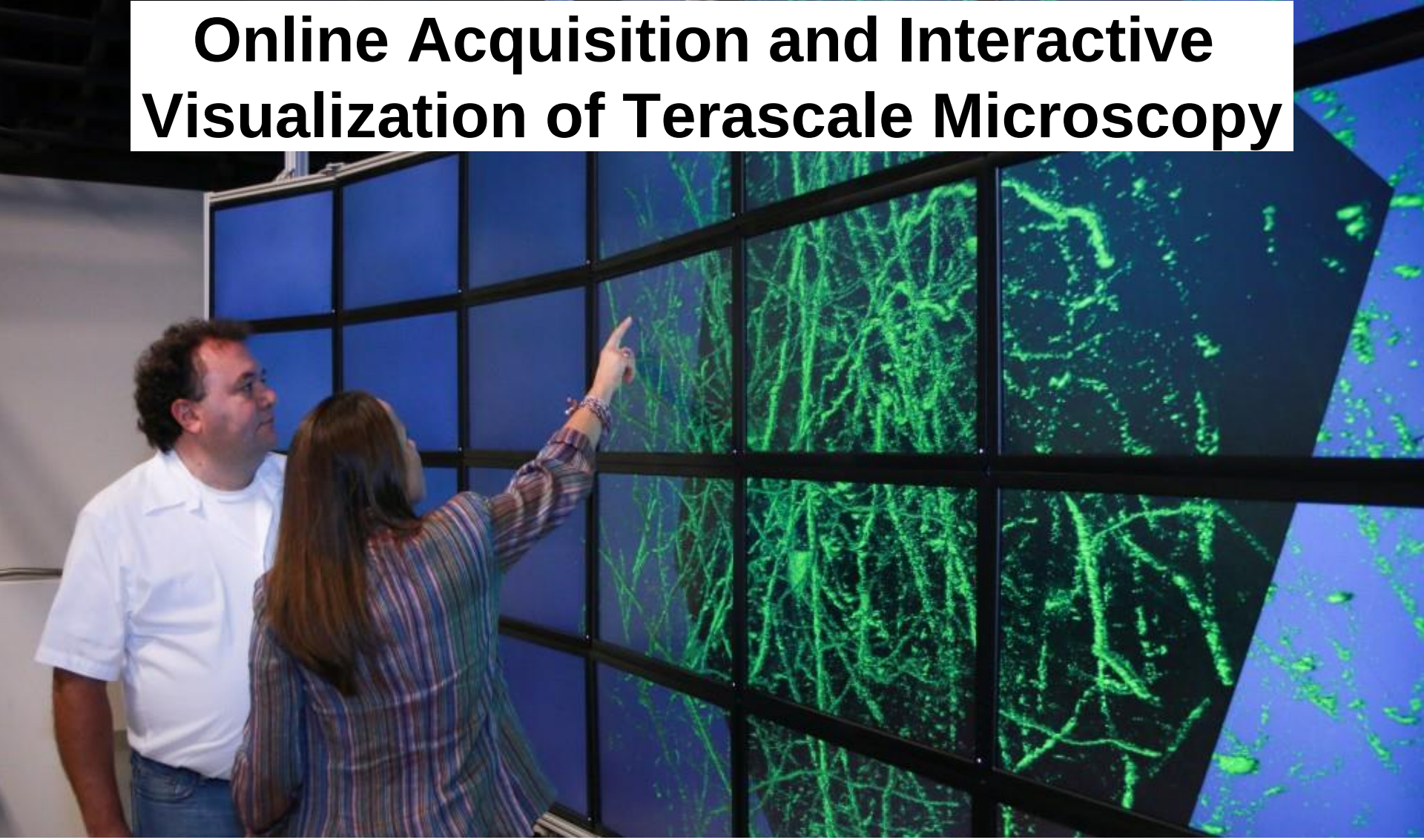
Online Acquisition and Interactive Visualization of Terascale Microscopy



Valerio Pascucci

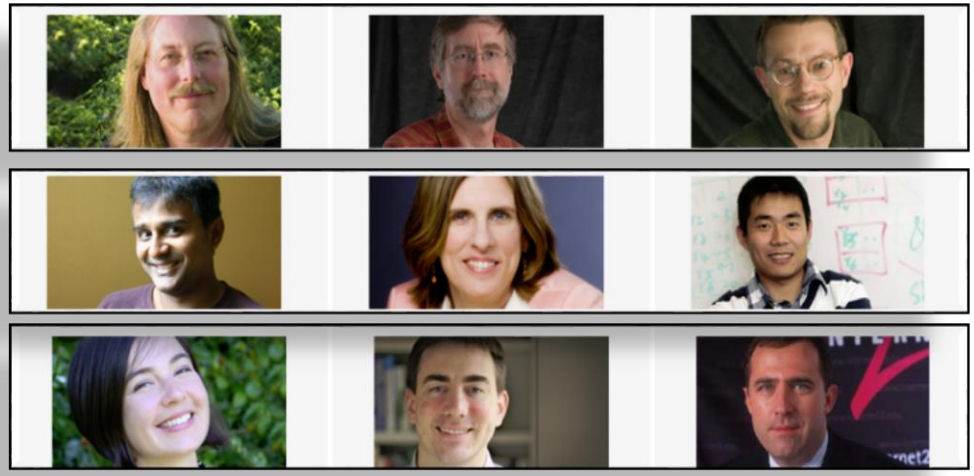
Director, Center for Extreme Data Management Analysis and Visualization

Professor, SCI institute and School of Computing, University of Utah

Laboratory Fellow, PNNL, CEO, Visus LLC

Center for Extreme Data Management, Analysis, and Visualization

- 10 Faculty + scientists, developers, students, …
- Primary partners: UU & PNNL
- Other partnerships: NSA, INL, LLNL, ANL, Battelle, ….
- Involvement in national Initiatives



SDAV Scalable Data Management, Analysis, and Visualization

EXACT

SUPER

The University of Utah Pacific Northwest National Laboratory

Lawrence Livermore National Laboratory

Sandia National Laboratories

INL Idaho National Laboratory

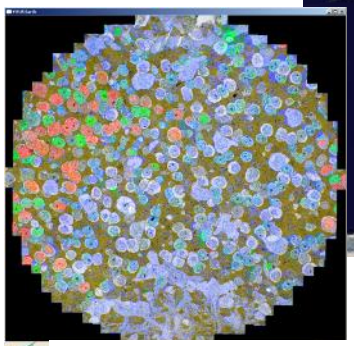
National Security Agency

$1.6B NSA data center
(1.5 million-square-foot facility)

Massive Simulation and Sensing Devices Generate Great Challenges and Opportunities

BlueGene/Q

Satellite



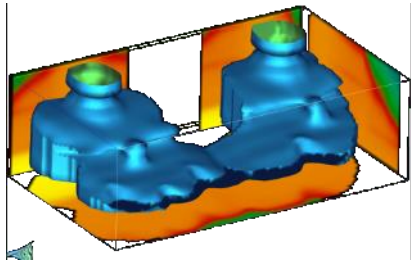
Earth Images

Retinal Connectome

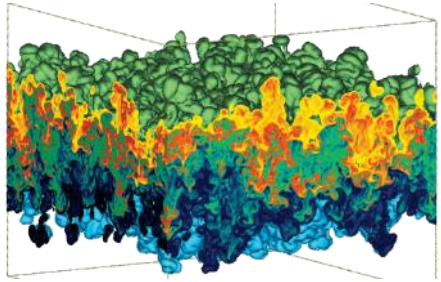
Cameras

EM

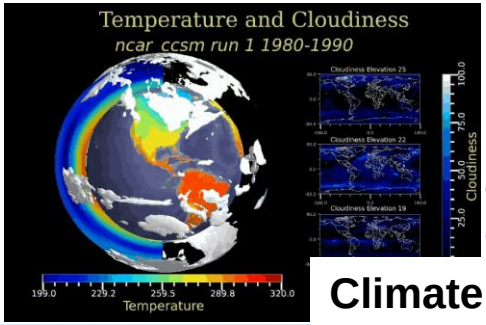
Jaguar



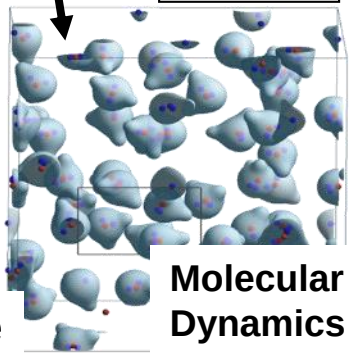
Carbon Seq. (Subsurface)



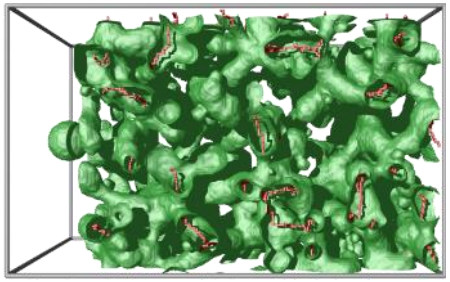
Hydrodynamic Inst.



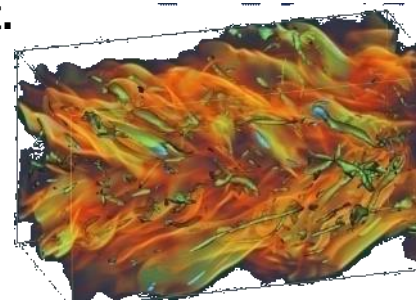
Climate



Molecular Dynamics



Porous Materials



Turbulent Combustion

Photography

Specific Goal in Neuroscience Applications

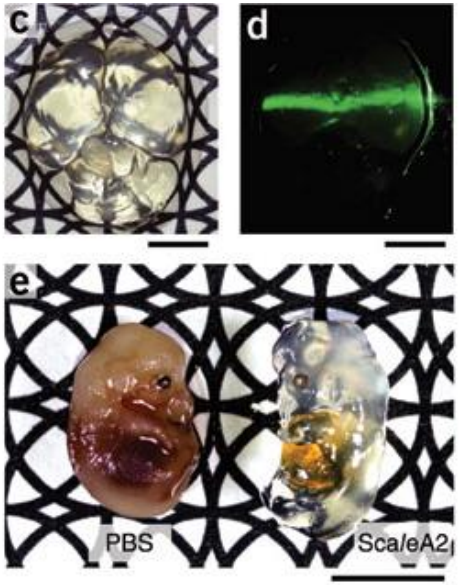
Image, reconstruct, and visualize labeled cortical circuits of the primate visual system in 3D for detailed anatomical (and functional) analysis.

Must scale to work with large primate brains.

Neuroscience Problems: Moran Eye Center

Computer Science Problems: CEDMAV

Clearing the Brain: Sca/e

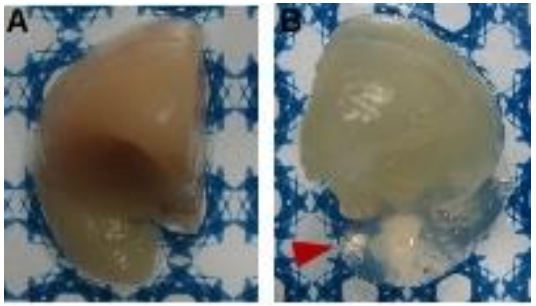


Hama et al., Nat Neuro 2011

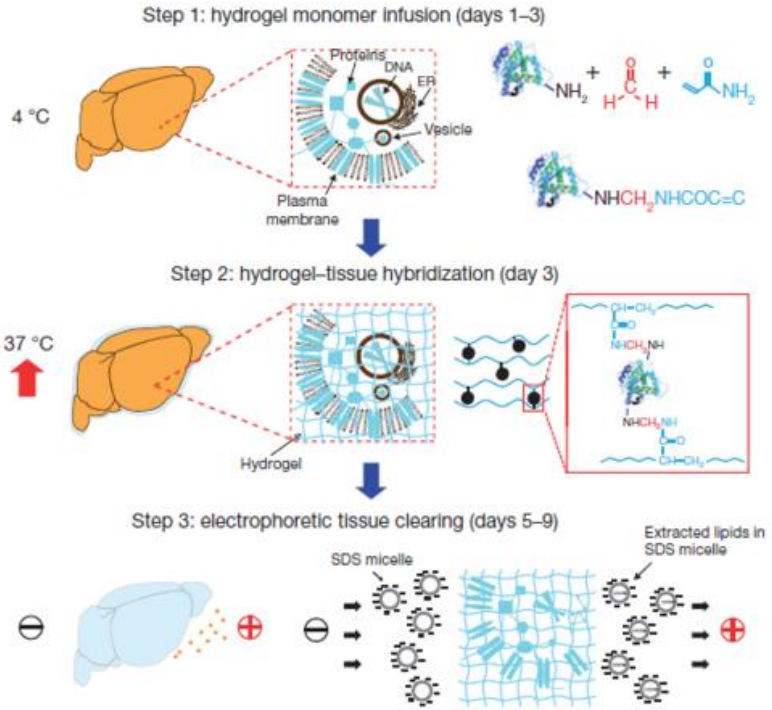
- 4 M Urea + 10% glycerol
- 2+ weeks to clear mouse brain

Disadvantages of Scale Clearing Method

- Large tissue expansion
- Takes a month or more to clear large primate tissue blocks
- Even when cleared…it wasn't that clear
- Can not do immune staining
- Difficult to image over long time courses



Clearing the Brain: CLARITY

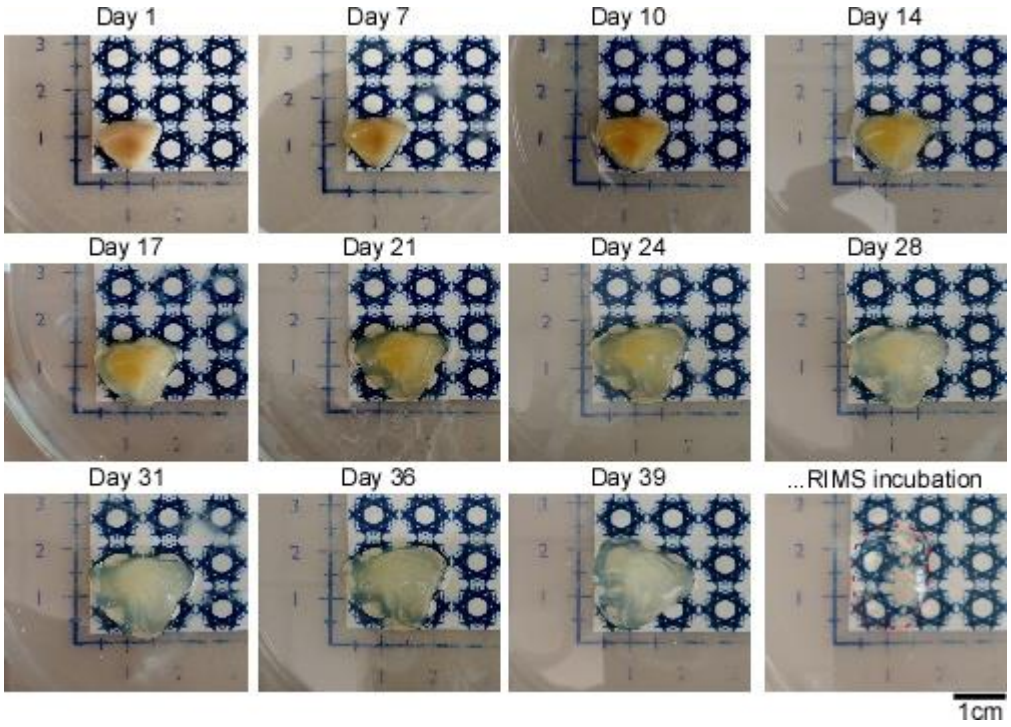


Chung et al., Nature 2013

Diesseroth Lab CLARITY Method Disadvantages

- Difficult initial set up and execution
- Expensive (as published)
- Tissue can be damaged in the electrophoresis process

PACT and RIMS on Primate Cortex



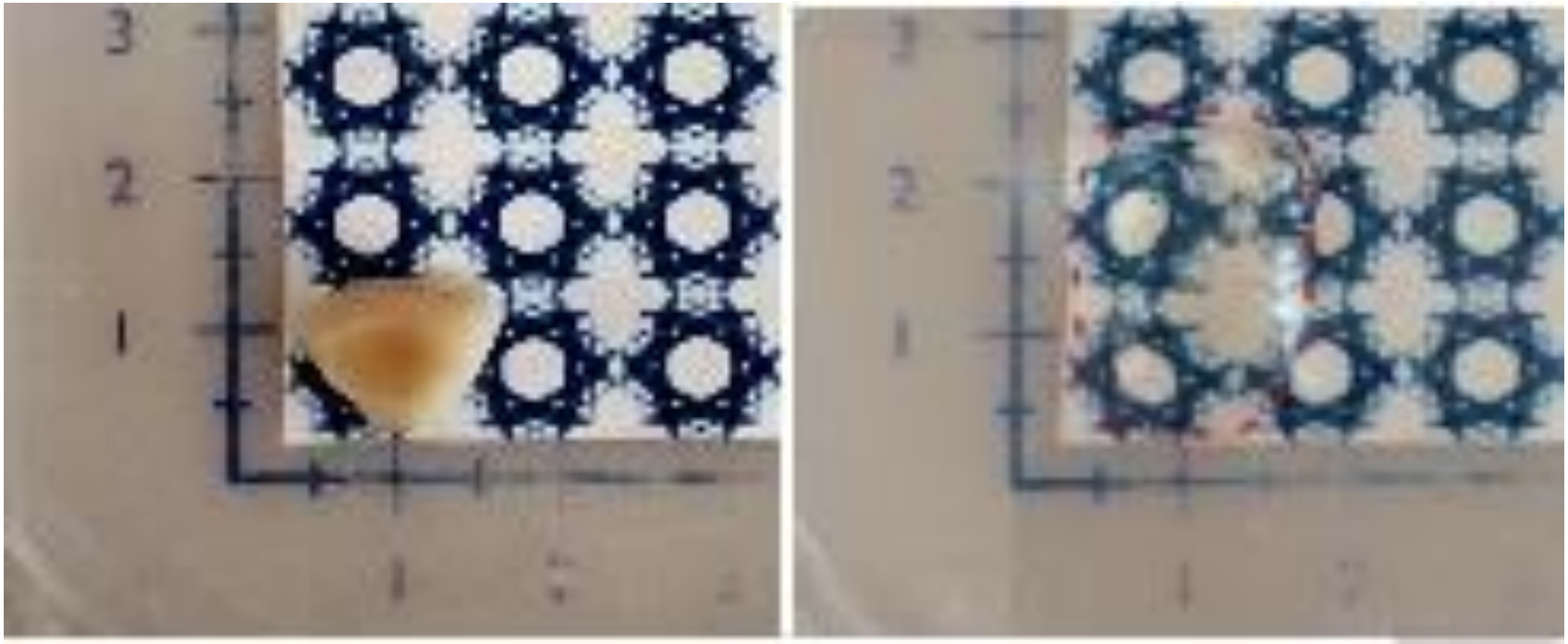
-Time consuming, but simple to execute and tissue unlikely to be damaged

-RIMS mounting medium produces the clearest tissue of any we have used (vs. Scale and FocusClear)

-Approx 15x15x15mm piece of cortex (3375 mm^3)

-Interested in imaging approx 5x3x2mm (30 mm^3)

PACT and RIMS on Primate Cortex2`



2-photon Microscope Generates 2kx2kx4k Images in one Acquisition

Imaging Volume Data Sizes

(Brain pictures: DeFilipe 2011)

30 TB

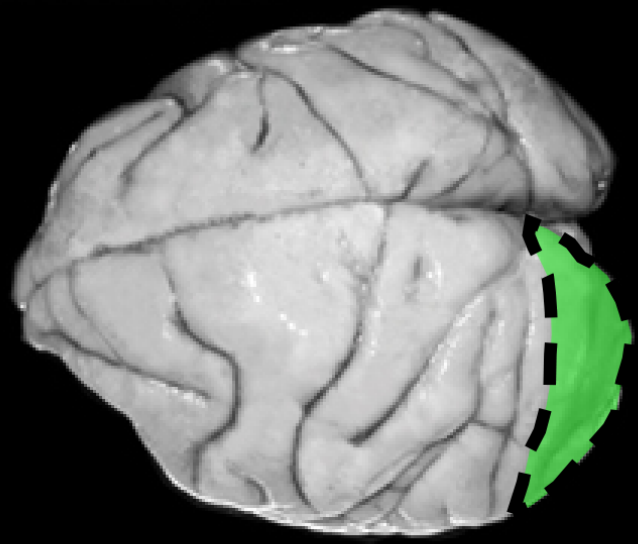
Penny-Sized Volume

5 cm

31 TB

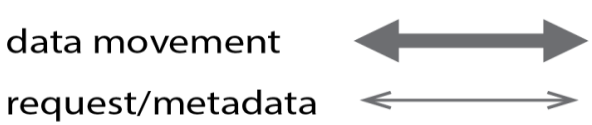
Whole Mouse Brain

318 TB



Macaque Primary Visual Cortex

On-the-fly Conversion



Client

Data Source

Image data

Transfer & Convert HTTP Request

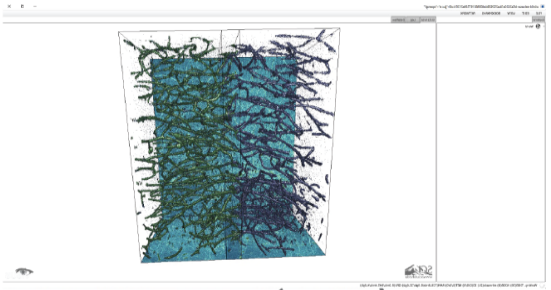
Create HTTP Request

ViSUS Web Service

Add IDX to ViSUS config

Create IDX Metadata

ViSUS Viewer



idx data

remote query

VISUS Server

idx data

local query

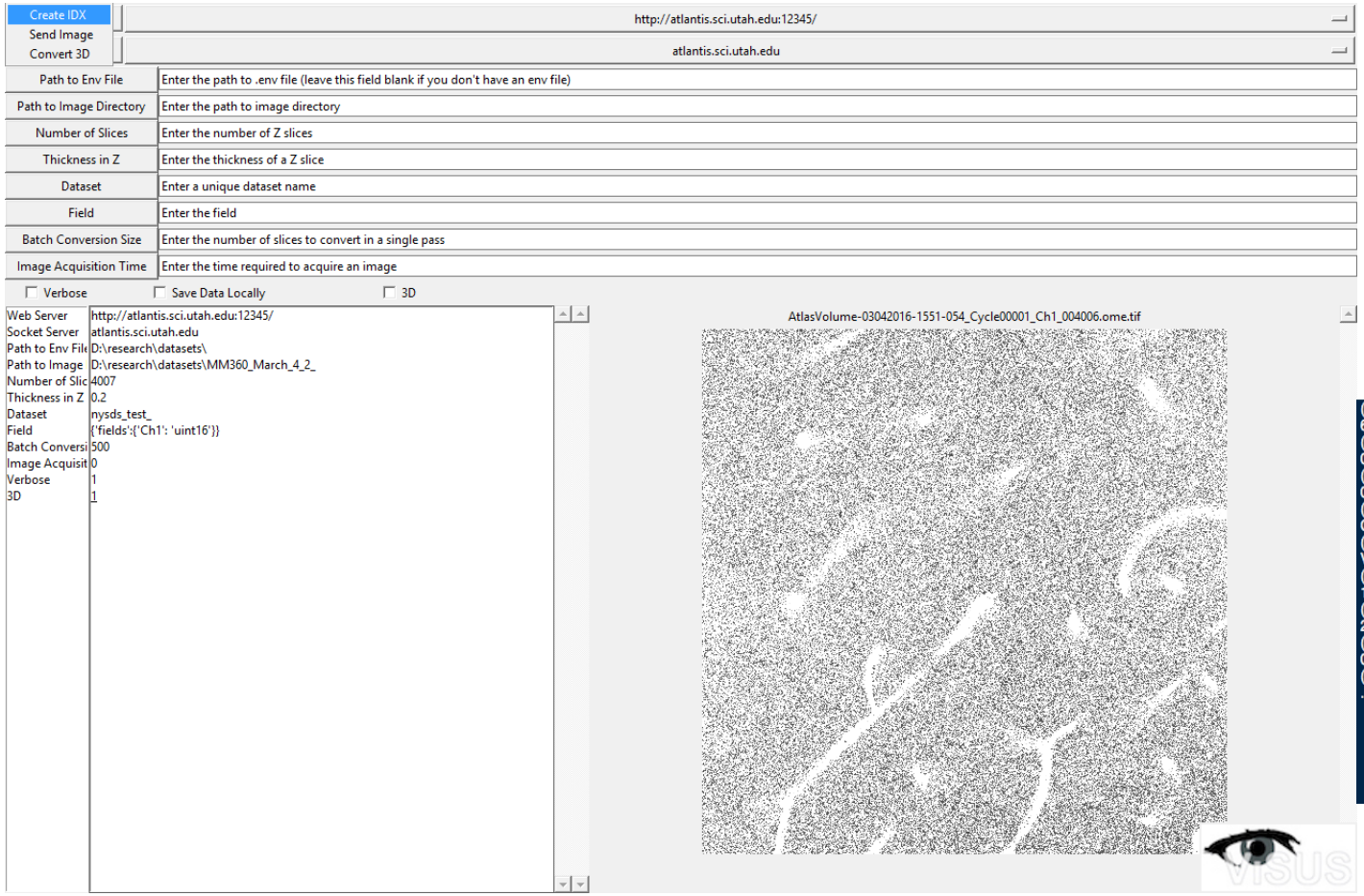
Database

Image data

Image To IDX Converter

idx data

Step 1: Create IDX Metadata

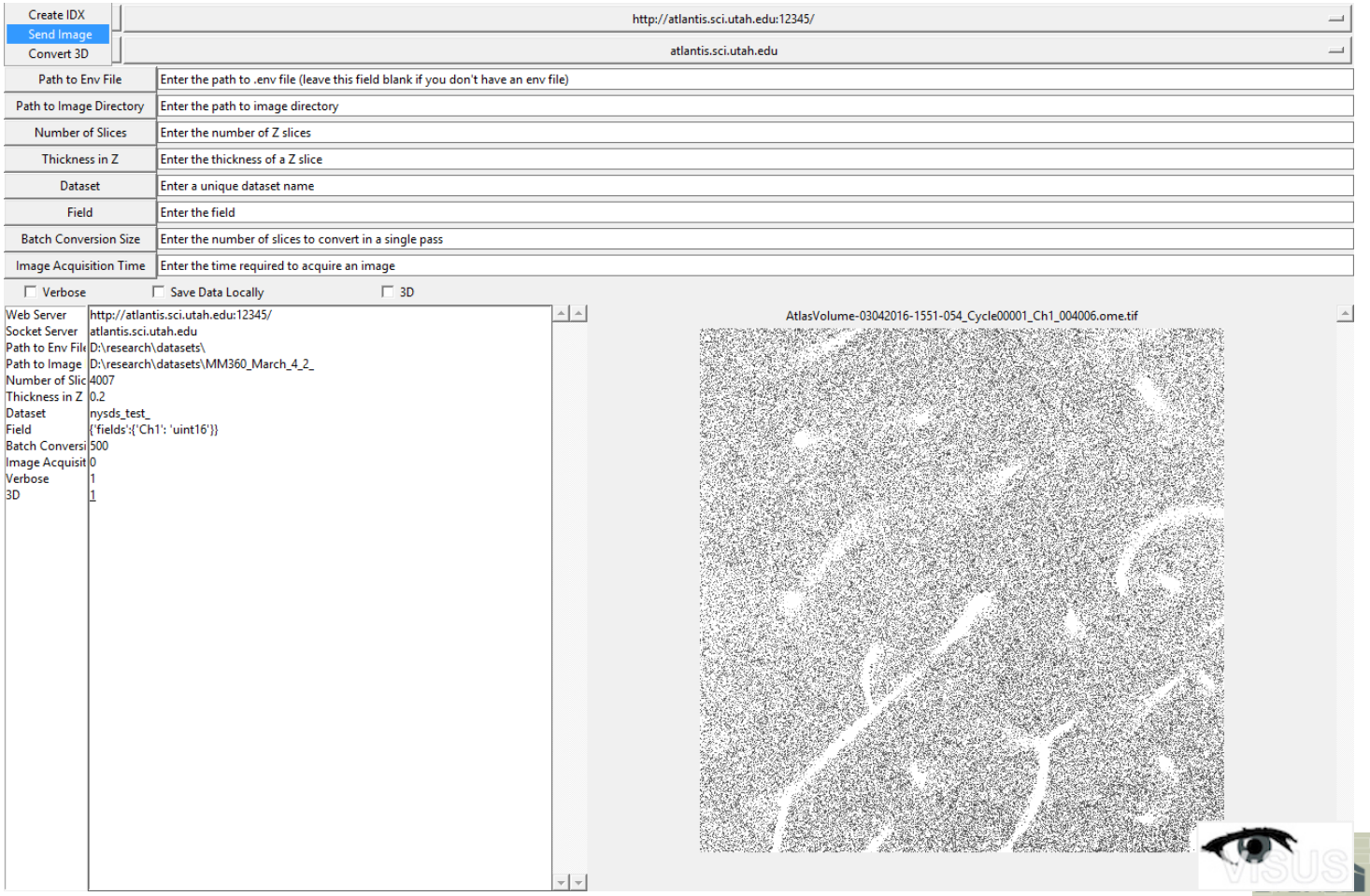


Creates IDX Metadata

```
(version)
6
(logic_to_physic)
0.267345 0 0 -397.64 0 0.267345 0 -145.3 0 0 0.2 0 0 0 0 1
(box)
0 2047 0 2047 0 4006 0 0 0 0
(fields)
Ch1 uint16 format(0)
(bits)
V2012012012012012012012012012012012
(bitsperblock)
16
(blocksperfile)
256
(interleave block)
0
(filename_template)
./MM360_March_4_2_00000/%02x/%04x.bin
```

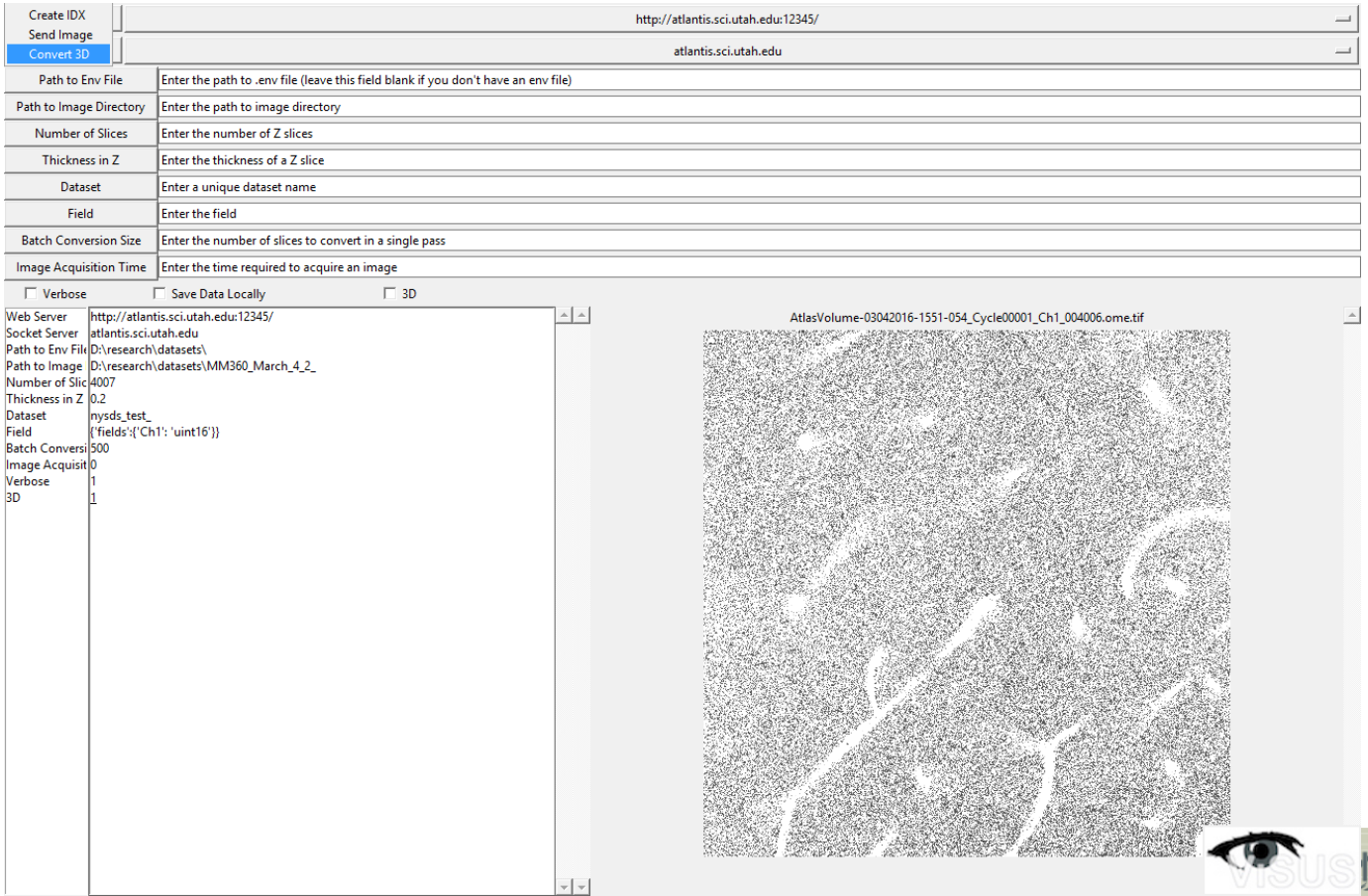
Step 2: Send Image

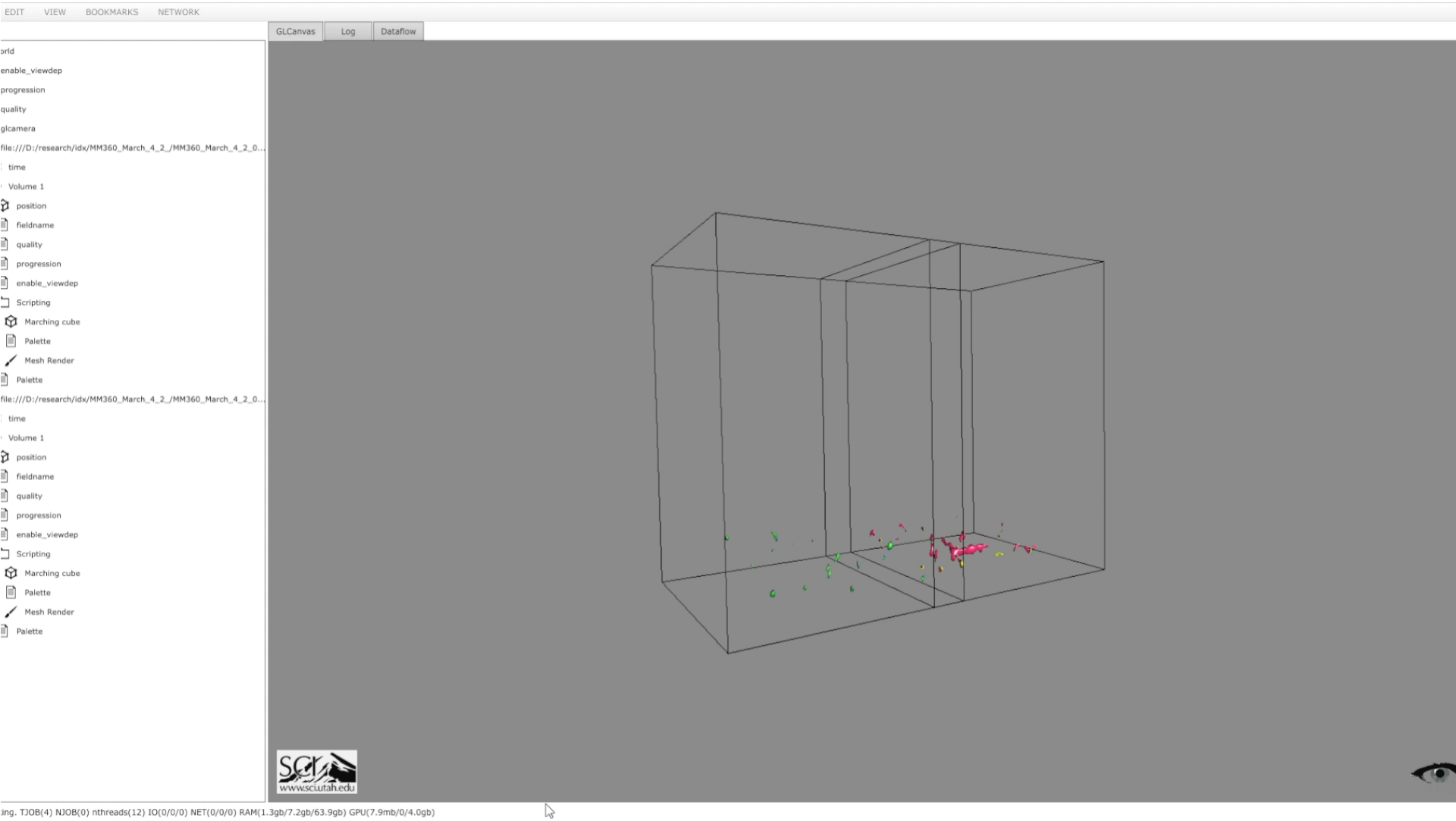
Transfers the acquired image data from the microscope to a remote server eliminating manual data offloading operations and expensive storage solutions at the microscope.



Step 3: Convert Data

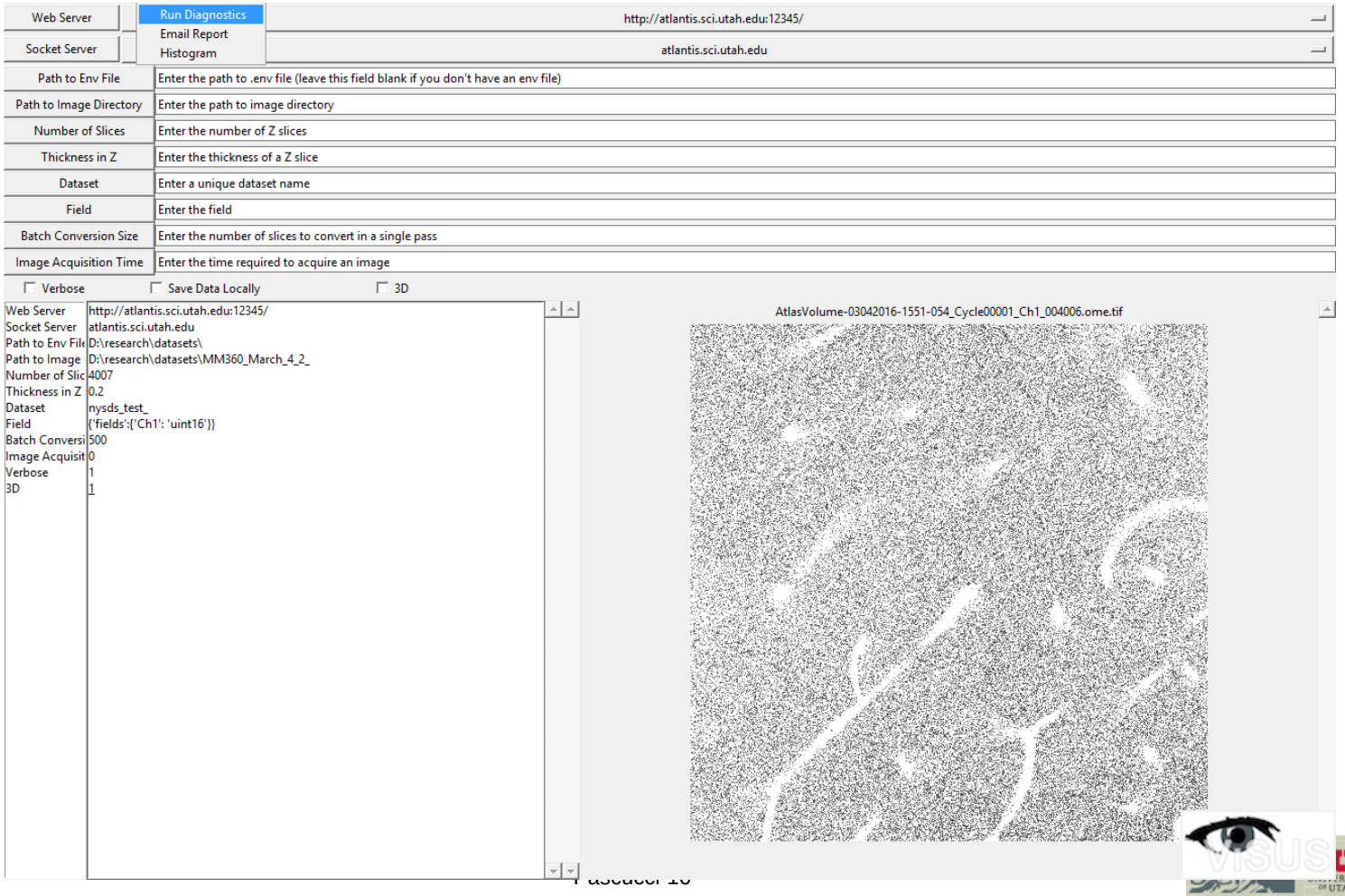
- **Converts raw data into multiresolution IDX streaming data format.**
- **Conversion process occurs as the data is being moved without creating processing delays.**
- **Converted data is accessible in real-time to ViSUS viewer client.**




EDIT
VIEW
BOOKMARKS
NETWORK
GLCanvas
Log
Dataflow
orld
enable_viewdep
progression
quality
glcamera
file:///D:/research/idx/MM360_March_4_2_/MM360_March_4_2_0...
time
Volume 1
position
fieldname
quality
progression
enable_viewdep
Scripting
Marching cube
Palette
Mesh Render
Palette
file:///D:/research/idx/MM360_March_4_2_/MM360_March_4_2_0...
time
Volume 1
position
fieldname
quality
progression
enable_viewdep
Scripting
Marching cube
Palette
Mesh Render
Palette
www.sci.utah.edu
ing. TJOB(4) NJOB(0) nthreads(12) IO(0/0/0) NET(0/0/0) RAM(1.3gb/7.2gb/63.9gb) GPU(7.9mb/0/4.0gb)

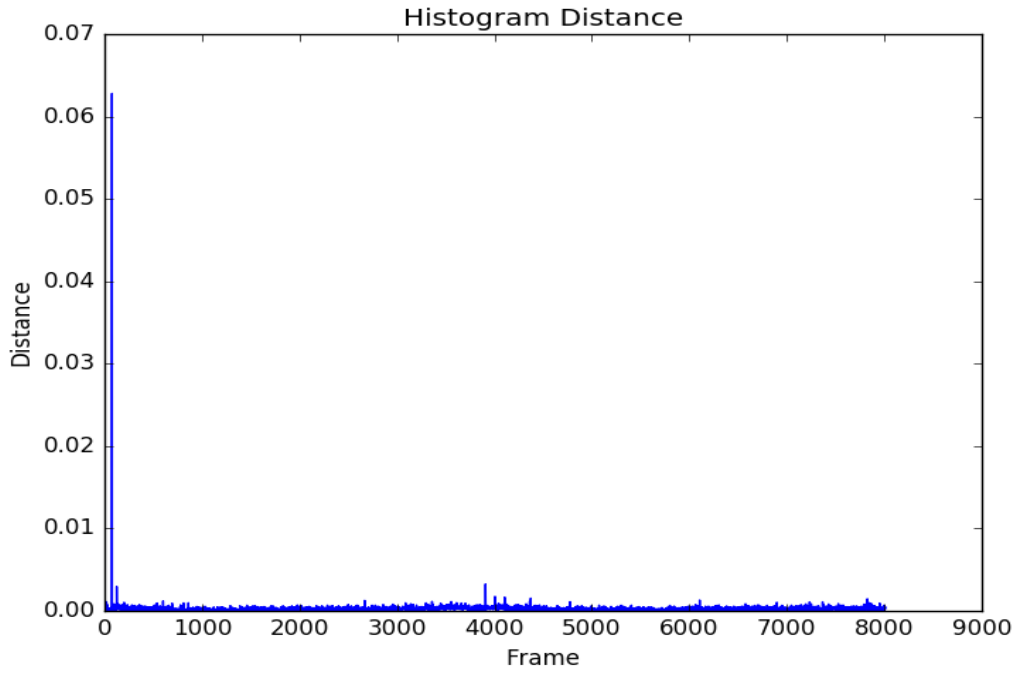
Step 4: Diagnostics

Run Diagnostics to automatically test for image intensity variations due to an increase in laser power or light-capture gain settings.

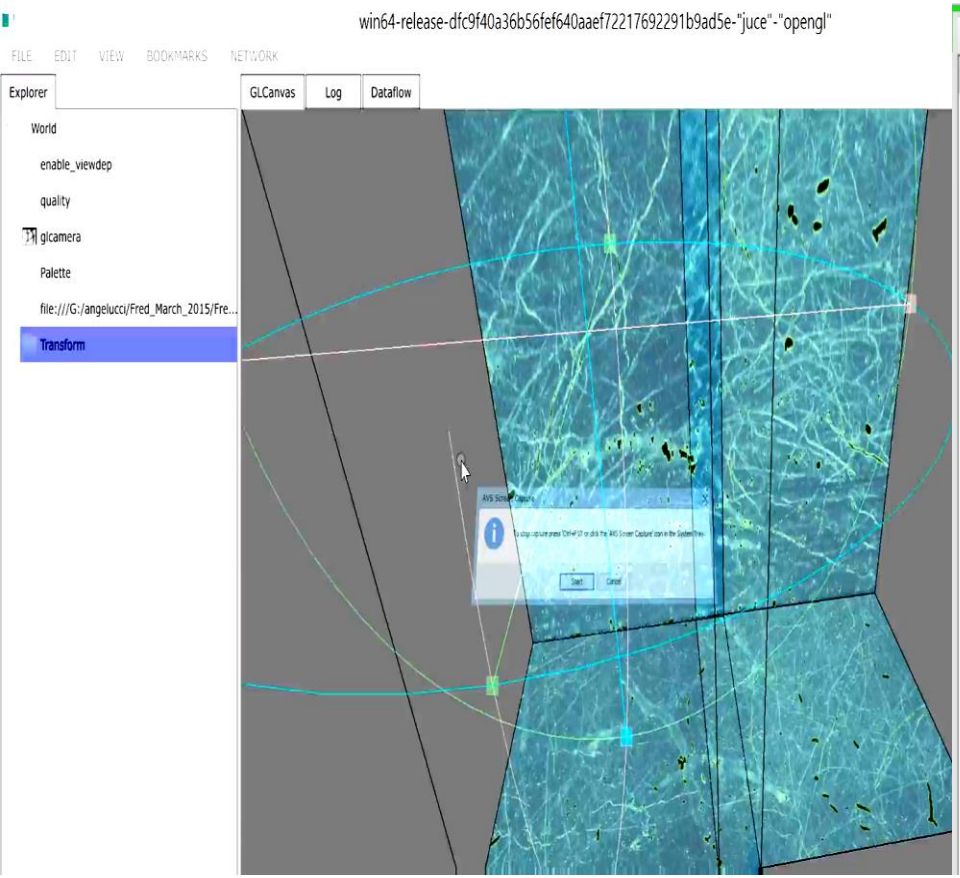


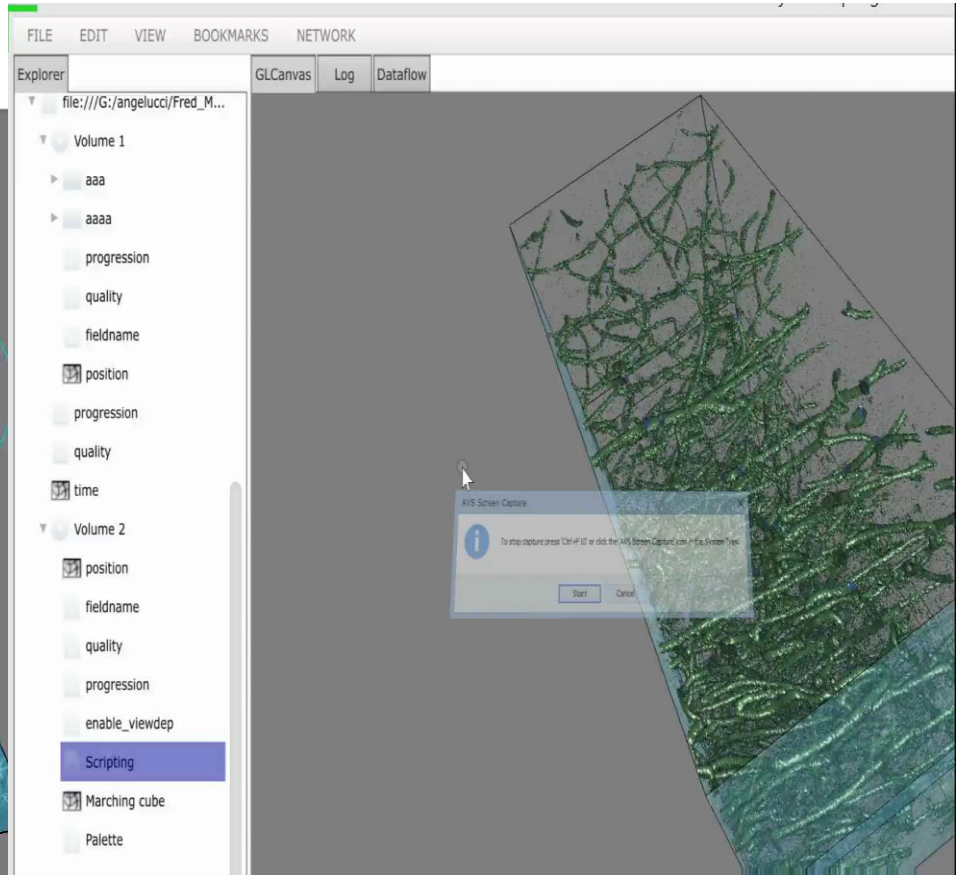
Step 5: Diagnostics Report

- **Diagnostics computes histogram distance between successive images to find variations in the image intensity.**
- **Plot shows the histogram distance for 8012 slices. A spike in the plot indicates that an image could be corrupt due to intensity variations.**
- **A thresholding on image distance is used to detect and report any spikes to the user (other diagnostics can be added).**

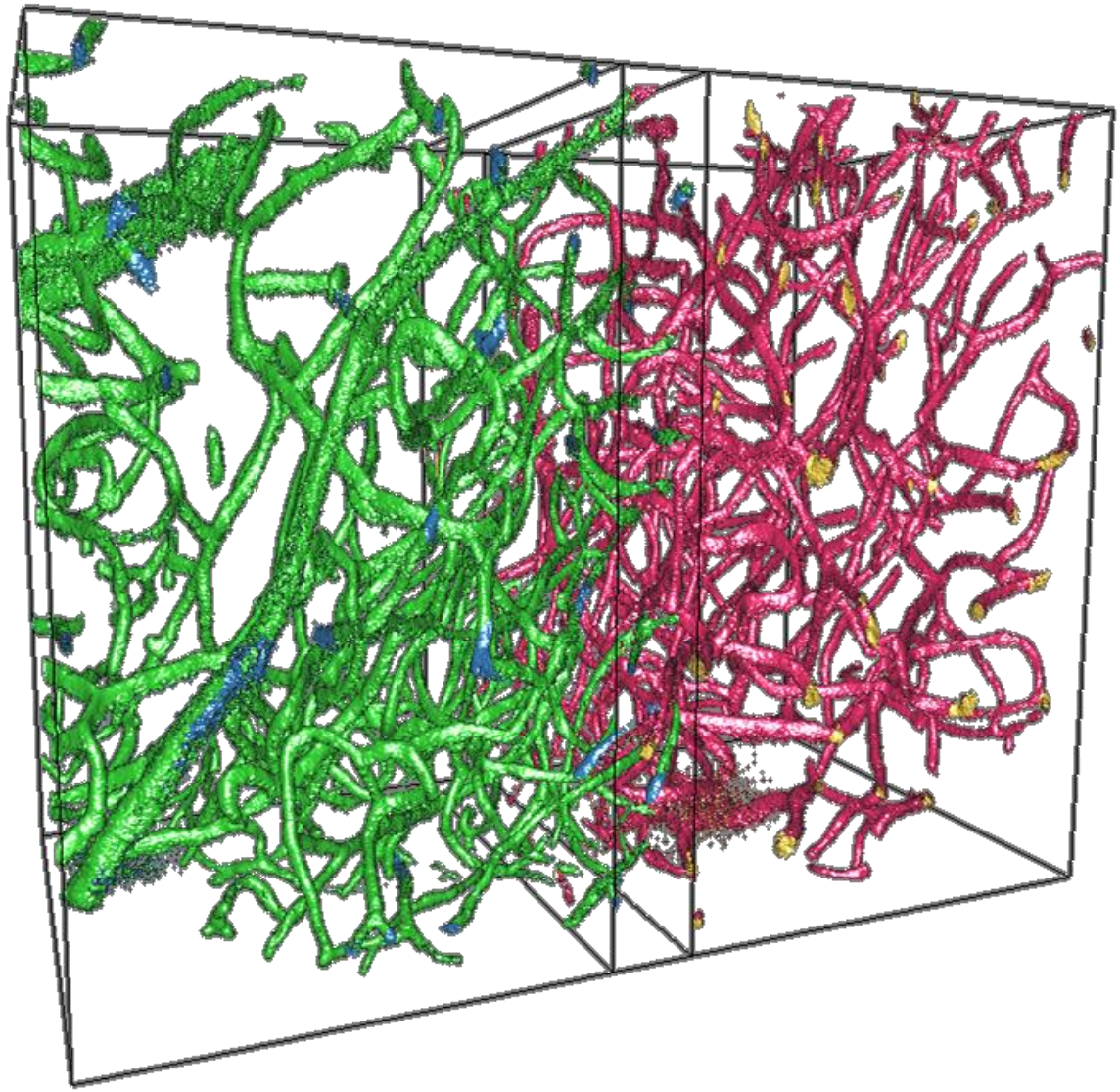


Mesoscopic Scale Anatomical Imaging and Visualization of Primate Cortical Circuits



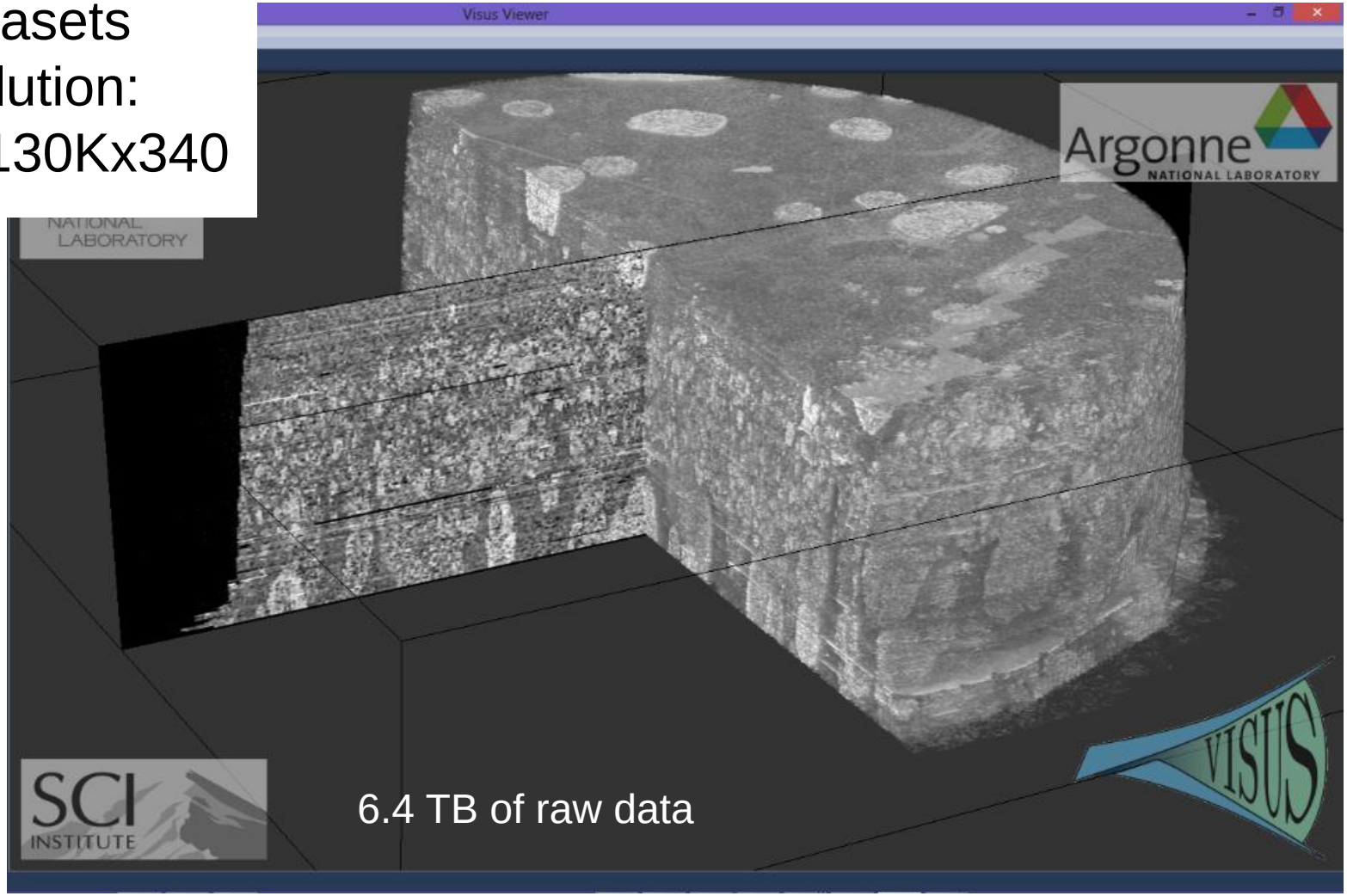


Live Demo



Scalability: Interactive Remote Analysis and Visualization of 6TB Imaging Data

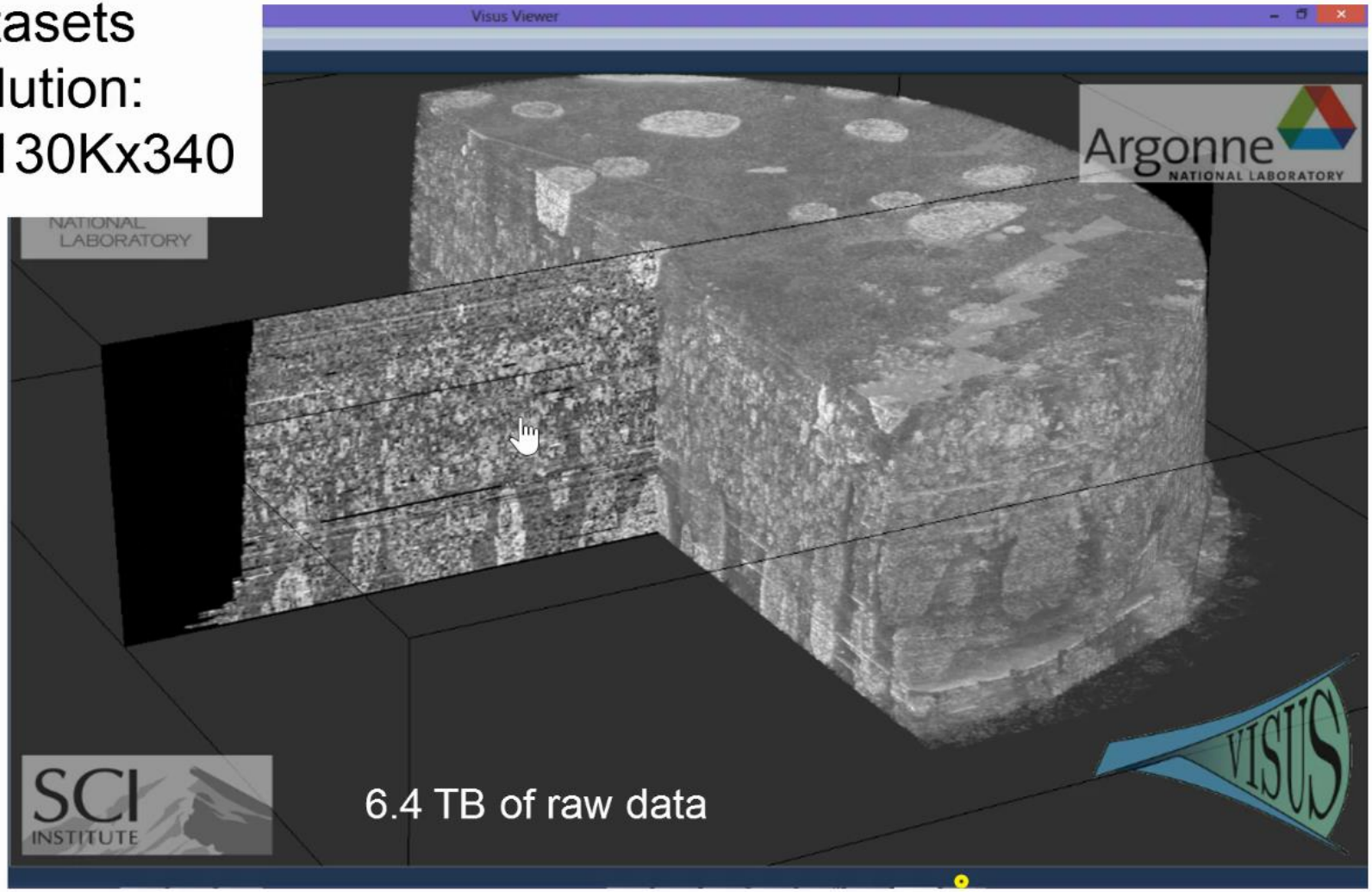
- EM datasets of resolution: 130Kx130Kx340



Web Server

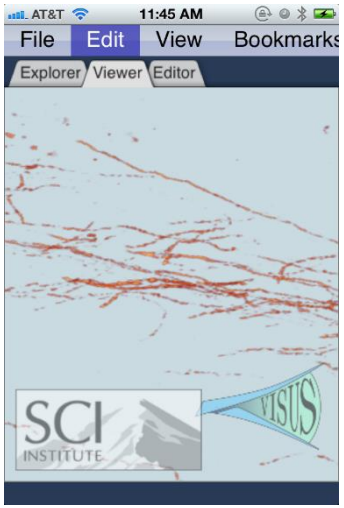
Demo: Interactive Remote Analysis and Visualization of 6TB Imaging Data

- EM datasets of resolution: 130Kx130Kx340

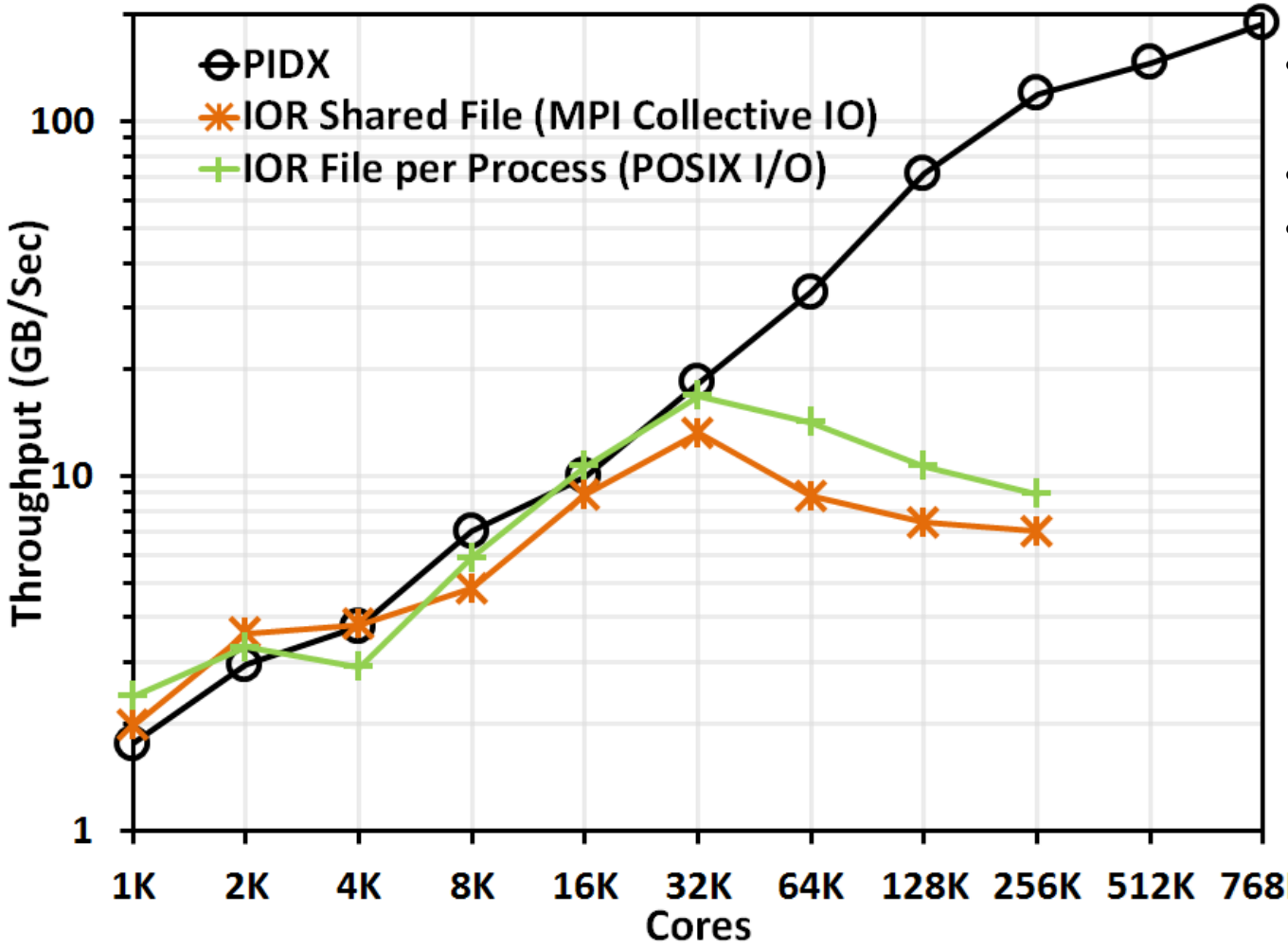


Web Server

Scalability: Output Device Can Range from a Handheld Device to a Powerwall Display

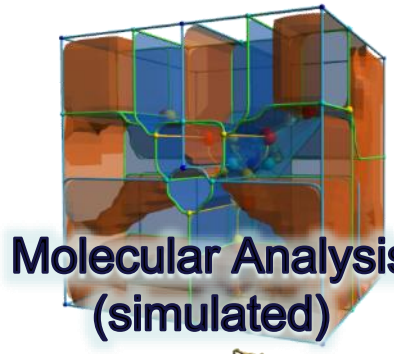


Scalability: Data Movements for Real-Time Monitoring of Large Scale Simulations

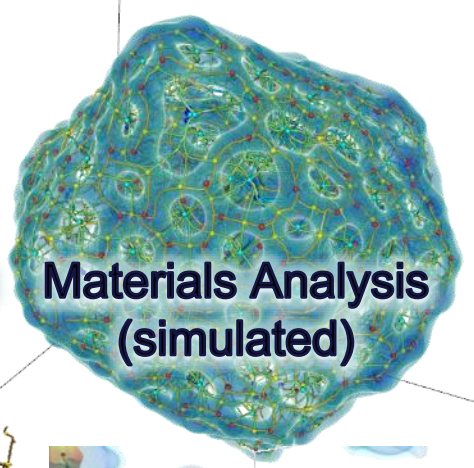


- 6 Terabytes per timestep
- 190 GB/sec
- 80% of max declared bandwidthq

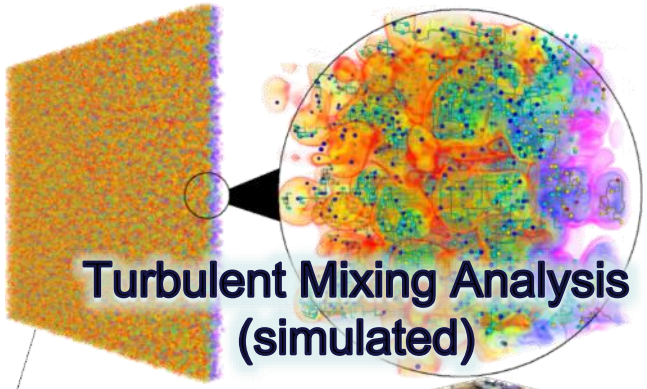
Topology Has Been Successful for Analysis and Visualization of Massive Scientific Data



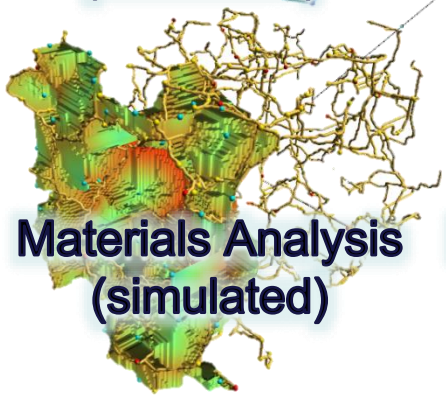
Molecular Analysis (simulated)



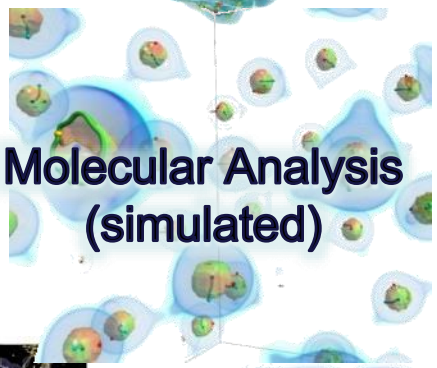
Materials Analysis (simulated)



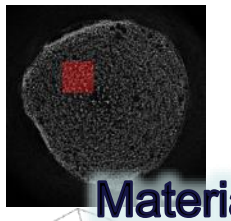
Turbulent Mixing Analysis (simulated)



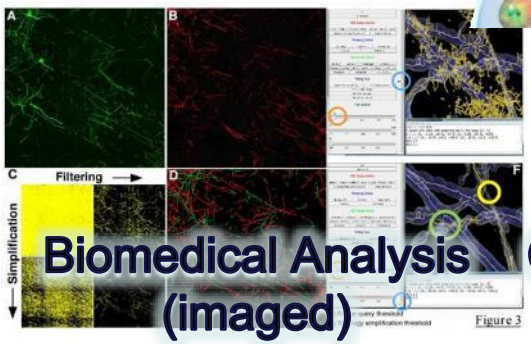
Materials Analysis (simulated)



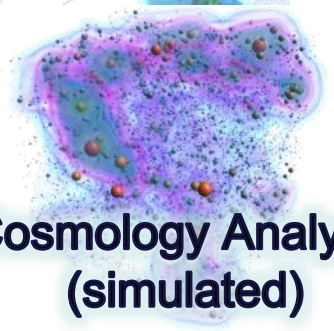
Molecular Analysis (simulated)



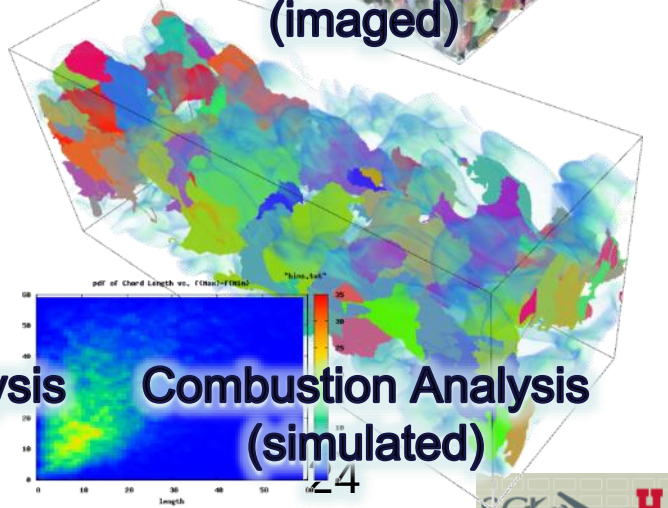
Materials Analysis (imaged)



Biomedical Analysis (imaged)

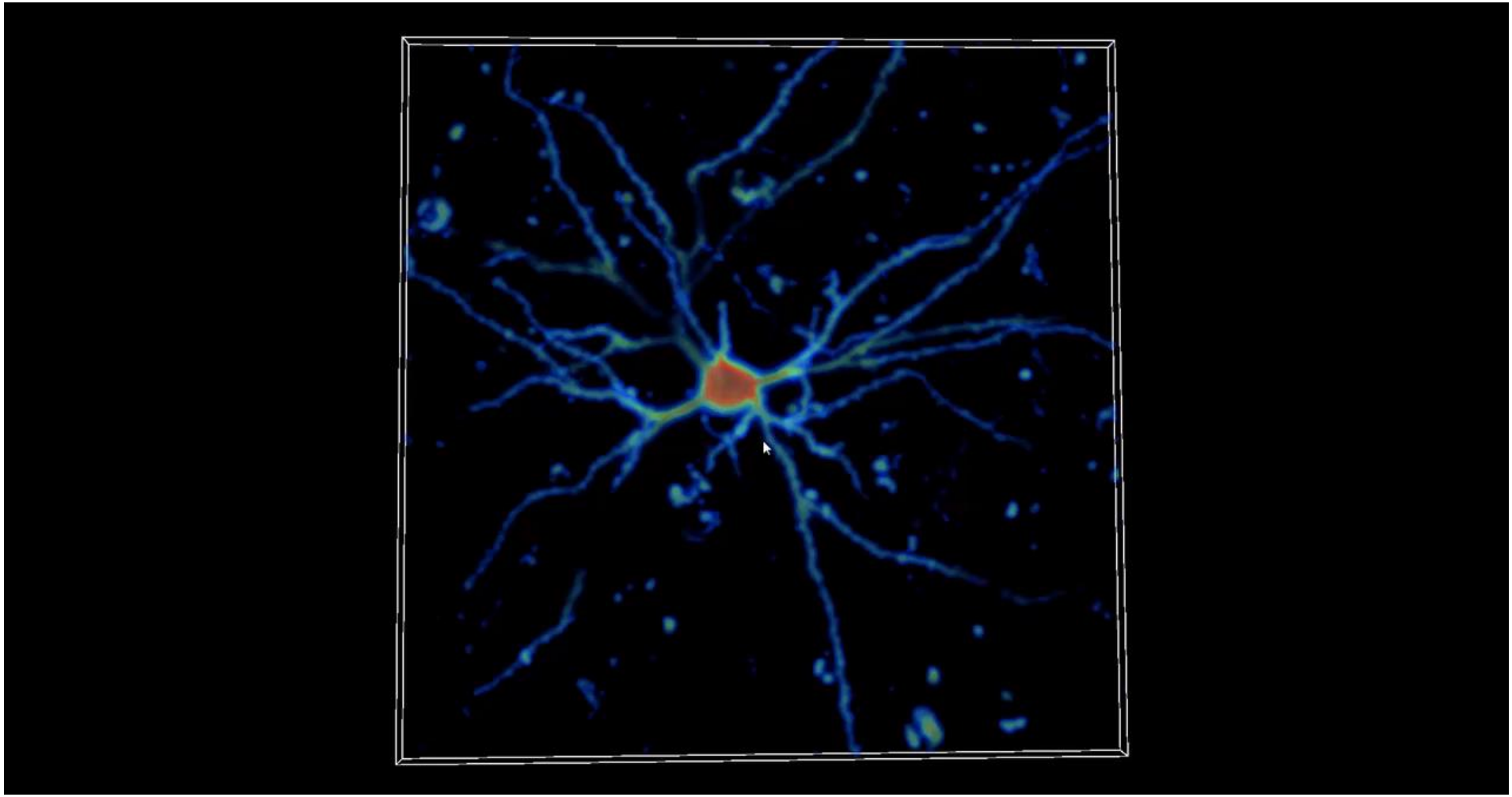


Cosmology Analysis (simulated)



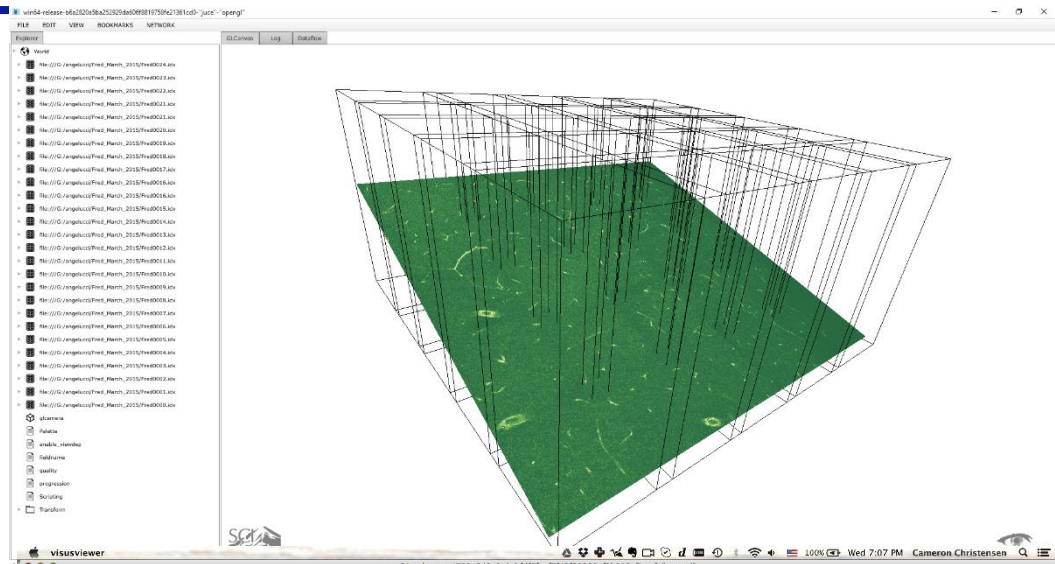
Combustion Analysis (simulated)

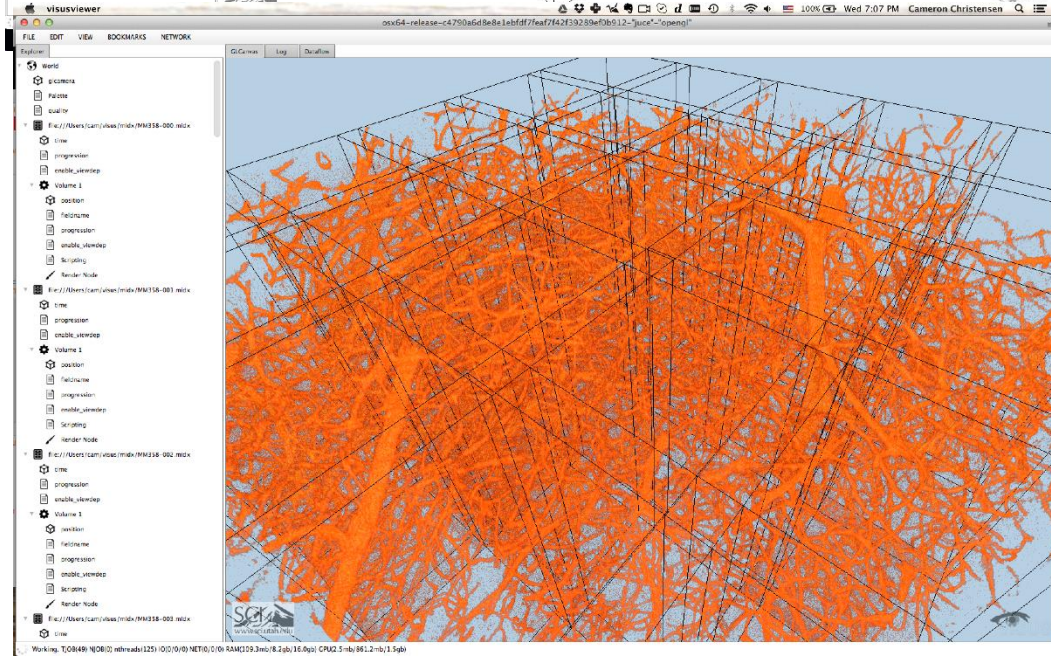
We are Working on Extraction of Neurons of a Primate (Macaque) with Topological Methods



Current Activities and Future Plans

- **Semi automated annotation**
- **Manual and automatic registration and segmentation**
- **Use of parallel computational infrastructure for large data collections**





CEDMAV

Computational Infrastructure for Scientific Discovery is a Highly Interdisciplinary

- **Performance**
  - remote data access
  - Scaling for HPC
  - Progressive techniques
  - In-situ analytics
  - Compression
  - Asynchronous computing
- **User Access**
  - Usability
  - Platform portability
  - Collaboration
  - Data abstractions
  - Visual metaphors
  - User interactions
- **Analytics**
  - Statistic
  - Topology
  - Geometry
  - Data mining
  - Machine learning
  - Feature extraction/tracking
- **Applications**
  - Experiments: microscopy, light sources, tomography, …
  - Simulations: climate, combustion, astrophysics, …
  - Data Collection: performance, twitter, climate, healthcare, ...