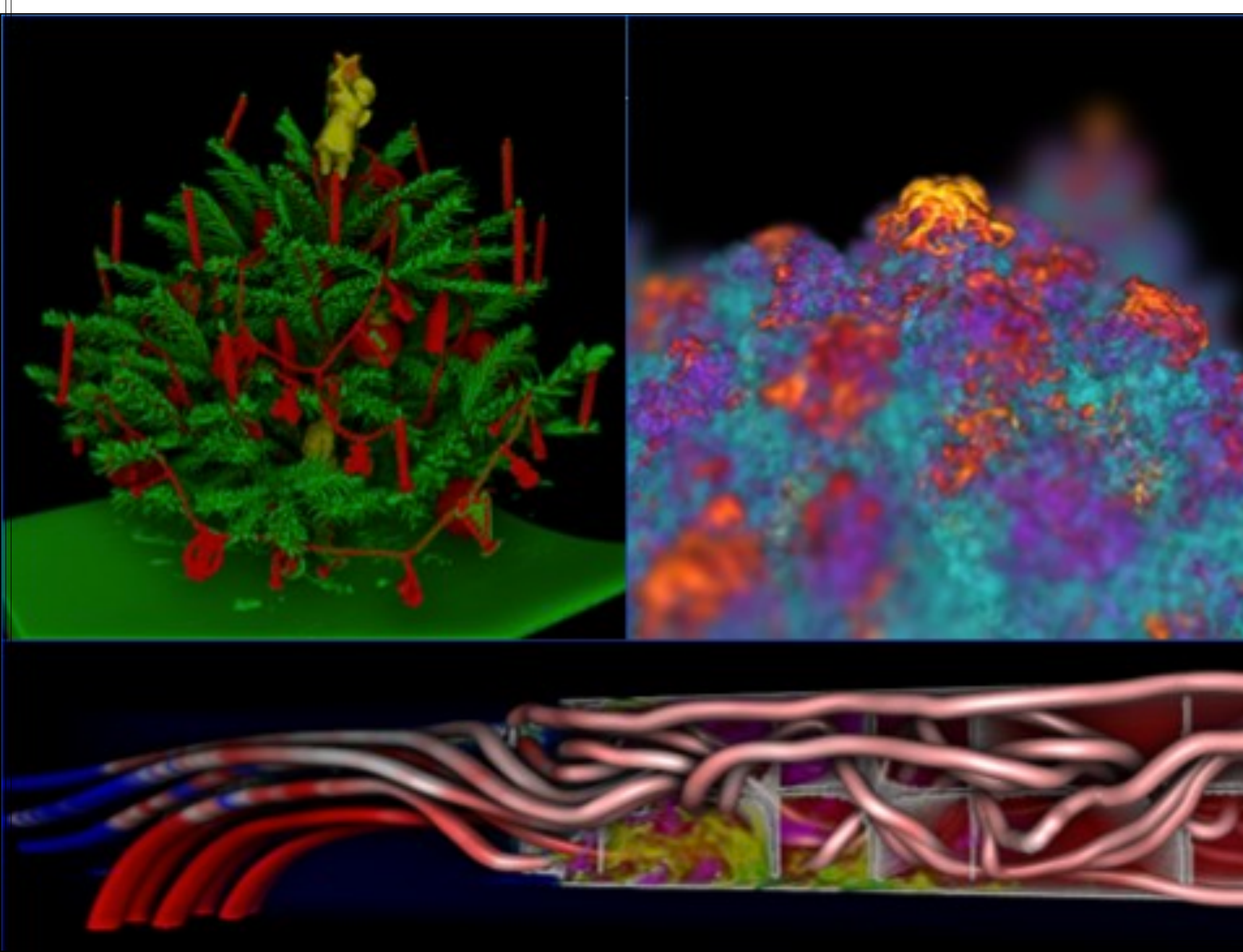


Indirect Shading of Particles: Ray tracing

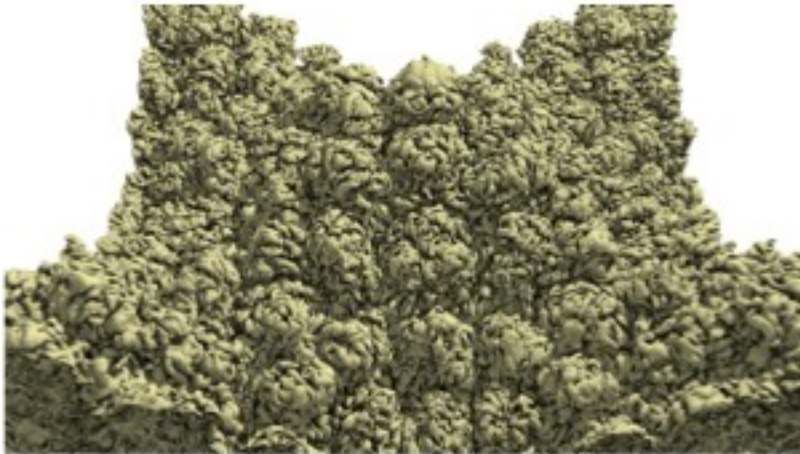


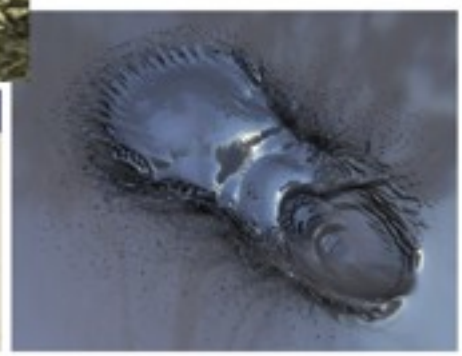




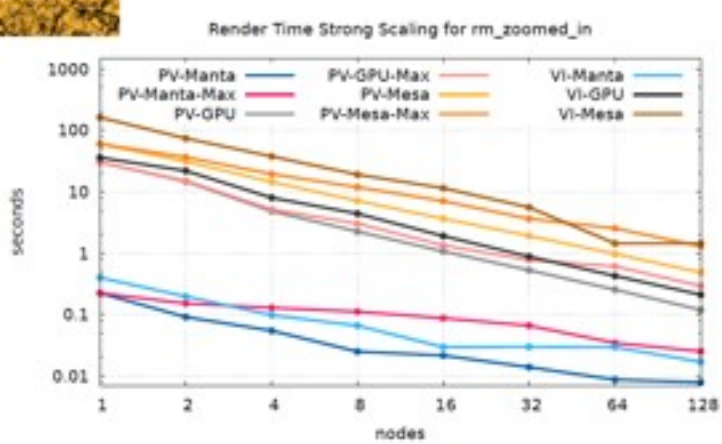


Current Cluster Ray Tracing Visualization





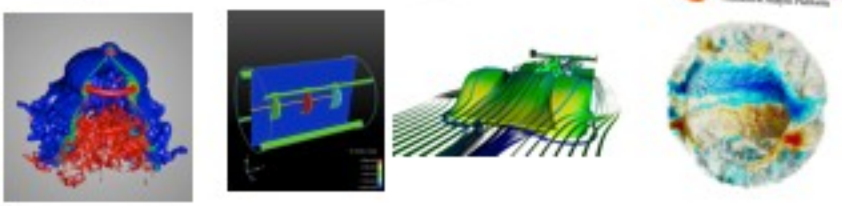
Ray-tracing over-riding VTK renderer



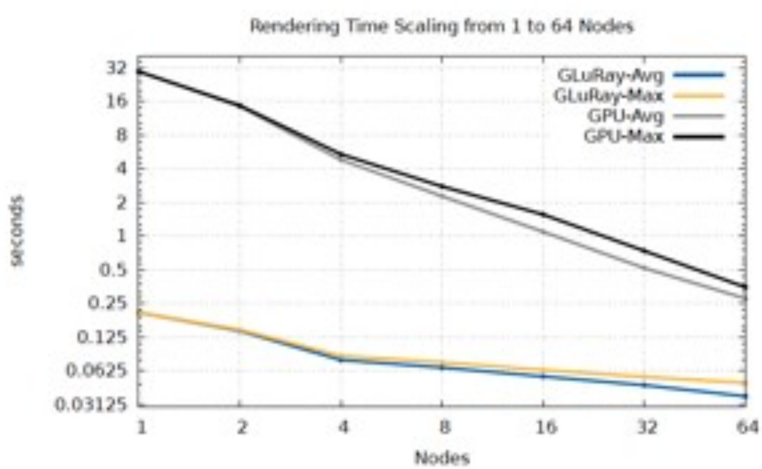
Current Visualization Software

OpenGL

VisIt ENSIGHT ParaView VAPOR



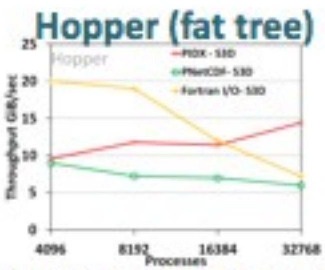
GLuRay

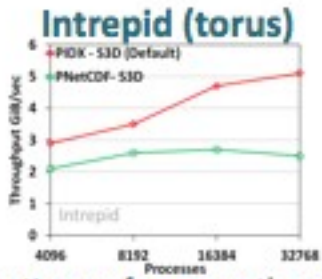


Michelangelo's David

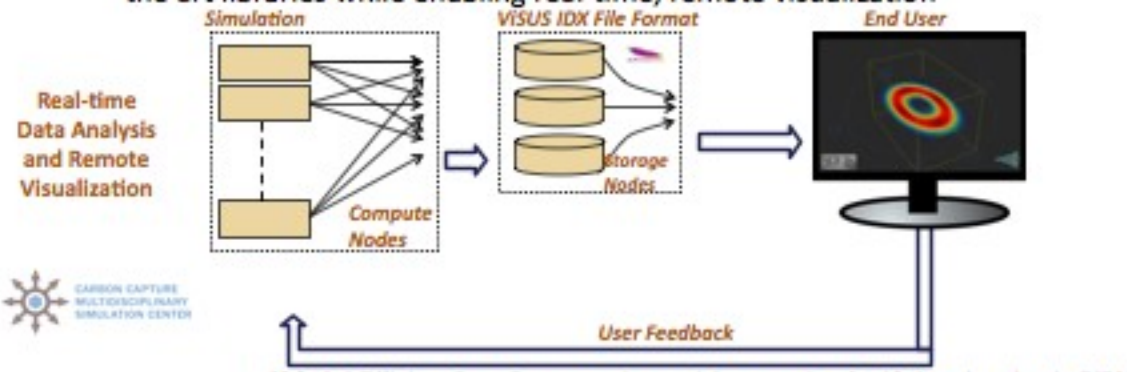
Exascale Challenges: Data Movement

High Performance Data Movements for Real-Time Monitoring of Large Scale Simulations





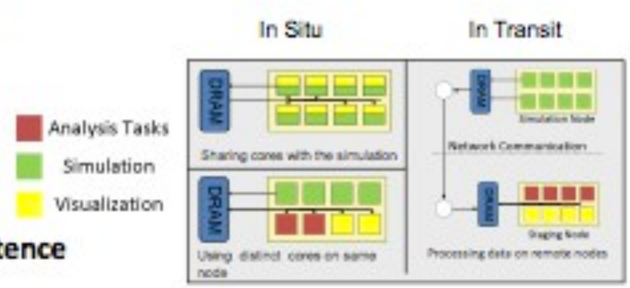
Scale simulation dumps to 130K cores with better performance than state of the art libraries while enabling real-time, remote visualization

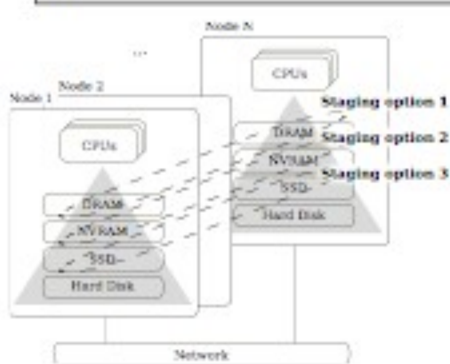


[SC12a] **Efficient Data Restructuring and Aggregation for IO Acceleration in PIDX**

Flexible DMAV Architecture that Allows Exploiting the "Possible" Exascale Hardware Available

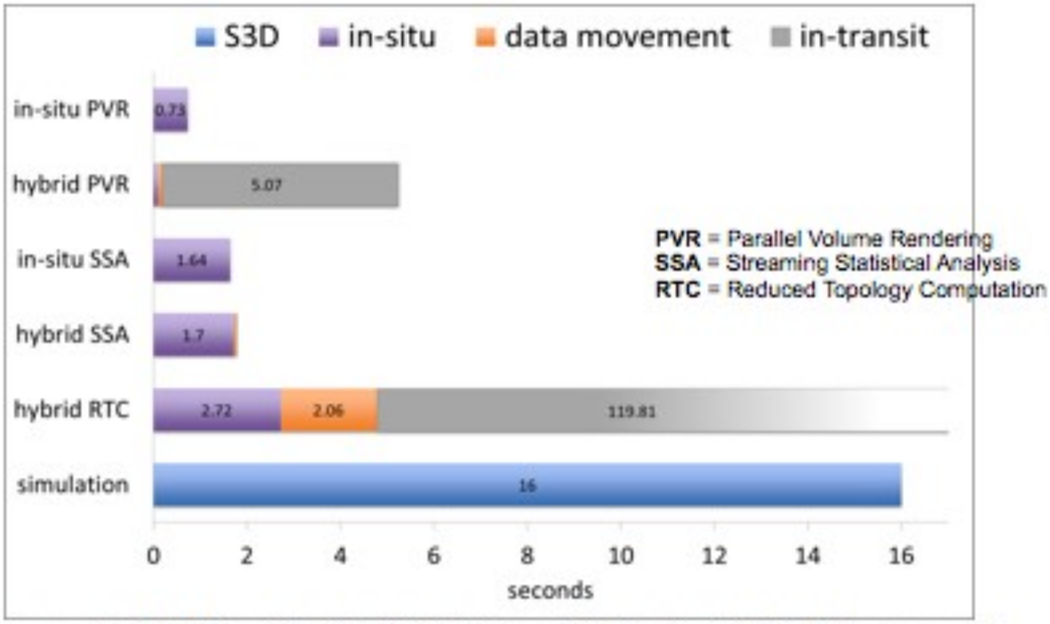
- **Location of the compute resources**
  - Same cores as the simulation
  - Dedicated cores
  - Dedicated nodes
  - External resource
- **Data access, placement, and persistence**
  - Shared data structures
  - Shared memory
  - Non-volatile on node storage (NVRAM)
  - Network transfer
- **Synchronization and scheduling**
  - Execute synchronously with simulation every n^{th} simulation time step
  - Execute asynchronously





Exploring algorithm design and task allocation

in-situ+in-transit workflows enable matching algorithms with architectures



- 4896 cores total (4480 simulation/in situ; 256 in transit; 160 task scheduling/data movement)
- Simulation size: 1600x1372x430 ; All measurements are per simulation time step

[SC12a] **Combining In-Situ and In-Transit Processing to Enable Extreme-Scale Scientific Analysis**

Exascale Challenges: Data Movement

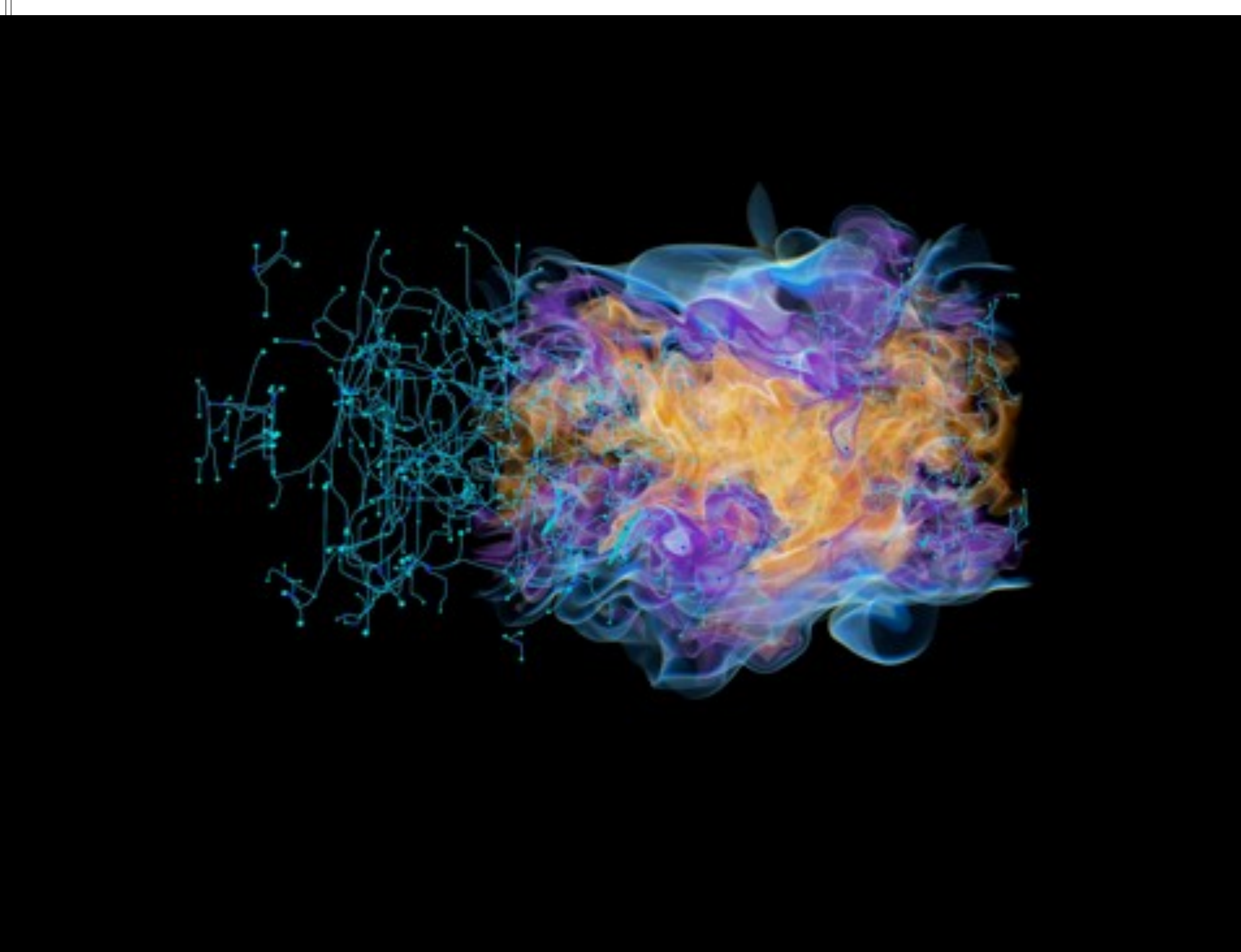
In Situ or In Transit

For Either:

Visualization is really just interpolation and sampling (think higher order)

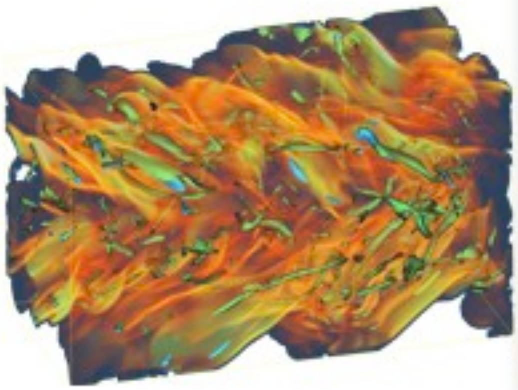
Why not use the data structures/ algorithms applications use?

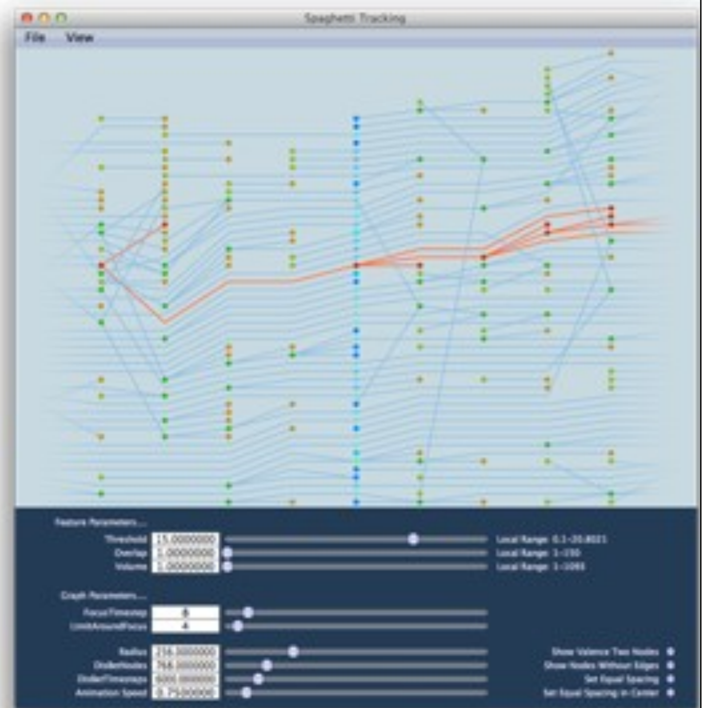
Exascale Challenges: Analysis



Topological Analysis of Massive Combustion Simulations

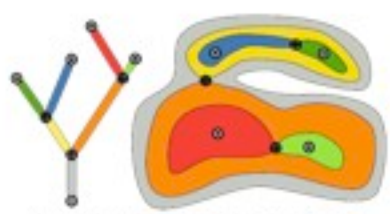
- Non-premixed DNS combustion (J. Chen, SNL): Analysis of the time evolution of extinction and reignition regions for the design of better fuels

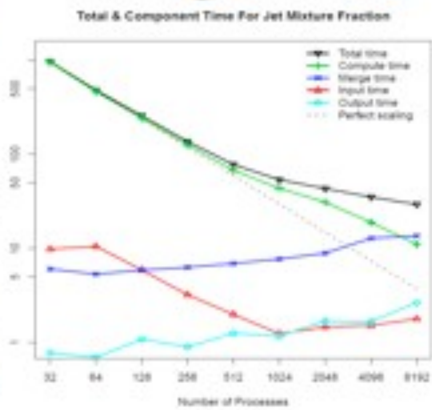


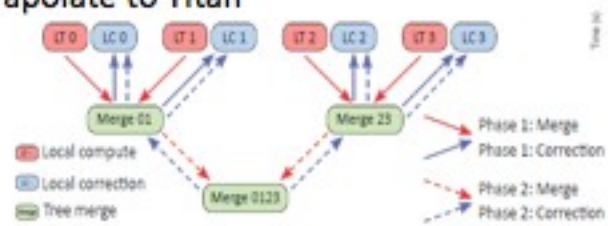


We Evaluate the Feasibility of Exascale *In-Situ* Feature Extraction and Tracking and its Impact on Power Consumption

- Analysis algorithms vs simulation:
  - different communication/instruction profile/memory access patterns
  - different communication patterns
- Develop power models with machine independent characteristics (Byfl and MPI)
- Validate power model using empirical studies with instrumented platforms
- Extrapolate to Titan



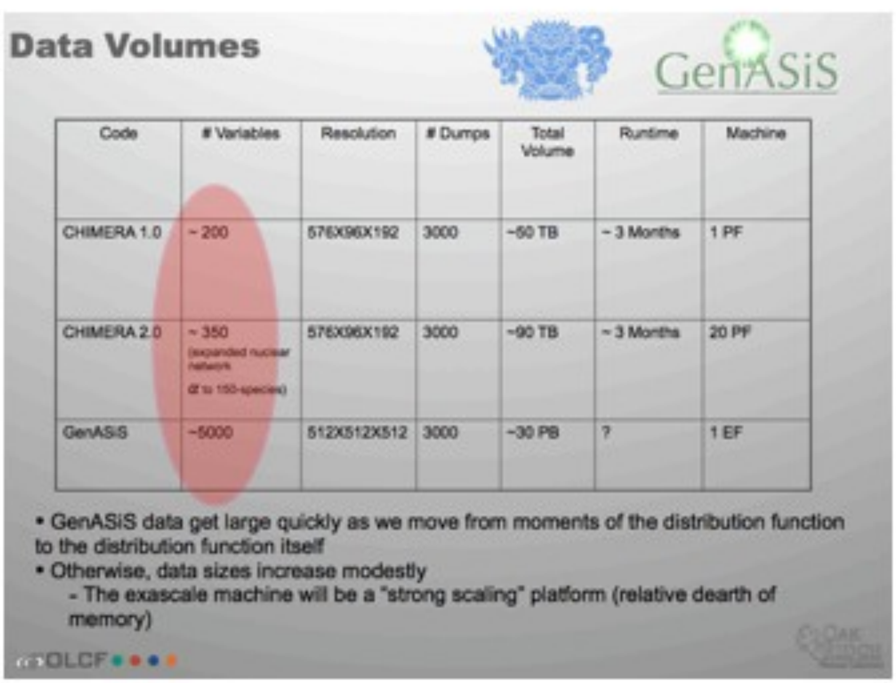




CARBON CAPTURE MULTIDISCIPLINARY SIMULATION CENTER

[SC13b] **Exploring Power Behaviors and Trade-offs of In-situ Data Analytics**

Exascale Challenges: Multifield / Multivariate



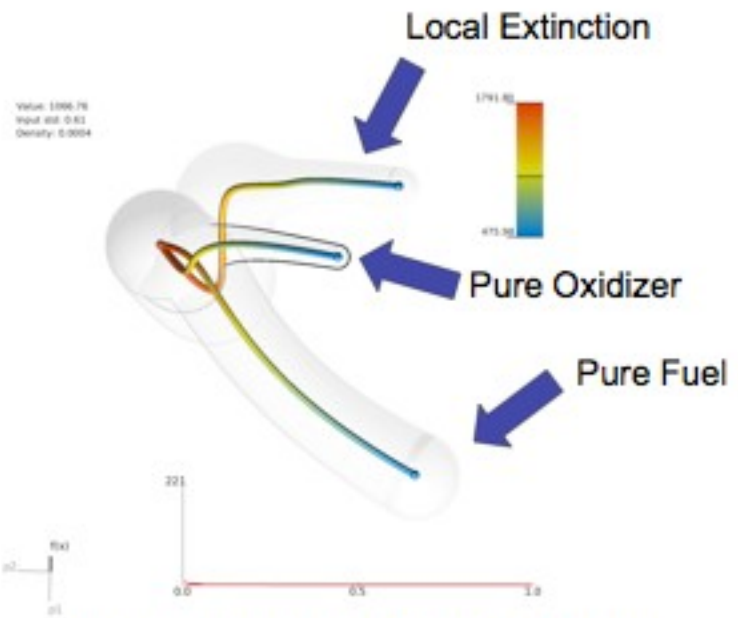
Data Volumes

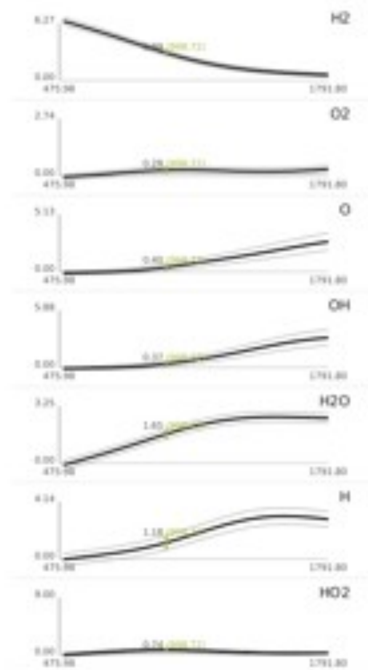
Code	# Variables	Resolution	# Dumps	Total Volume	Runtime	Machine
CHIMERA 1.0	~ 200	576X96X192	3000	~50 TB	~ 3 Months	1 PF
CHIMERA 2.0	~ 350 (expanded nuclear network 𝜃 to 150-species)	576X96X192	3000	~90 TB	~ 3 Months	20 PF
GenASiS	~5000	512X512X512	3000	~30 PB	?	1 EF

- GenASiS data get large quickly as we move from moments of the distribution function to the distribution function itself
- Otherwise, data sizes increase modestly
  - The exascale machine will be a "strong scaling" platform (relative dearth of memory)

Bronson Messer

Visualization of 10D Combustion Simulation of Jet CO/H2-Air Flames





10 dimensional data set describing the heat release wrt. to various chemical species in a combustion simulation

Analysis of Combustion Simulations

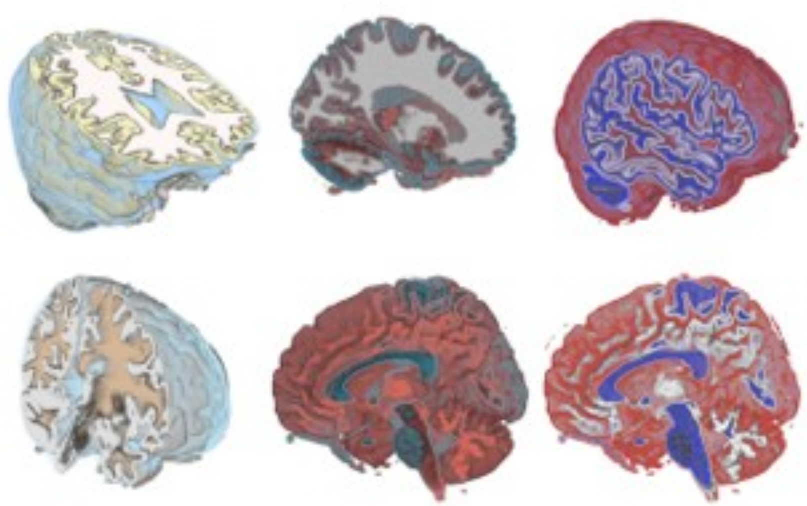
Combustion Simulation of Jet CO/H2-Air Flames

Input: Composition of 10 chemical species

Output: Temperature

Exascale Challenges: Uncertainty

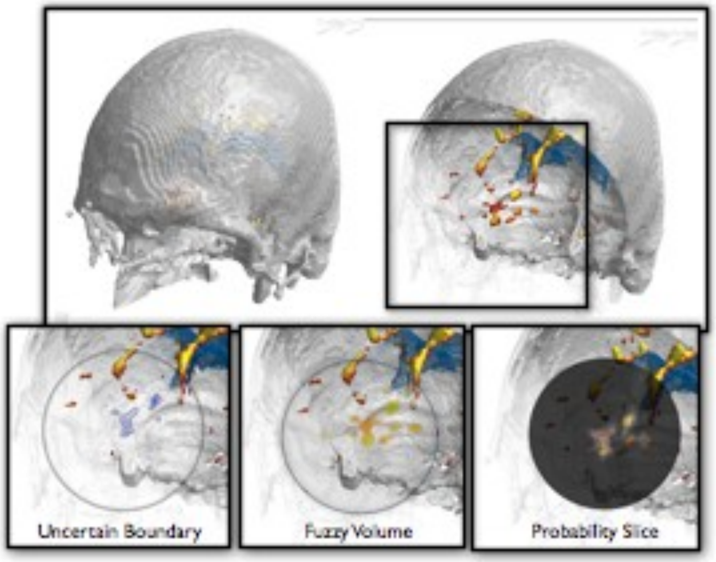
Visualizing Uncertainty



Fuzzy Sensitivity Confidence

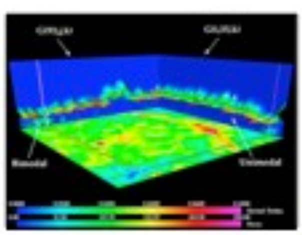
QuizLens: A Multi-lens approach for uncertainty exploration

- Global information important for qualitative evaluation & context
- Local information necessary for quantitative understanding
- Interchangeable lenses to explore various data characteristics

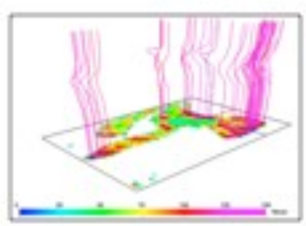


Ensembles

- **Multi-run/model simulations**
- **Distribution of data at every point**
- **Mean/std dev may not be appropriate**

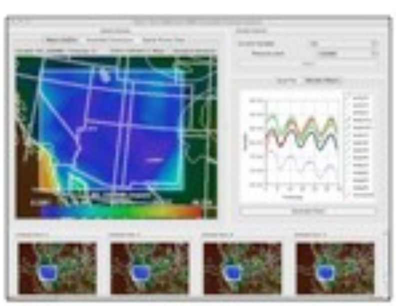


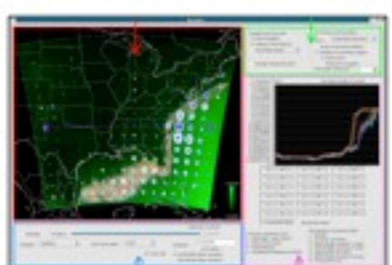
D. Kao, A. Luo, J. Dungan, A. Pang.
Visualizing Spatially Varying Distribution Data.
In Proc Information Visualization, 2002.



D. Kao, M. Kramer, A. Luo, J. Dungan, A. Pang.
Visualizing Distributions from Multi-Return Lidar Data to Understand Forest Structure.
In *The Cartographic Journal*, 42(1), 2005.

K. Potter, et al.
Ensemble-Vis: A Framework for the Statistical Visualization of Ensemble Data.
In *IEEE ICDM Workshop on Knowledge Discovery from Climate Data: Prediction*, 2009.





J. Sanyal, S. Zhang, J. Dyer, A. Mercer, P. Amburn.
Noodles: A Tool for Visualization of Numerical Weather Model Ensemble Uncertainty
In Proc *IEEE Vis*, 2010.

Visualization 2020

Scalable methods

Data analytics

Multivariate techniques

Minimize data movement

Power aware algorithms

Uncertainty visualization

Visualization 2020

E

X C

I T I

N G T I

M E S L O

T S O F R E

S E A R C H .

www.eyechartmaker.com

Visualization 2020

Exciting Times - Lots of Research