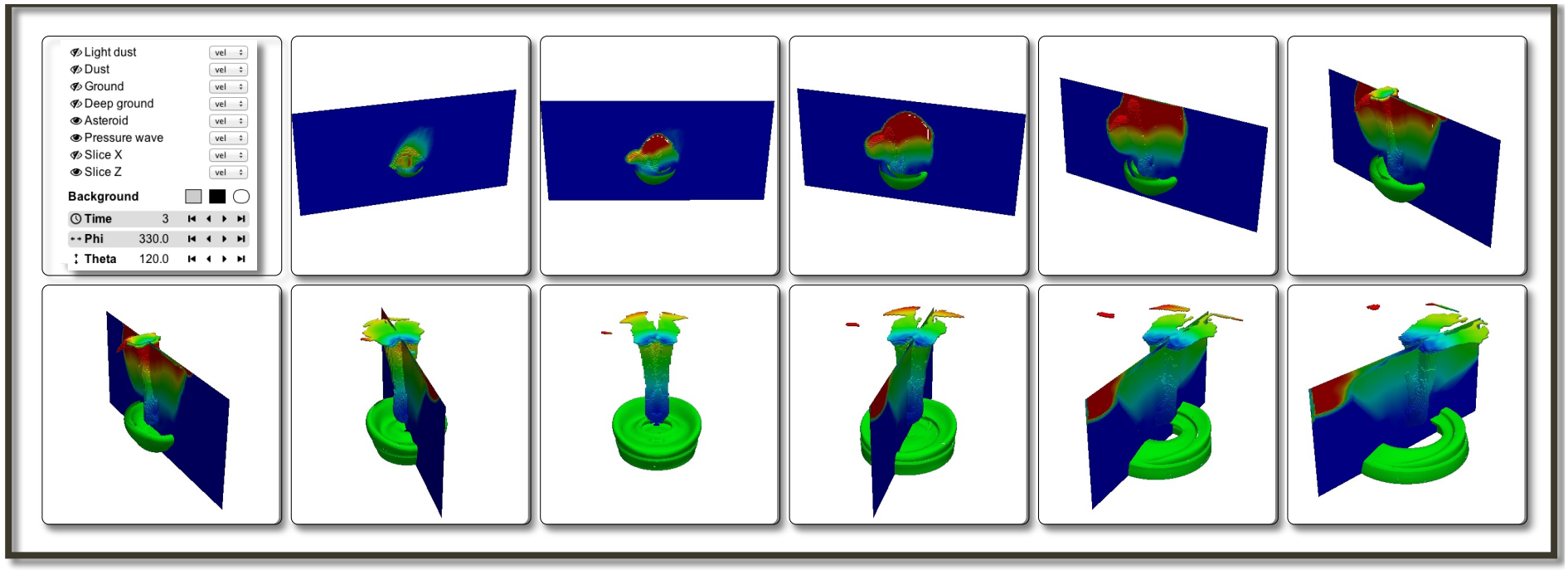


An Image-based Approach to Extreme Scale *In Situ* Visualization and Analysis

James Ahrens, Sebastien Jourdain, Patrick O'Leary,

John Patchett, David H. Rogers, Mark Petersen

Los Alamos National Laboratory

Kitware, Inc.

SC14 New Orleans, LA | hpc matters.

The traditional post-processing visualization and analysis approach is becoming unworkable

Why?

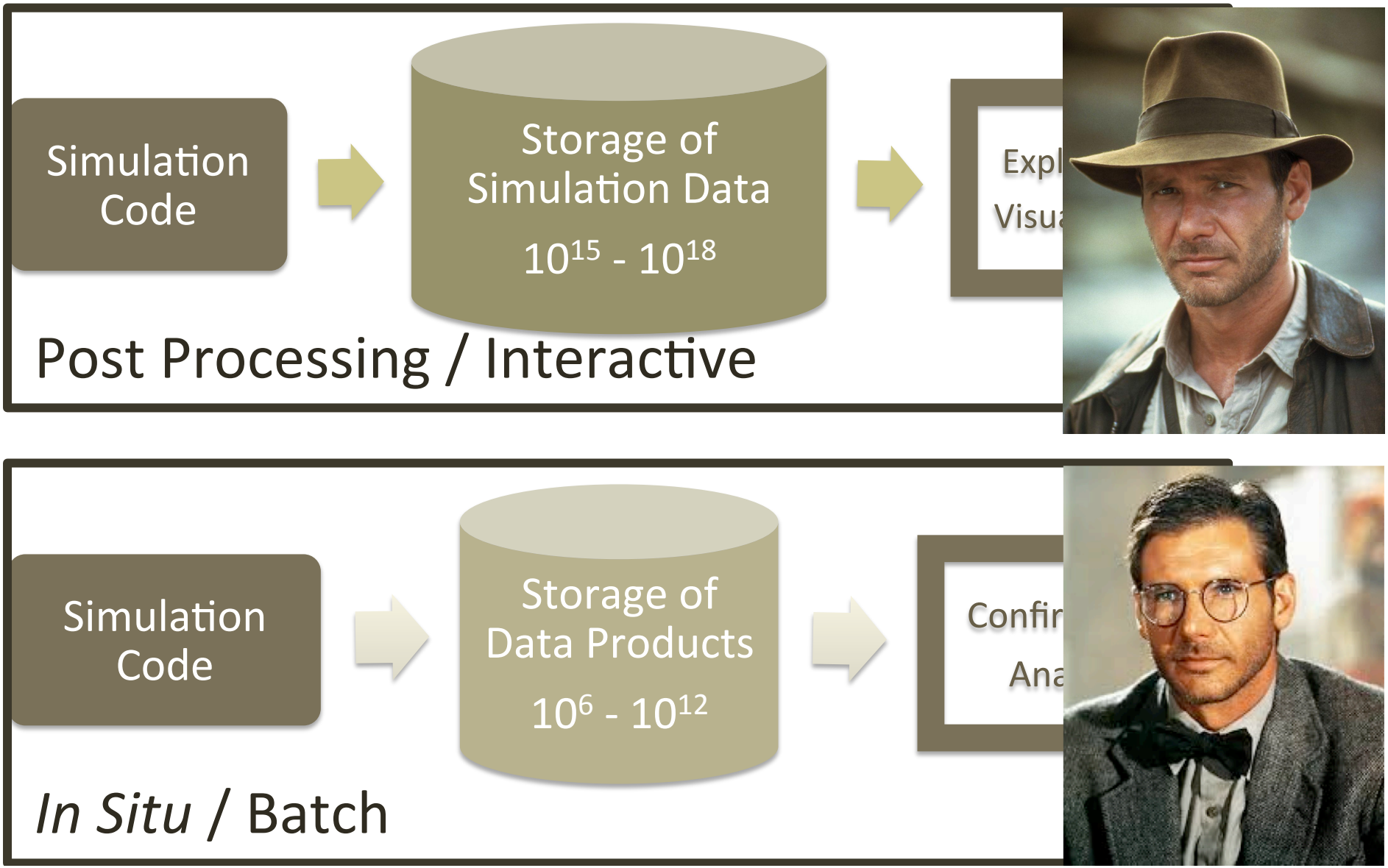
- Temporal simulation snapshots are saved at longer intervals
  - Full checkpoints are costly - less temporal data available for analysis
- Rate of improvement of rotating storage is not keeping pace with compute
  - Power, cost and reliability are becoming significant issues
- Extreme-scale supercomputing plans:
  - Peak performance to increase 3 orders of magnitude
  - System power to increase by a factor of 2
- Most expensive power operation is data movement

In Situ Approaches

- Benefits:
  - *In situ* saves reduced-sized data products during simulation run
    - Saving storage space
    - Saving time in post-processing analysis
    - Producing higher temporal fidelity results
  - Help manage cognitive and storage resource ***budget***
    - Prioritized by scientist's importance metrics
    - Answer specific analysis questions
- Challenge:
  - *In situ* analysis occurs during batch process
  - **Concern that exploratory aspect of analysis will be lost**
- Observation
  - Store **many images in the space needed for a single scientific data dump…**

Meta	Giga	Tera	Peta	Exa
10^6	10^9	10^12	10^15	10^18
Image size	Network bandwidth	Data size	Data size	Data size

Characterization of options for extreme scale data analysis



Contributions

A sampling-based approach to *in situ* visualization and analysis

- Sampling visualization parameter space
  - Cameras, operations, parameters
- Create an image database from *in situ* analysis
  - Post-processing exploration of image database

Use cases

1) Traditional interactive exploration

Spatial, temporal

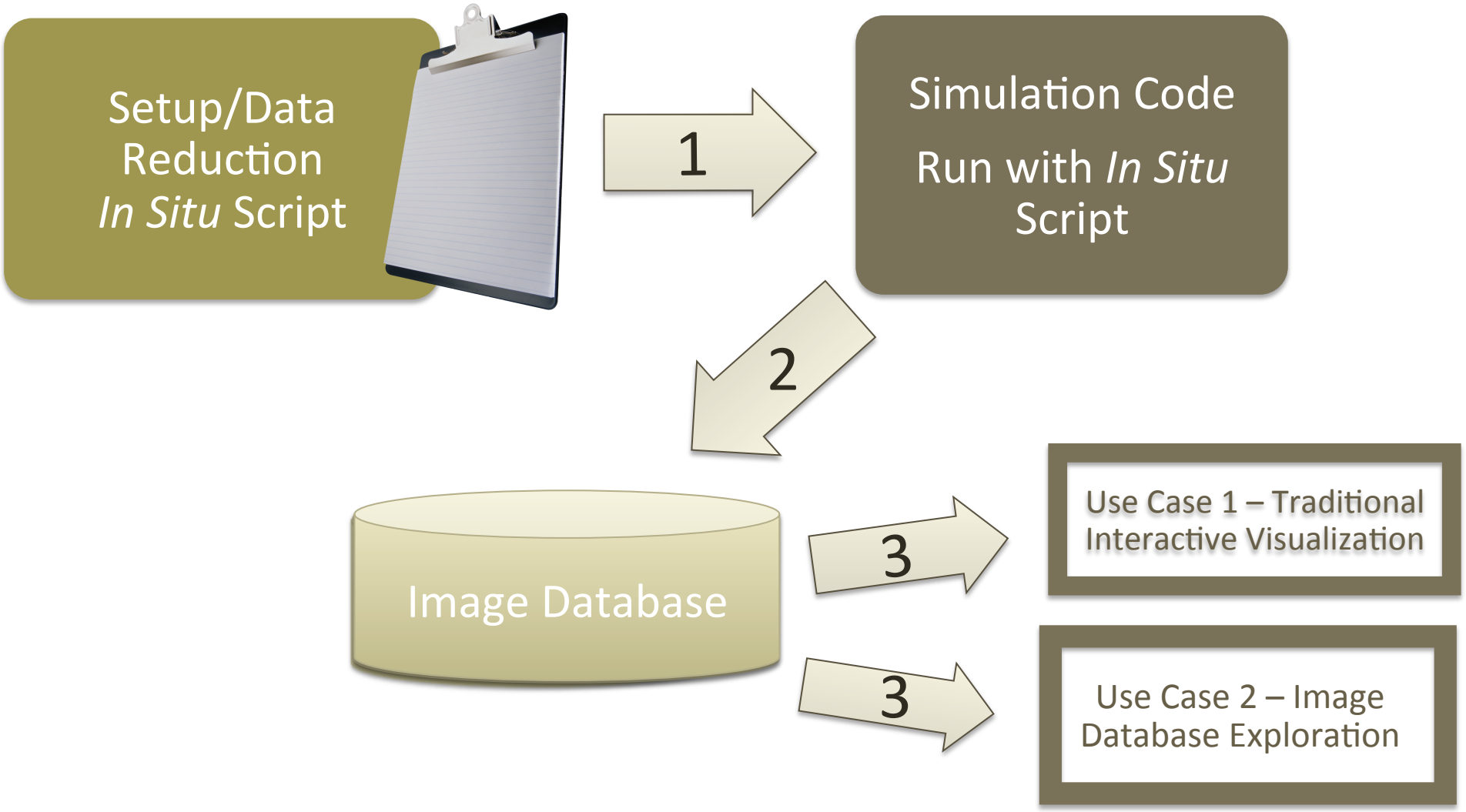
2) Image database exploration

Including content-based search

3) Creation of new visualizations

Composing operator results

Cinema Workflow



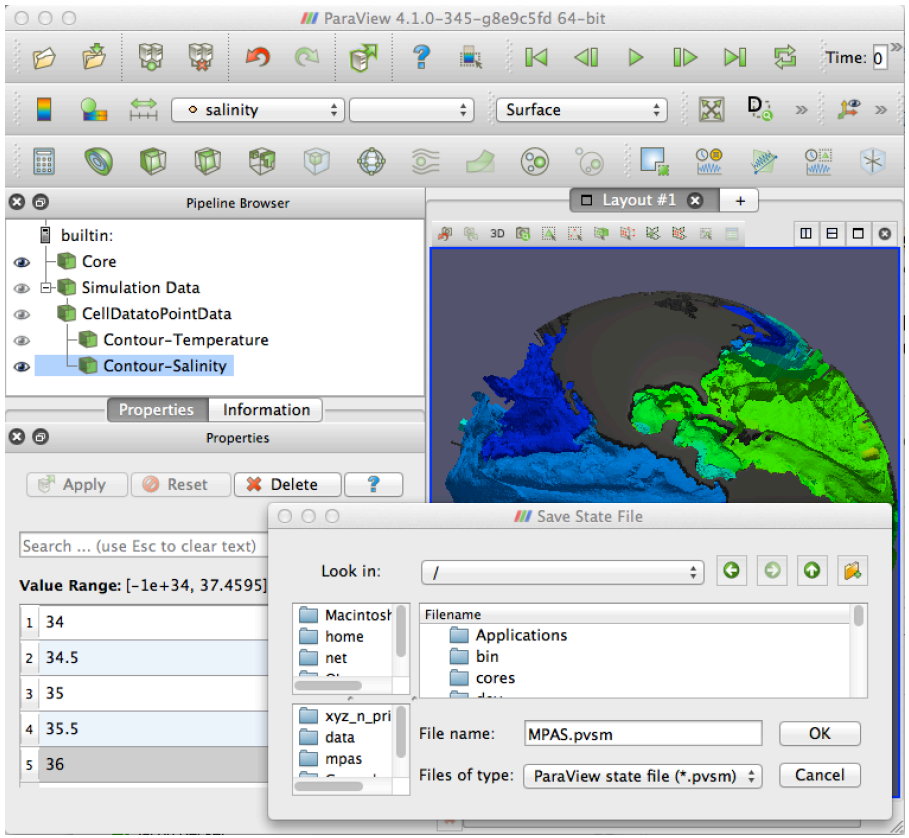
Setup /Data Reduction Phase

Setup/Data Reduction *In Situ* Script

1

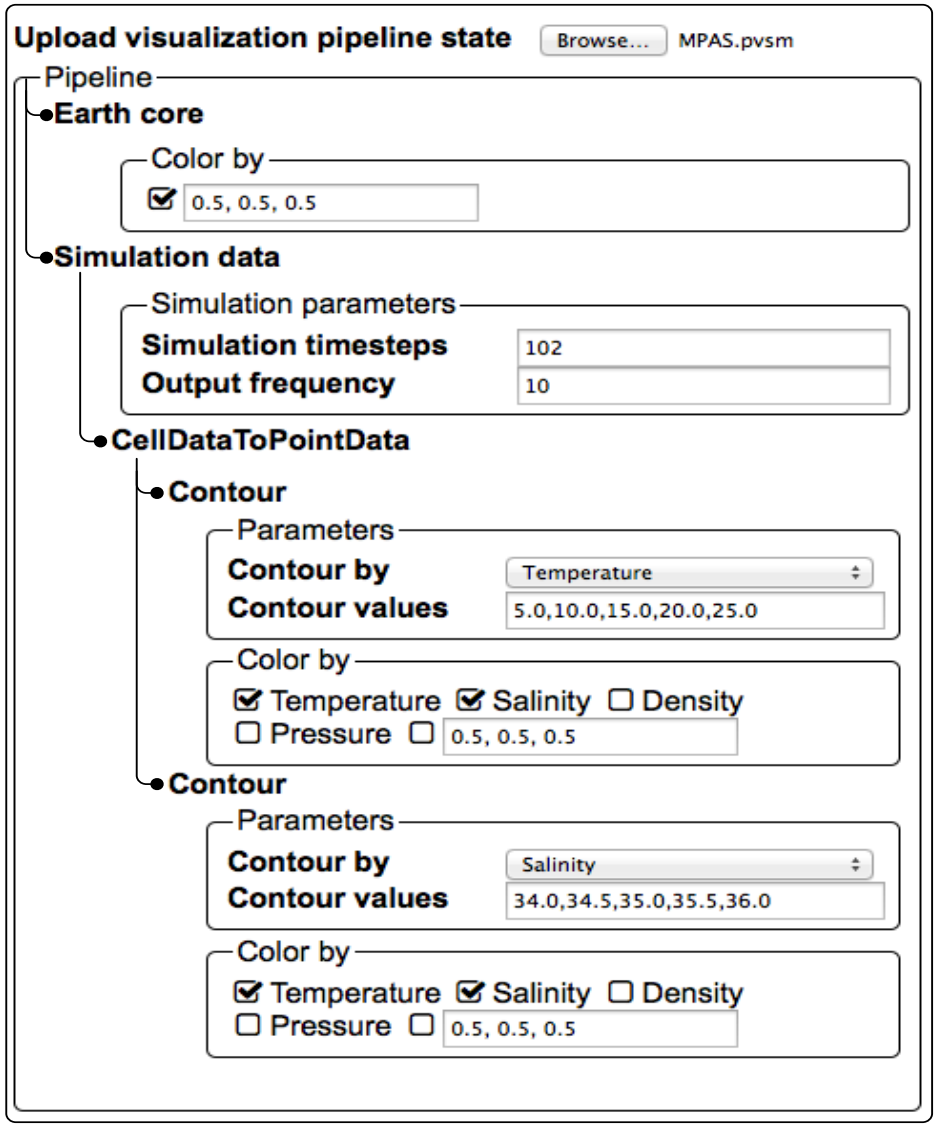
Simulation Code

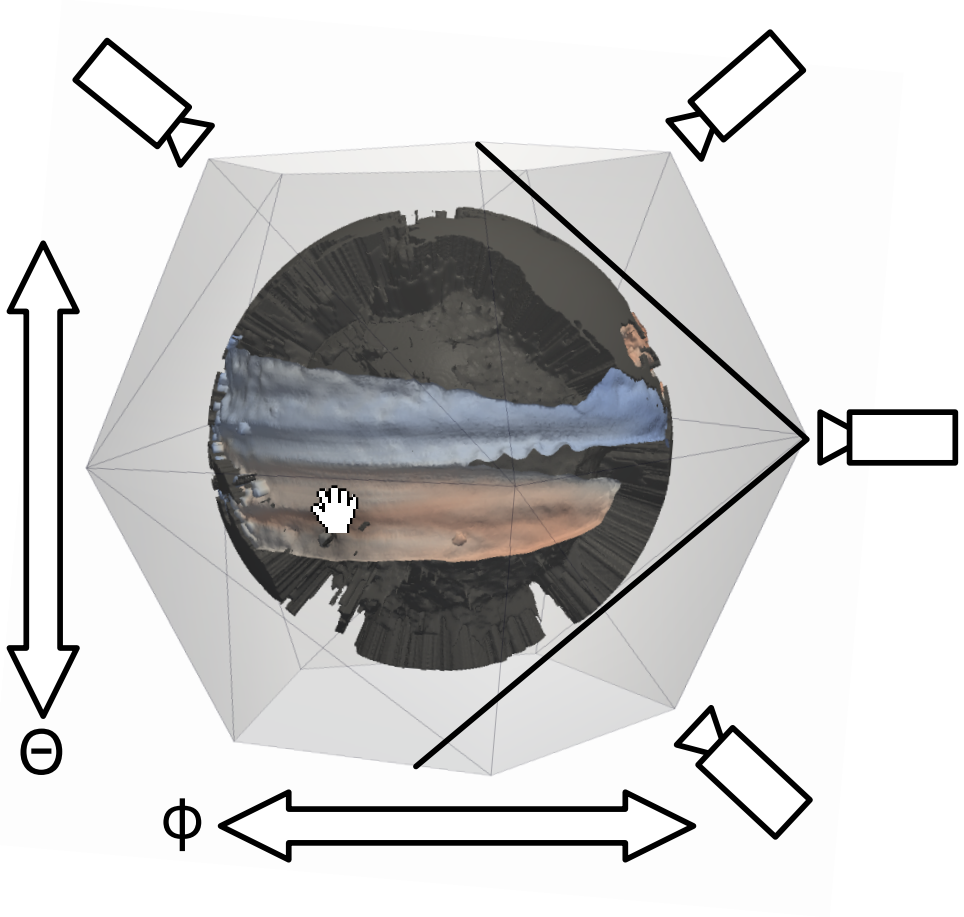
Run with *In Situ* Script



- Interactively create or reuse a visualization pipeline
  - Contains all operations
  - Specifies information needed to generate images for the database

Setup /Data Reduction Phase

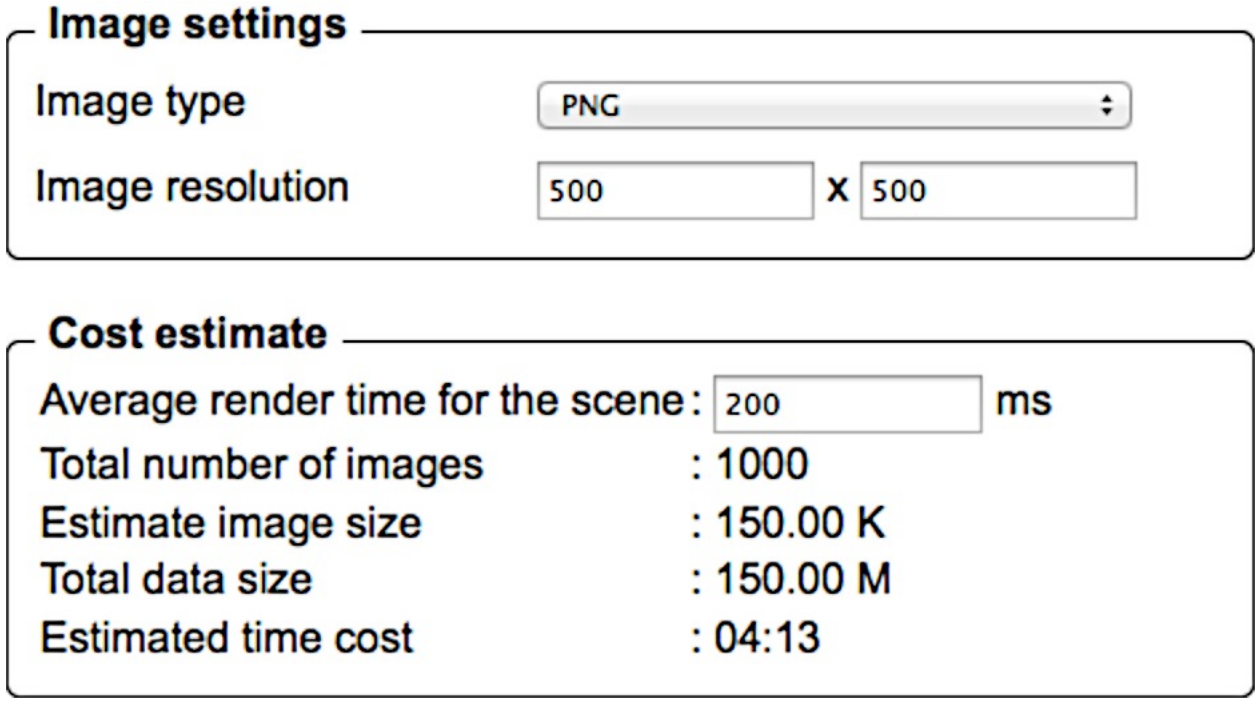




2. Set camera and operator parameters to visualize

Setup /Data Reduction Phase

3. Cost estimate section helps scientist to manage their computation and storage ***budget***



4. Create *in situ* script that generates images

Image Database

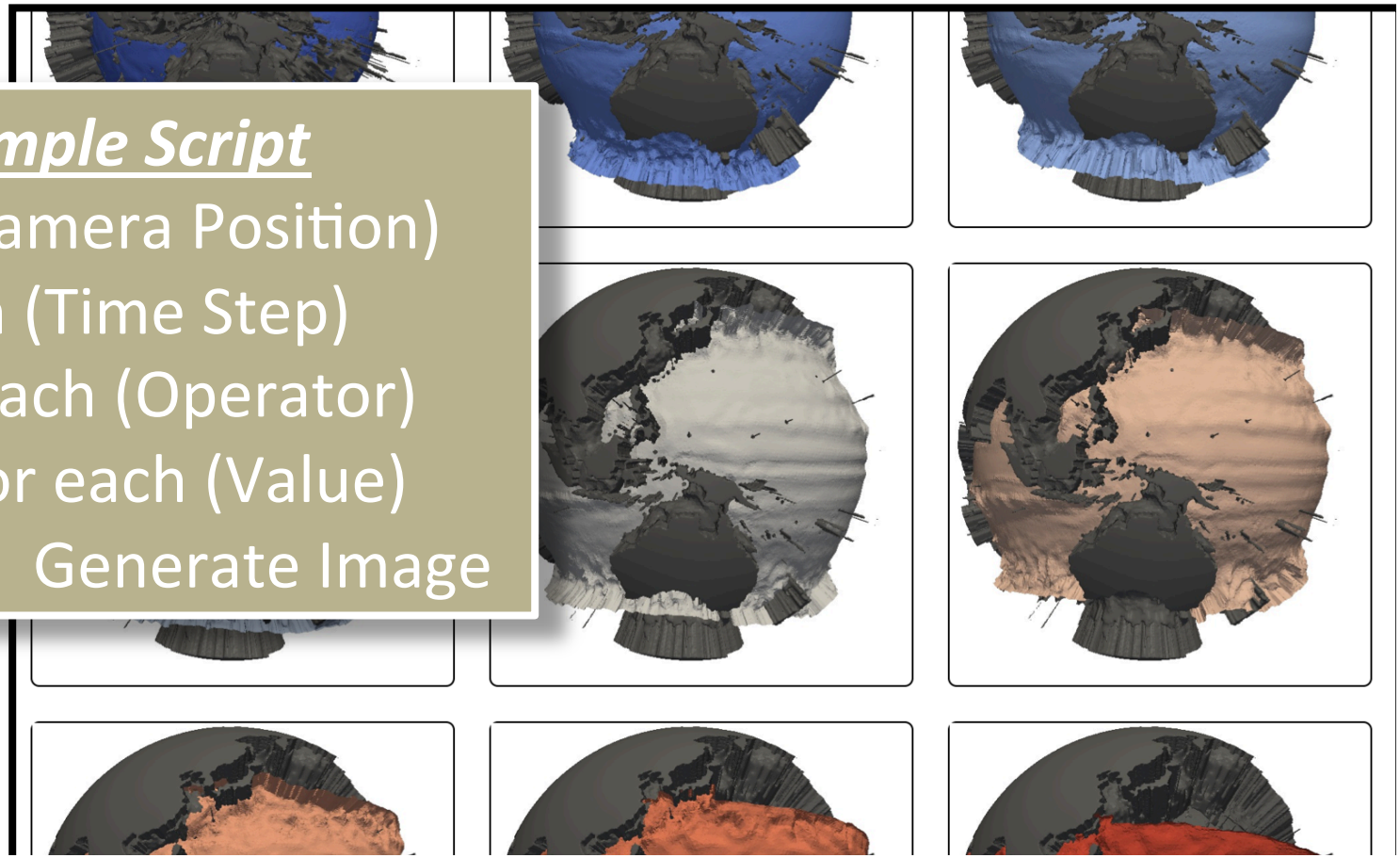
Simulation Code

Run with *In Situ* Script

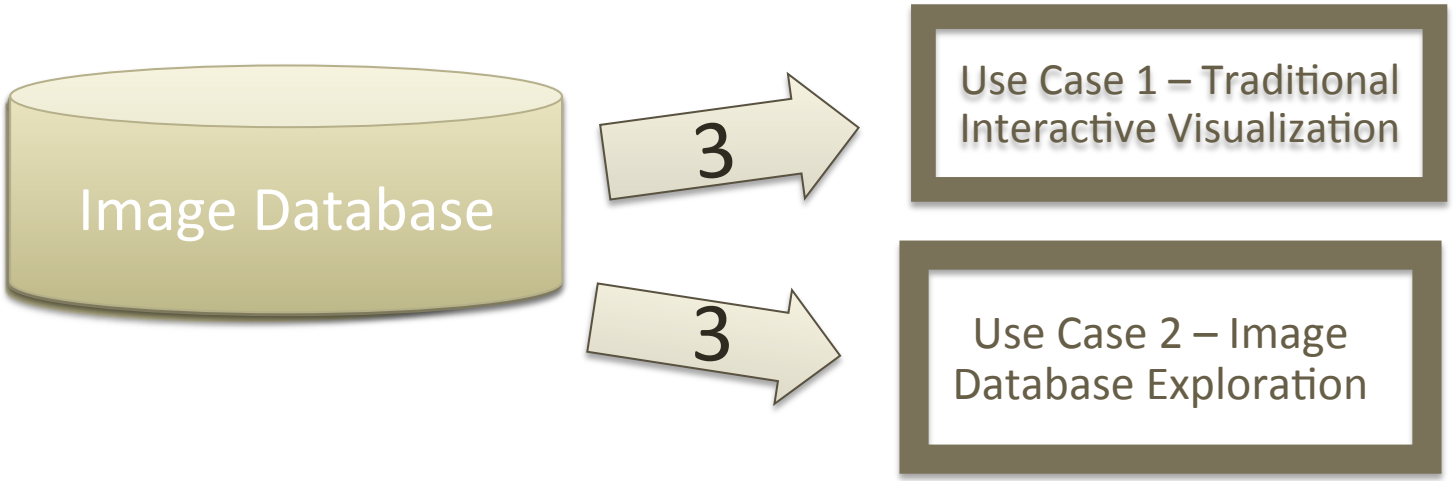
2

Example Script

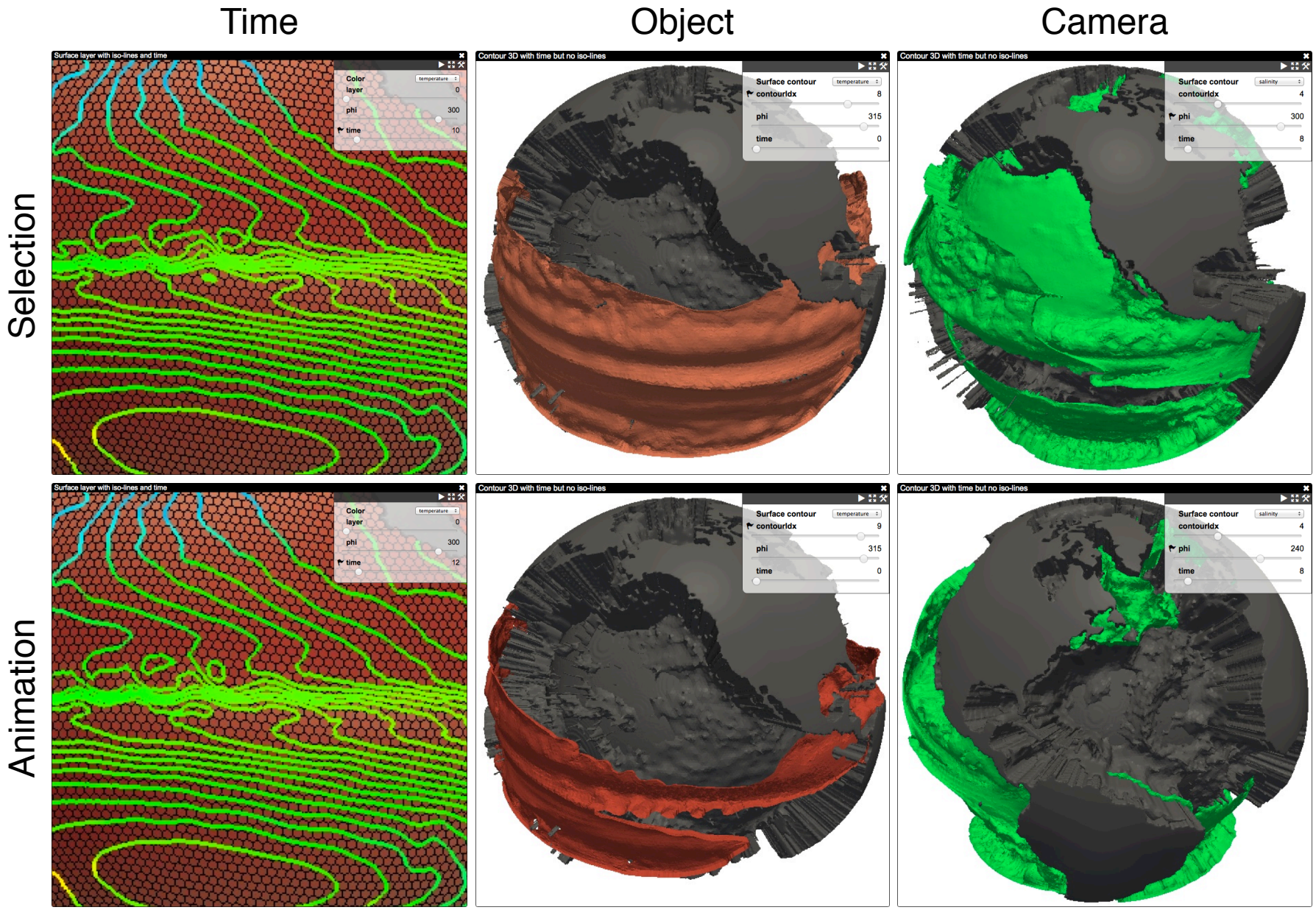
For each (Camera Position)
For each (Time Step)
For each (Operator)
For each (Value)
Generate Image



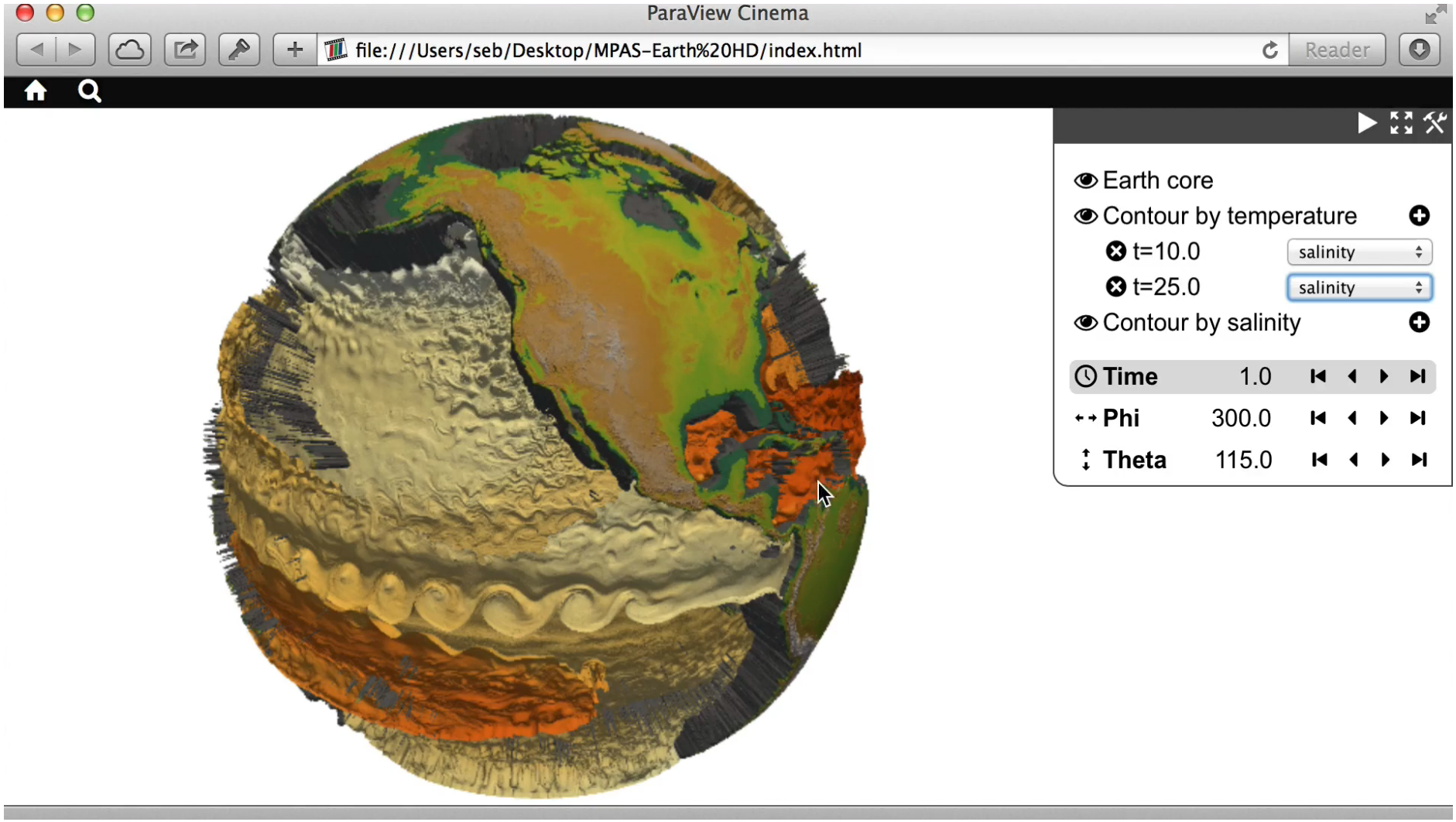
Use Case / Task-based Interfaces



Use Case 1 – Traditional interactive exploration



Use Case 1 – Traditional interactive exploration

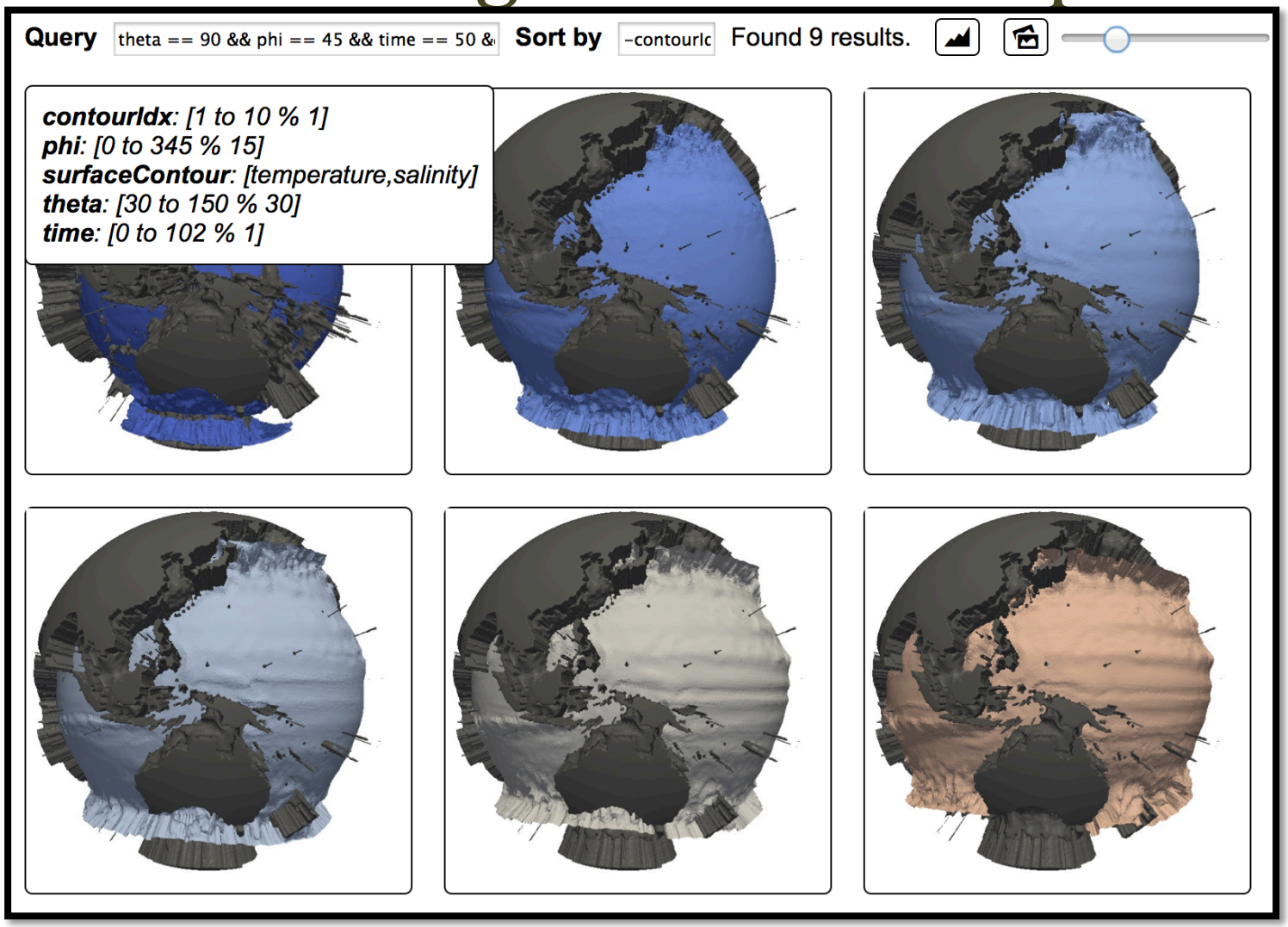


- In all videos in this presentation:

Processing, combining and showing images from the image database

- No raw scientific data is read, no geometry is created during viewing

Use Case 2 - Image database exploration



- Traditional key-value pair queries
  - Keys: Camera (phi, theta), time, operator parameters
    - Contour Index, Contour Variable

Use Case 2 – Image database exploration

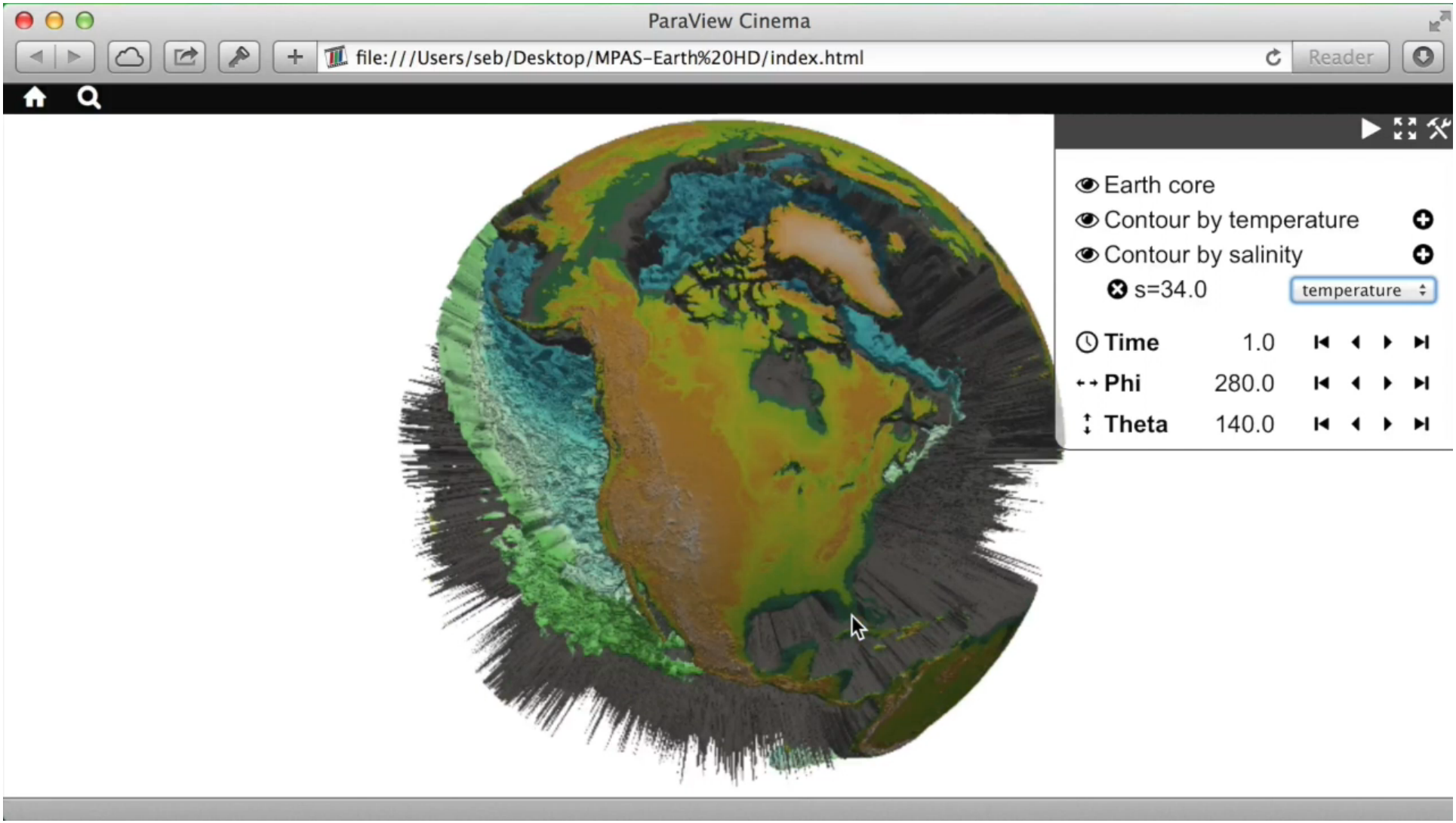
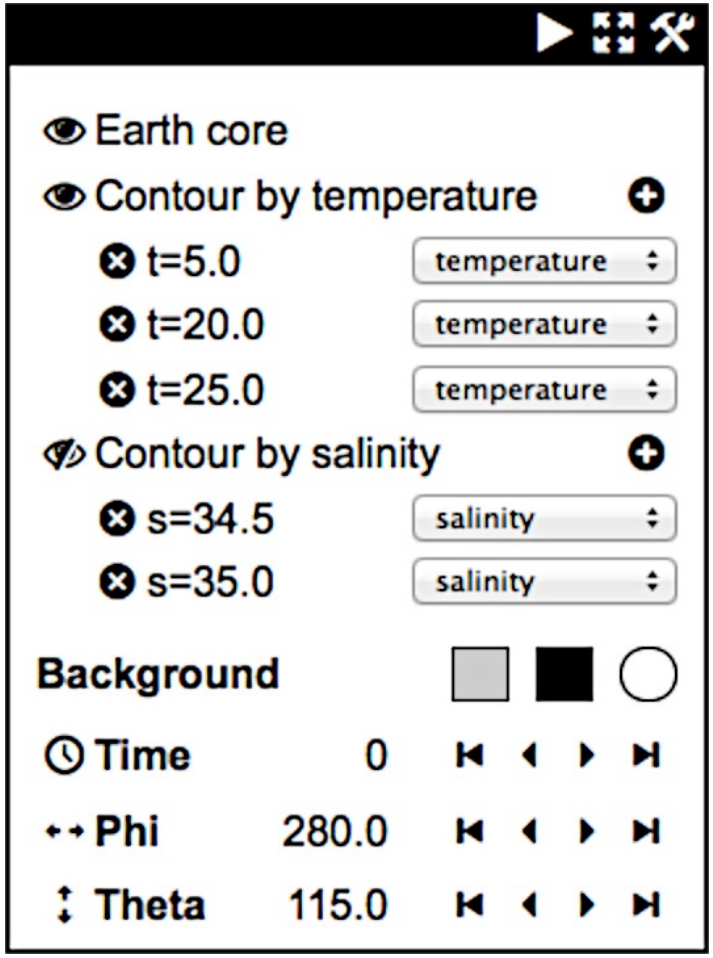
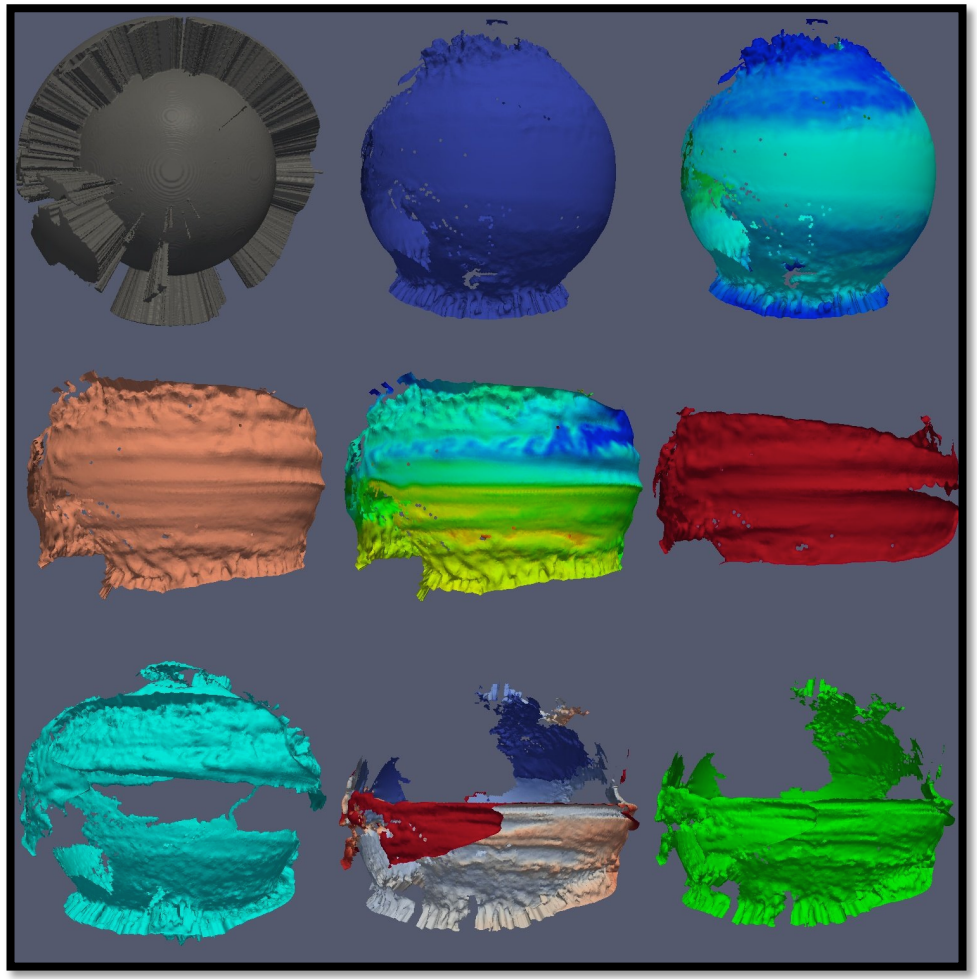


Image-based approach reduces analysis exploration bias

- Traditional post-processing approach
  - Generate visualization and analysis result upon user request
  - User wait time is extremely variable
    - Rendering (seconds)
    - File system accesses (minutes)
  - Creates inherent bias in what is explored
    - For example: little significant interactive temporal analysis
- For an image-based approach
  - All “operations” take the same amount of time
    - Reduces bias of what get explored

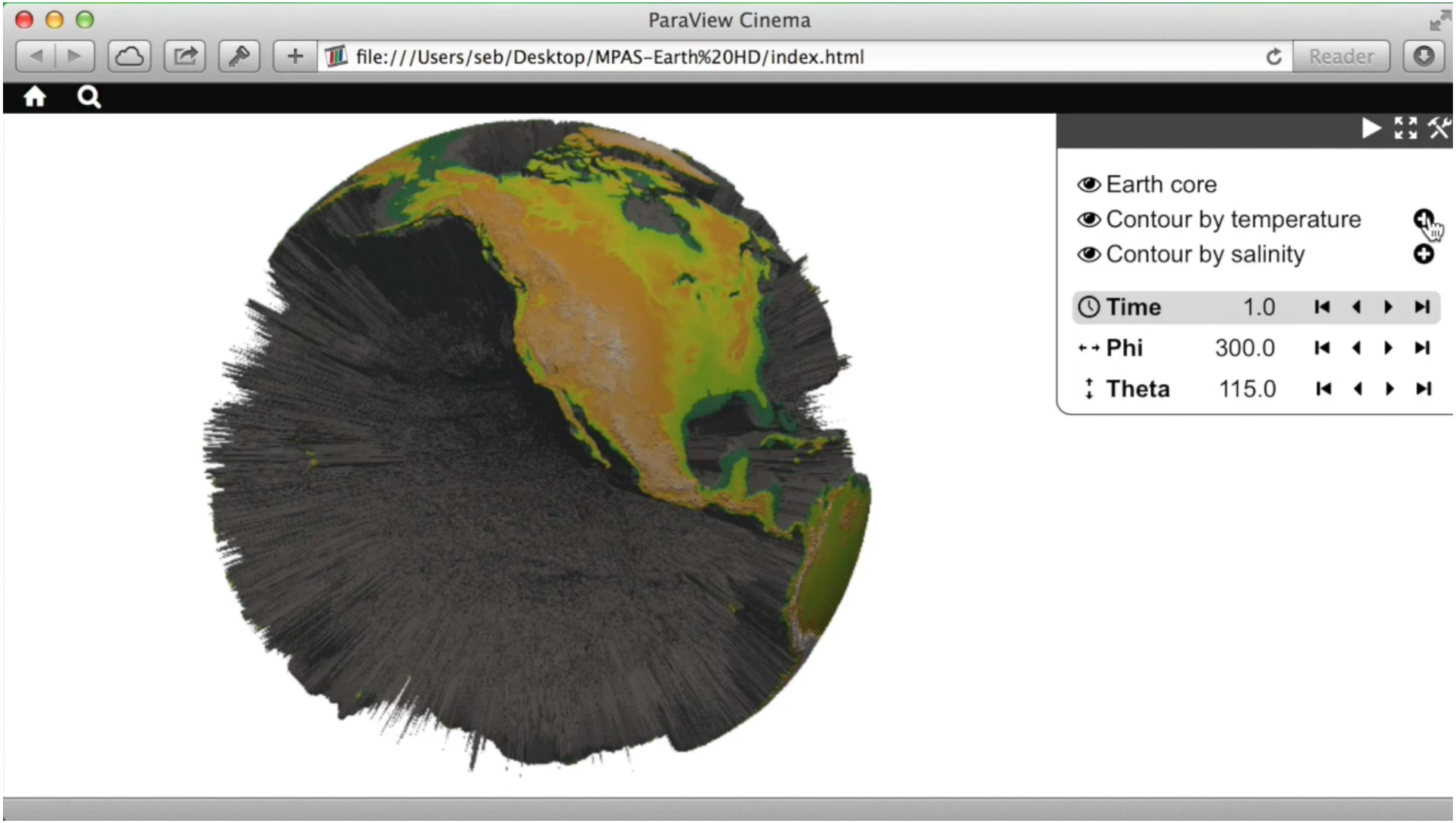
Use Case 3 – Creation of new visualizations





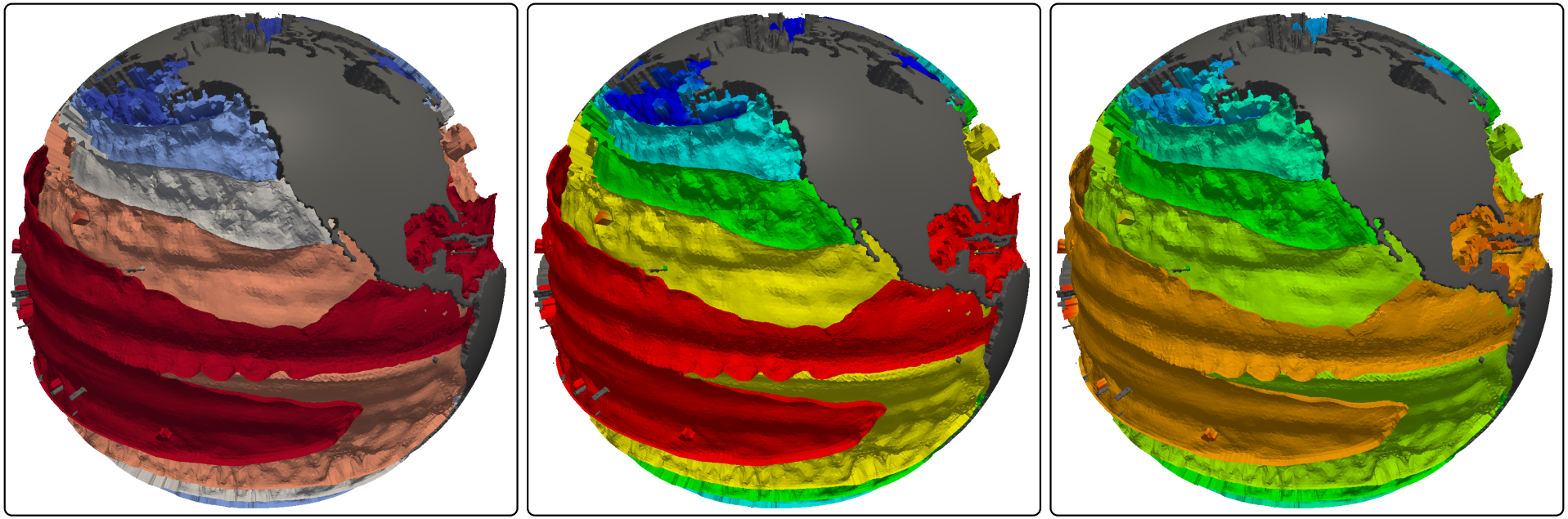
- Use real time image compositing to build new pipelines
  - Image representation: Color & depth buffer
  - Multitude of combinations/visualizations possible

Use Case 3 – Creation of new visualizations



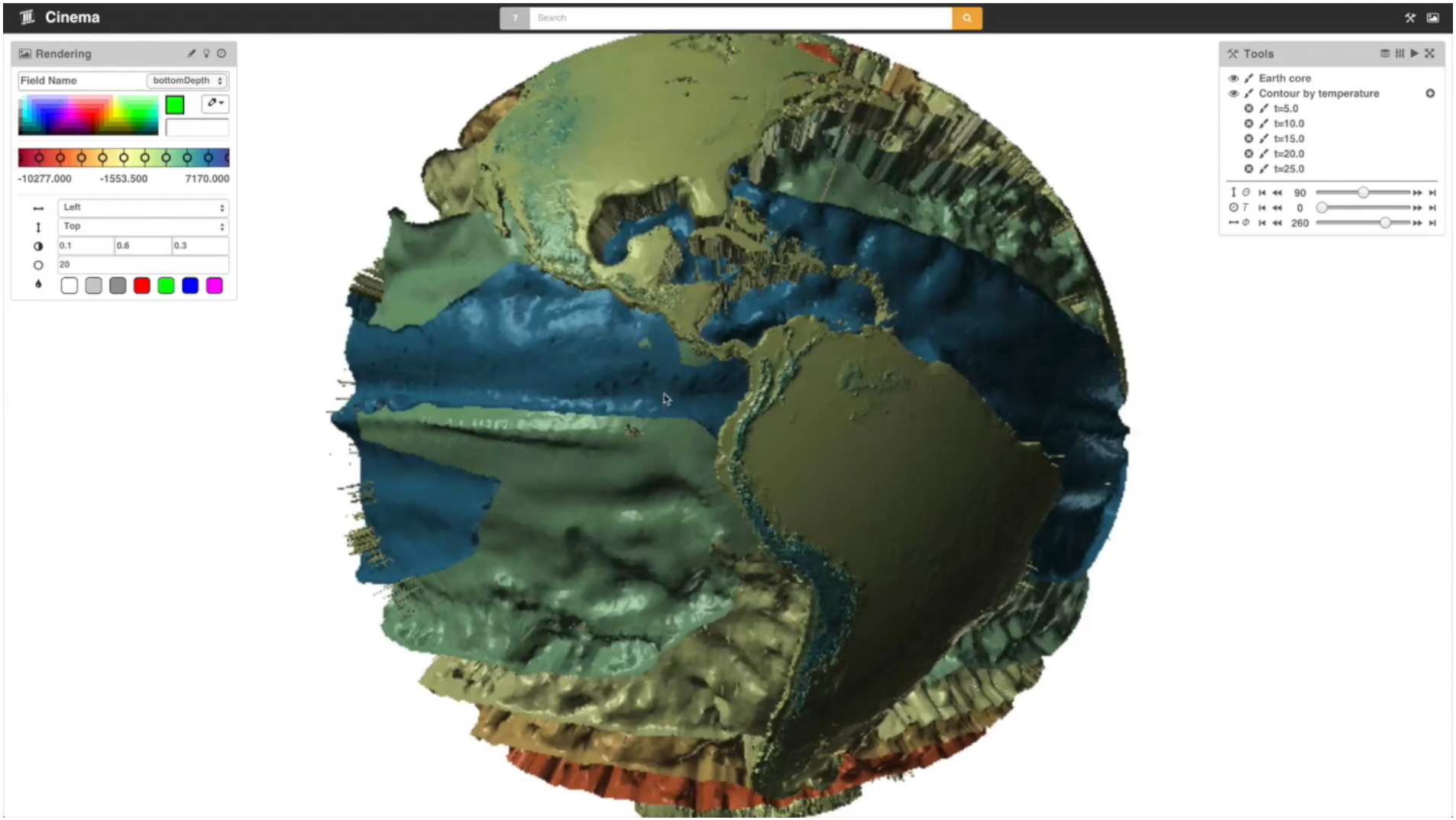
- Scientists can quickly create “arbitrary” pipelines to answer their analysis questions

Use Case 3 – Creation of new visualizations

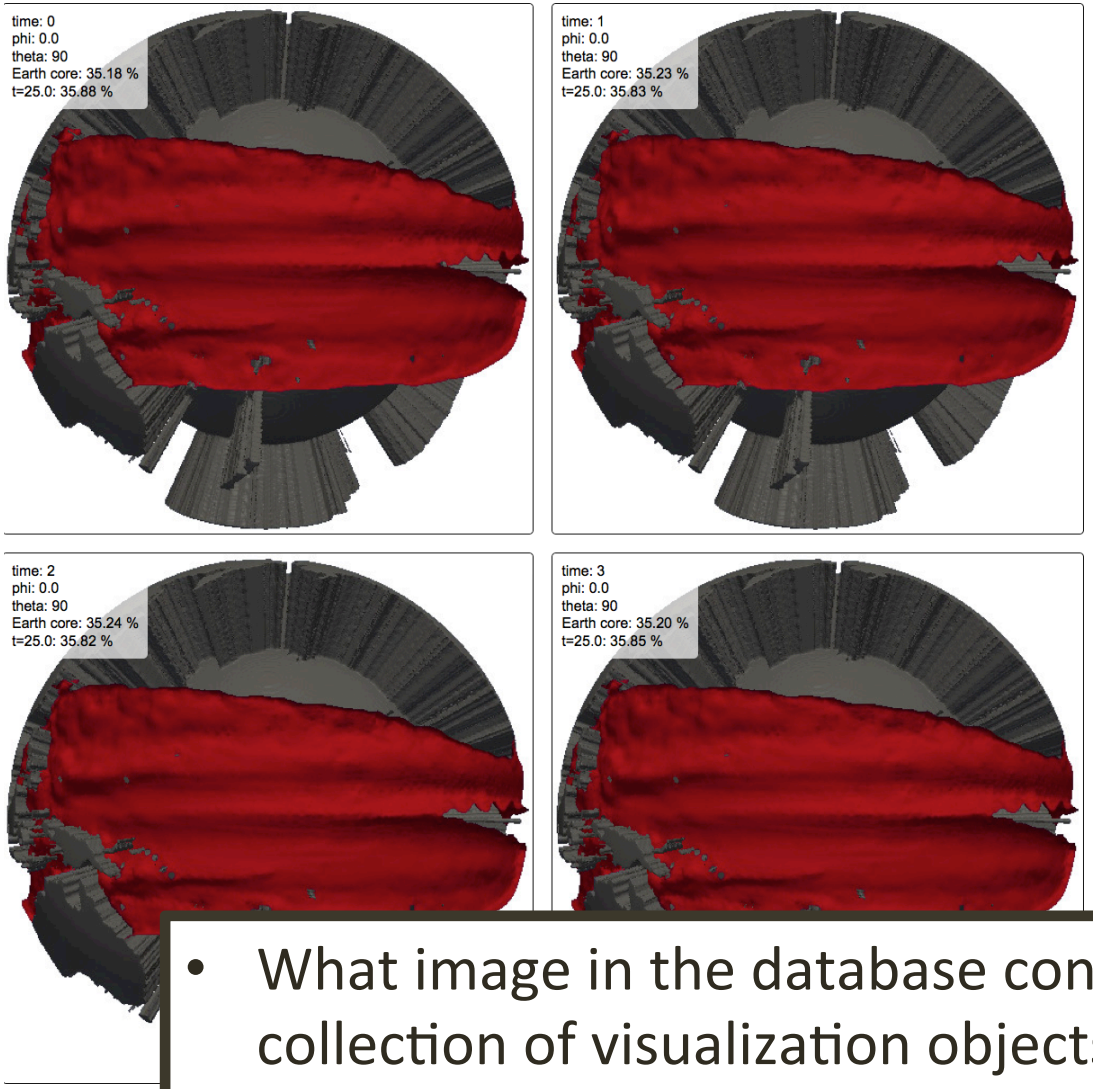


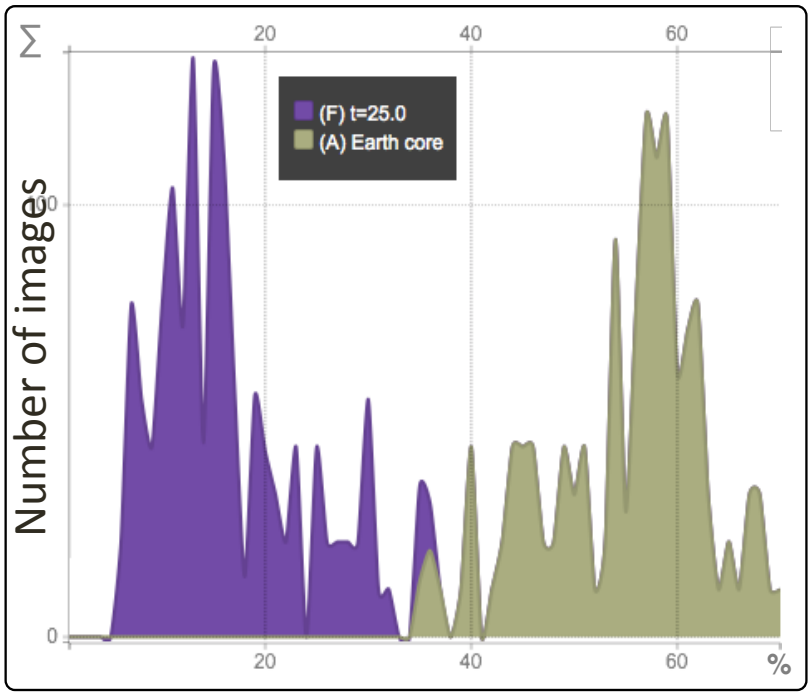
- Use real time image compositing to build new pipelines
  - Image representation: Raw scientific data, depth buffer, normals
  - Dynamically changing colormap and lighting possible

Use Case 3 – Creation of new visualizations



Use Case 2 & 3 – Content-based image search

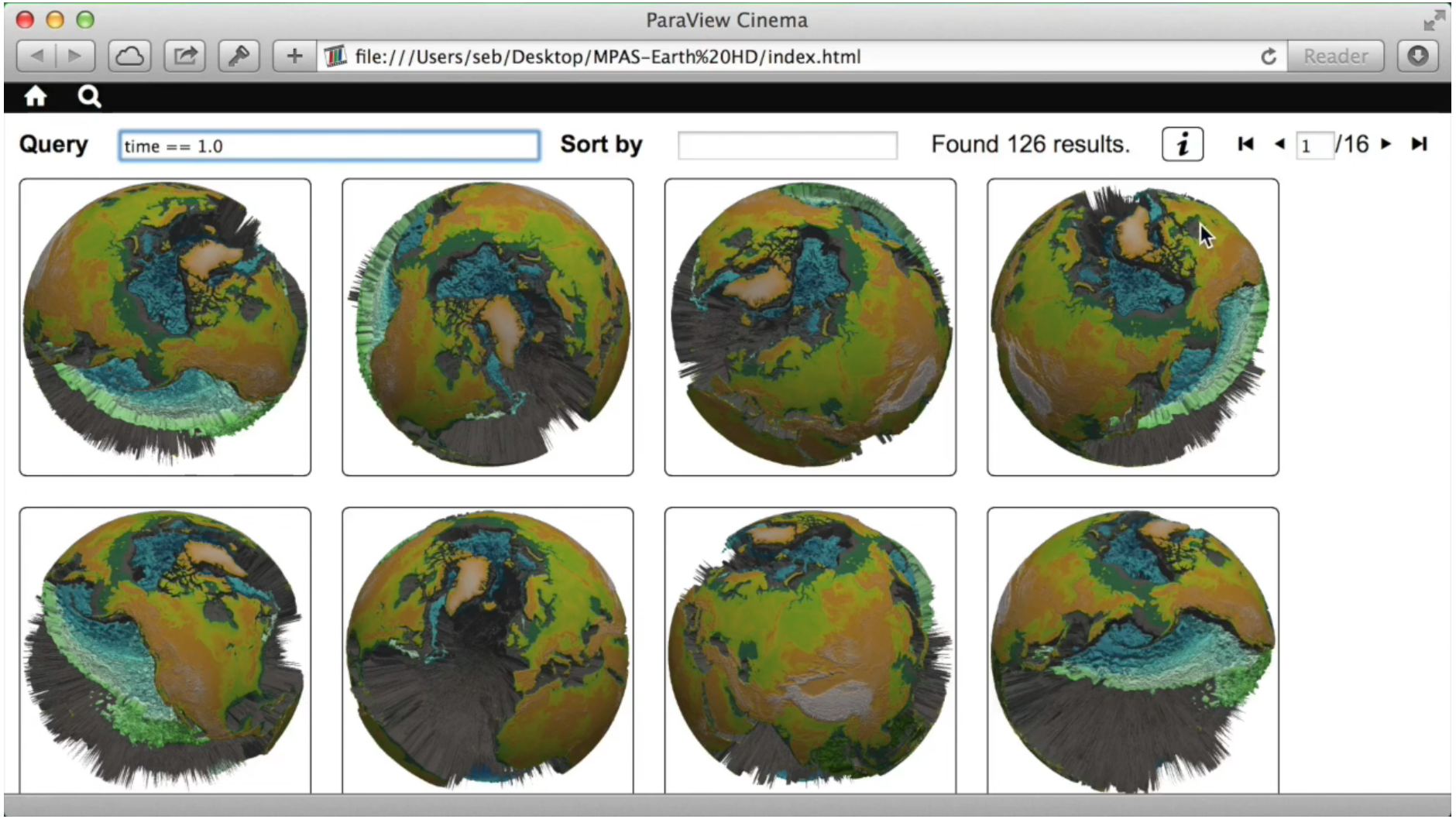




Percent of image covered

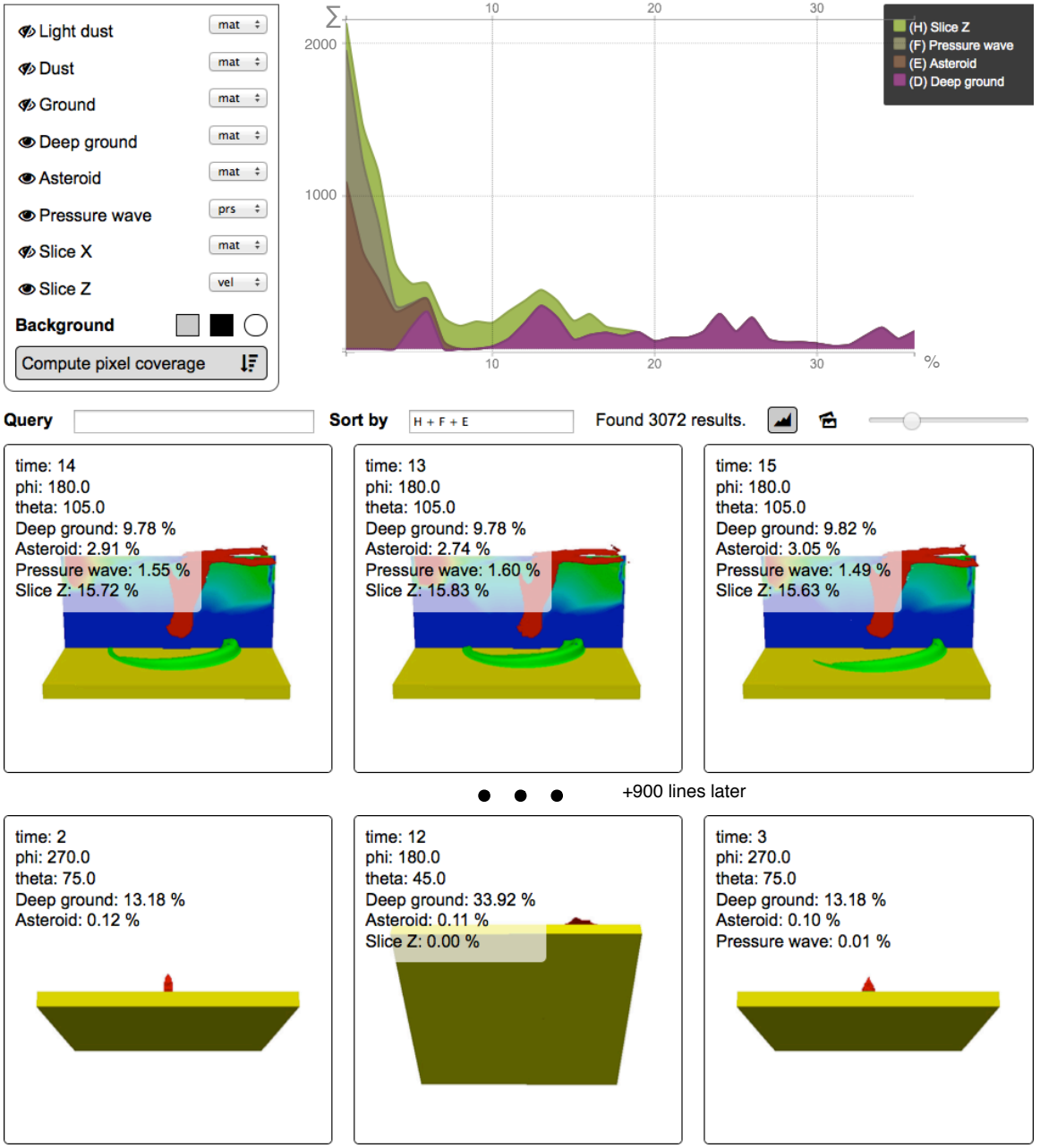
- What image in the database contains the “best” view of a collection of visualization objects?
  - Each image/pixel contains a list of the order/visibility of the objects for each view
  - Pixel coverage is calculate for all views and objects

Use Case 2 & 3 – Content-based image search

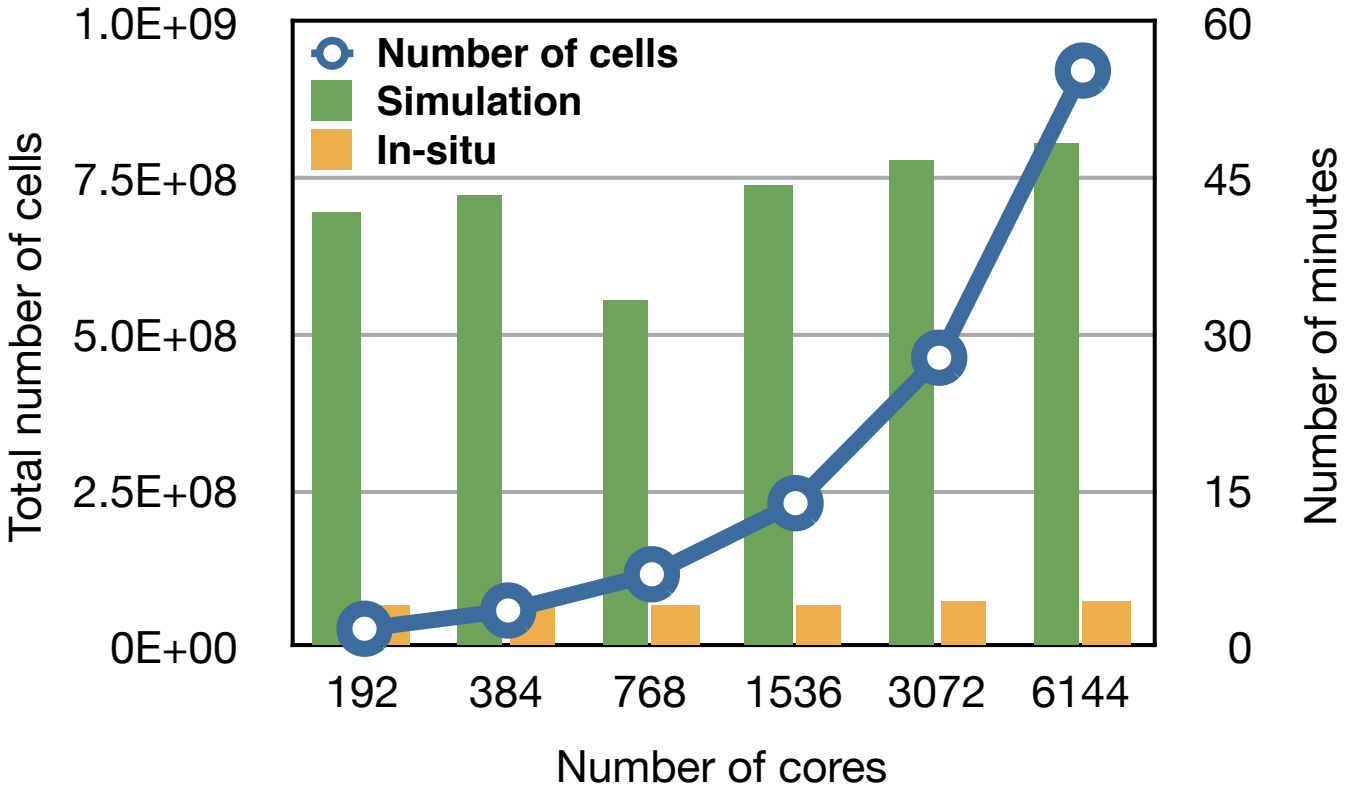


- Unique exploration interaction approach
  - Enabled by image database

Use Case 2 & 3 – Content-based image search

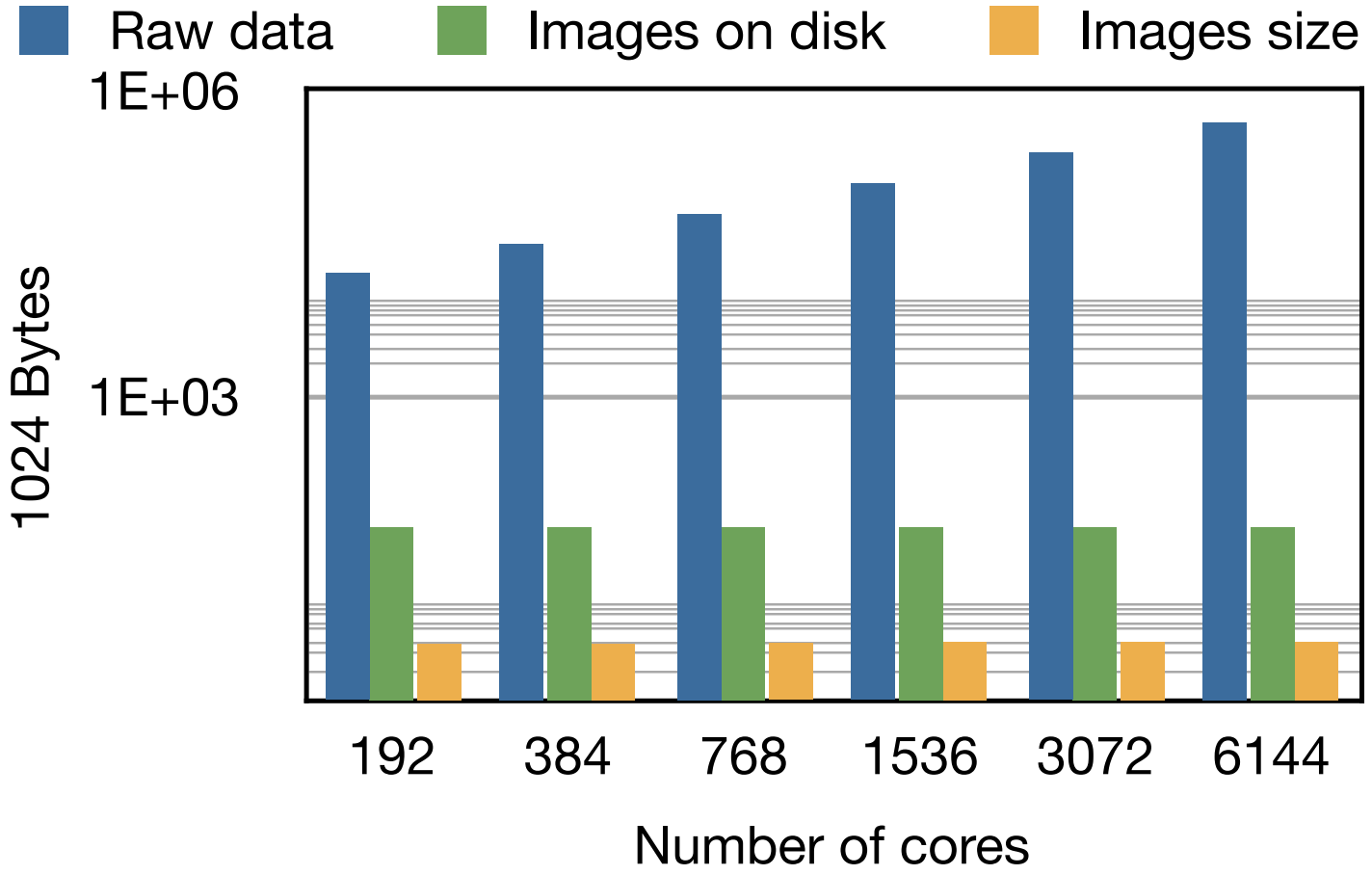


Weak Scaling of XRage Simulation and *In Situ* Analysis



- *In situ* analysis of 10 contour objects and background
  - image size of 500x500
- Summary: Scalable *in situ* performance to generate database

Disk usage reduction
Full XRage data files versus *in situ*



- Summary: Orders of magnitude data saving with Cinema approach

Conclusions

- Workflows will change at exascale
  - Image-based approach
    - Preserves important elements of the simulation
    - Significantly reduces data saved
    - Offering unique interactive exploration options

Acknowledgements

This work was funded by Dr. Lucy Nowell, Program Manager for the Advanced Scientific Computing Research (ASCR) program office in the Department of Energy's (DOE) Office of Science

Questions


Workbench
Run
Layout
Select a run from the Run menu in the header bar.
HACC Cosmo
Cinema Volume Visualization

An Image based Approach

Saving images is a very viable approach at extreme scale

- Simulation sizes are 10^15 going 10^18
- Image sizes are 10^6 and bounded by visual acuity

1. During *in situ* analysis save out massive image database

- Images contain values and depth
- Cartesian product of camera positions, operators, variables, timesteps
- Guided by budget and analysis questions

2. During post-processing

Visualization process is now the exploration of the image **database**

- Interactive visualization
  - Change camera, operators
    - Through image compositing and rendering techniques
- Meta-data and content-based image search
  - Explore and search from simple web based interface