Advanced Research Computing

Embracing heterogeneous simulation of complex fluid flows

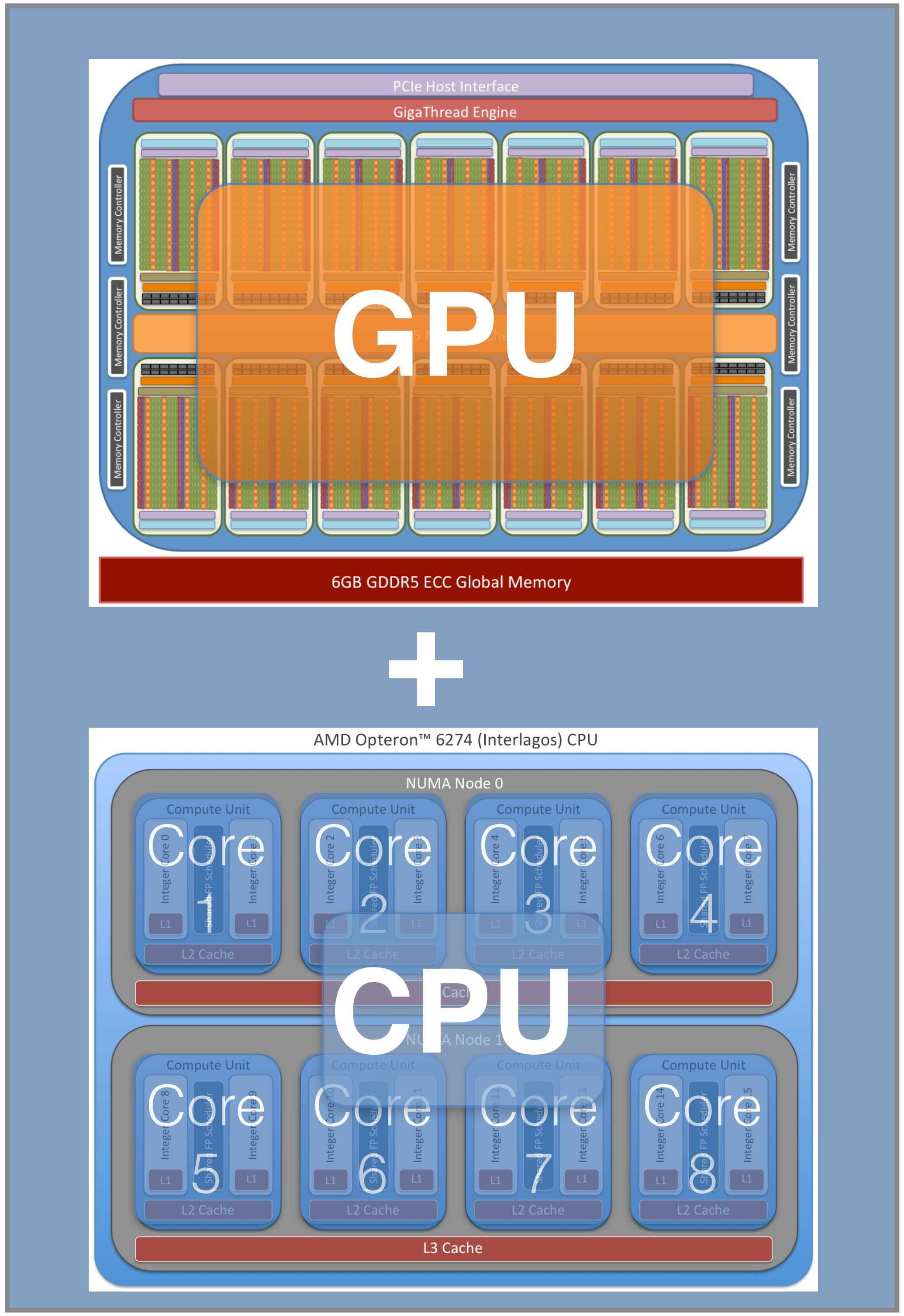
James E. McClure
Advanced Research Computing
Virginia Tech

Motivation

- Scientific opportunities and challenges associated with new data sources
  - *Digital Rock Physics*: experimental data + simulation to understand geologic materials

- Challenges associated with exascale
  - Theoretical performance gains due to accelerators
  - Massive parallelism, massive data

Complementary Processors



GPU

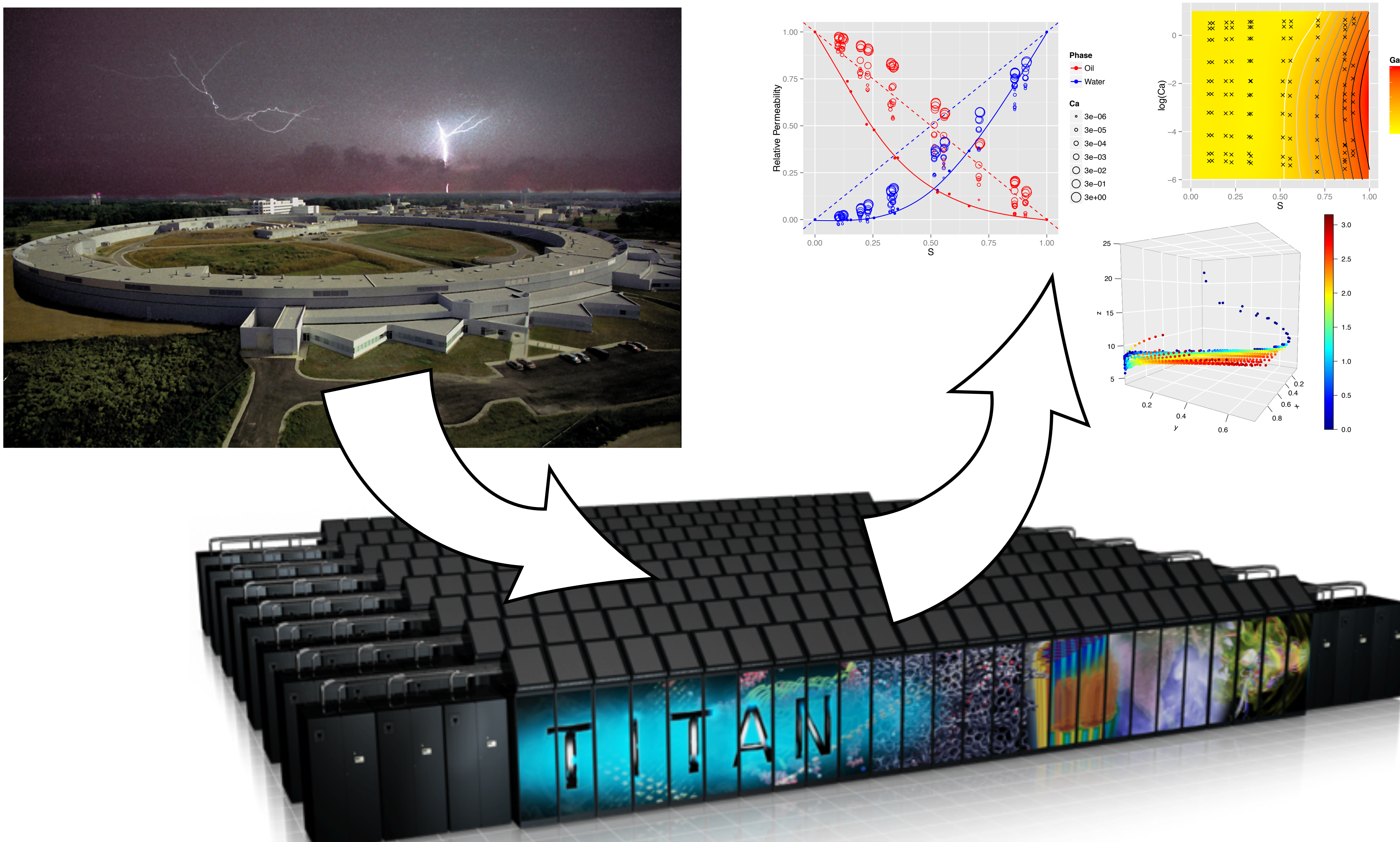
- Slow clock speed
- Massive parallelism

CPU

- Fast clock speed
- Modest parallelism

Advanced Research Computing

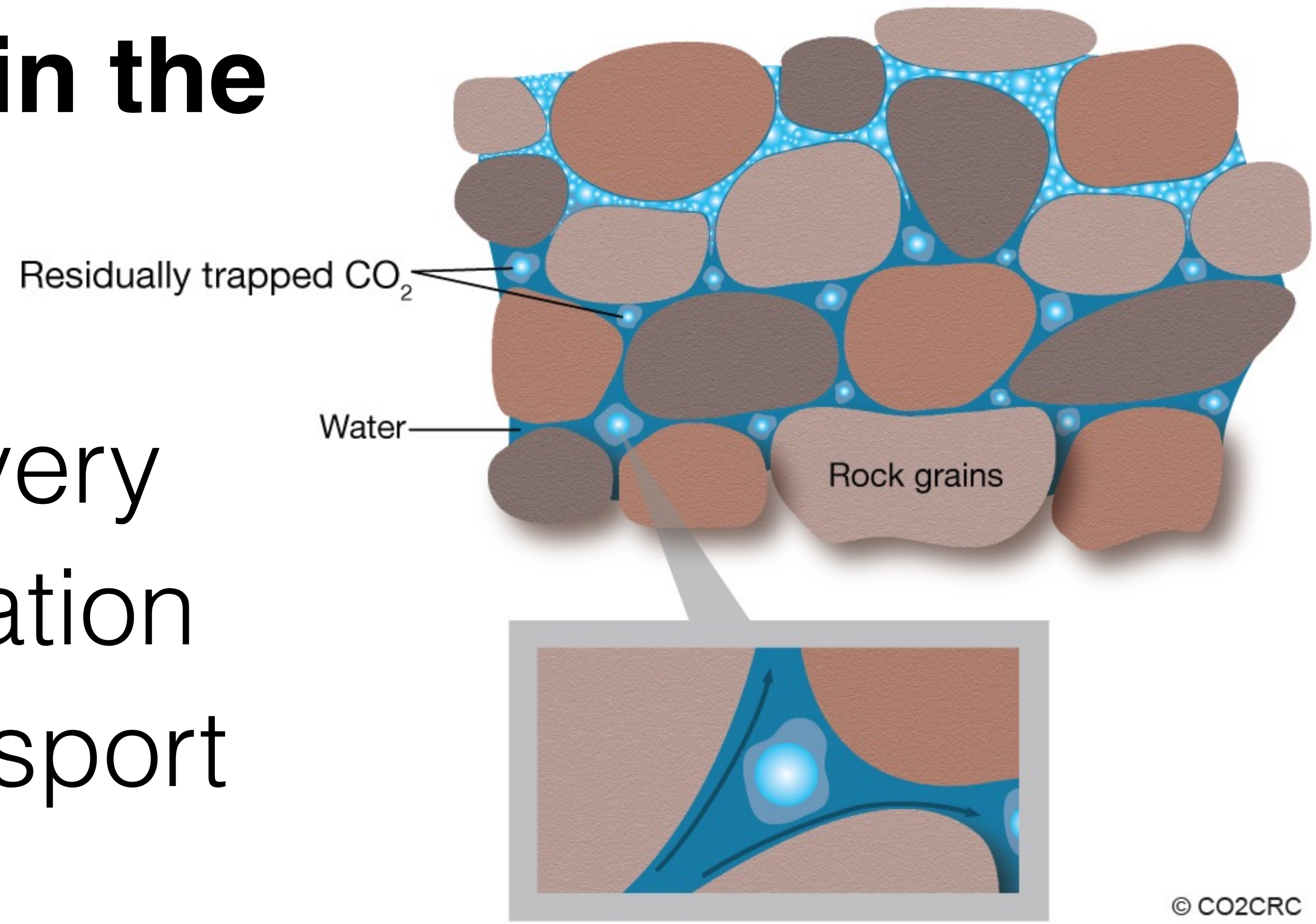
Simulation as Data Analysis



Geological Systems

Multiphase flow in the subsurface

- Oil and gas recovery
- Carbon sequestration
- Contaminant transport



X-Ray Tomography

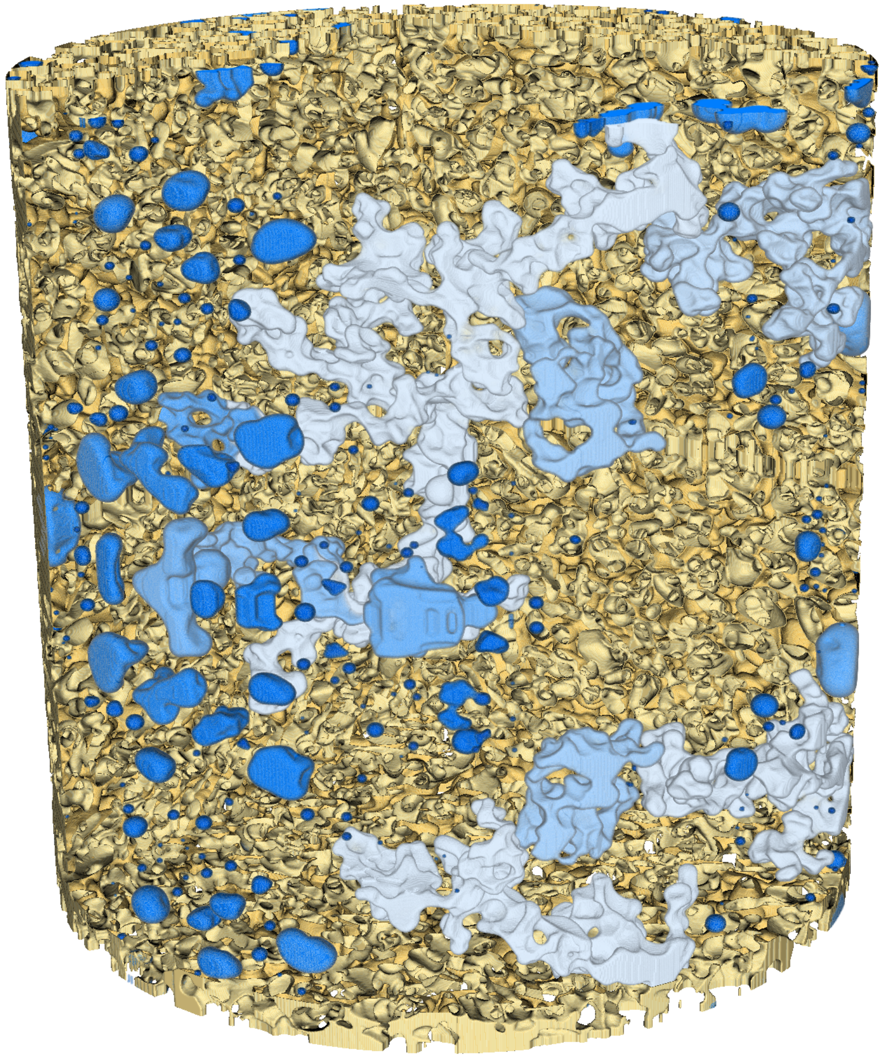
- X-ray tomography based on synchrotron light sources can be used to obtain three-dimensional images of complex microscopic systems
  - Observe fluid configurations within geologic materials such as rock (sub-micron resolution)
- Projected rate of data generation from light sources could reach 1 petabyte per hour by 2020
- Light source intensity outpacing Moore's law

Digital Rock Physics

- Complex 3D microstructure of rocks directly observed (nanometer to micrometer length scale)

- Fluid movement within these spaces determines transport at larger scales (meter to kilometer)

- Physics-based modeling to predict movement of fluids within microstructure



Application Parallelism

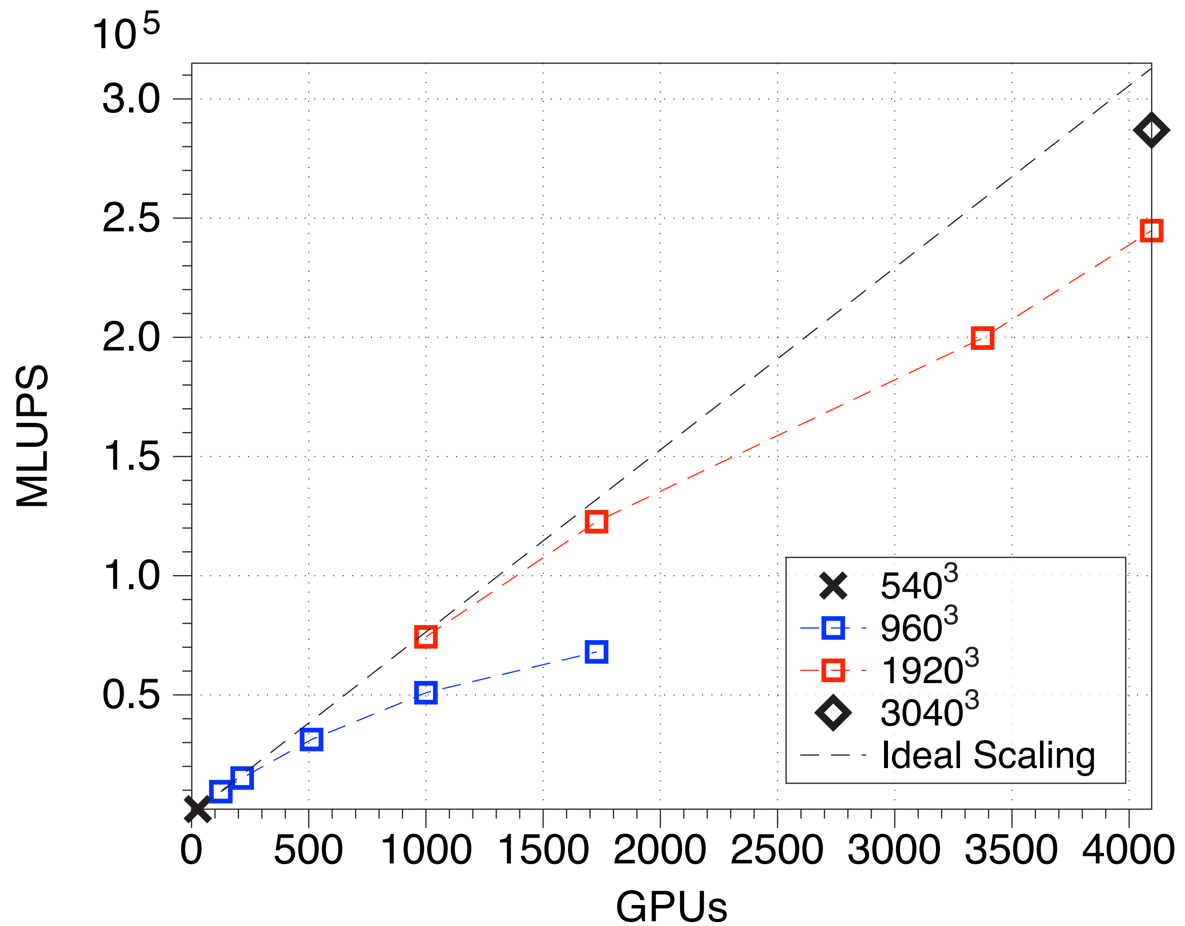
- Large 3D images (1000^3 - 2000^3 voxels)
  - 5 minutes of scan time required to produce image
  - hundreds to thousands of compute nodes per image
- Large number of simulations to deliver scientific value
  - Time sequences from synchrotron (many images)
  - Multiple cases per image

Lattice Boltzmann Methods

- Lattice Boltzmann methods (LBMs) have been devised to model a wide range of transport phenomena
  - Single phase flow, transport of multiple species
  - Two and three-phase flow
  - Complex boundaries can be accommodated easily
- Calculations of the LBM are typically local
  - Scale very well in distributed memory
  - Can be implemented efficiently on GPU and other architectures that rely on SIMD

Multiphase Simulation

- Lattice Boltzmann method: order of magnitude speedup using GPU
- Domain decomposition to distribute across nodes
- One MPI task per node

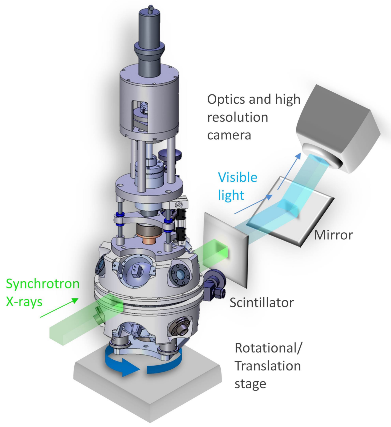


Parallel performance in million-lattice updates per second (MLUPS) for multiphase lattice Boltzmann simulator on Titan (90% parallel efficiency on 4,096 compute nodes)

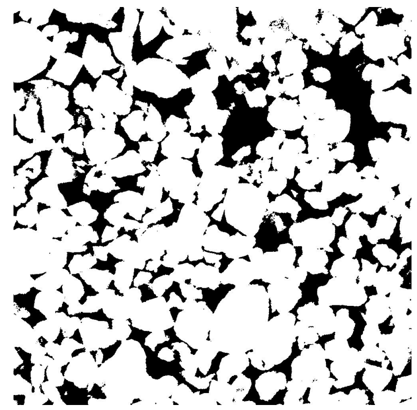
McClure et al. (2014) doi: 10.1109/ IPDPS.2014.67

Workflow

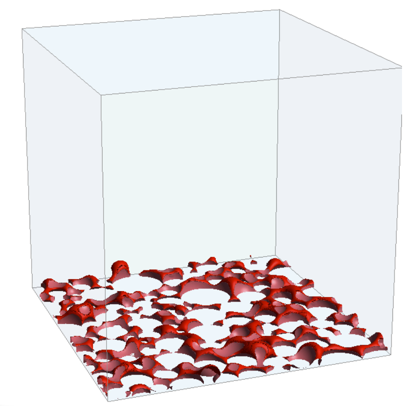
Data Acquisition



Pre-processing



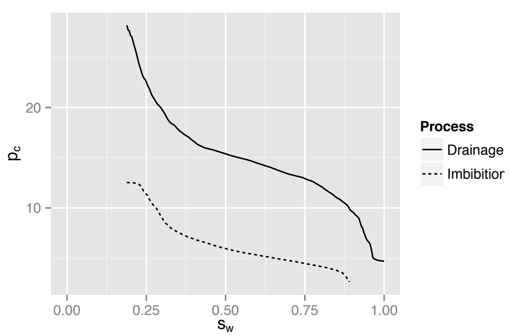
Simulation

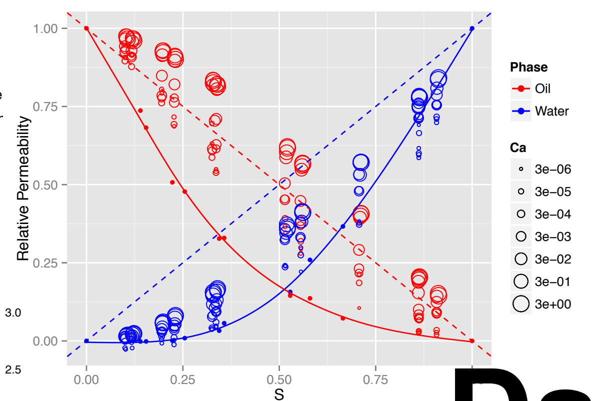


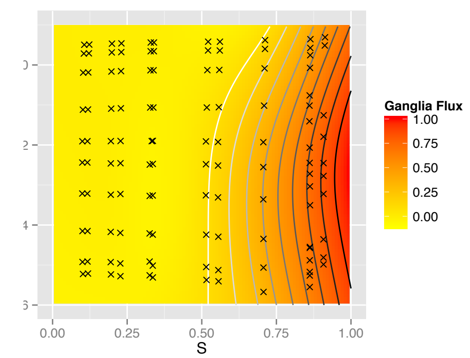
Data Analysis & Post-processing

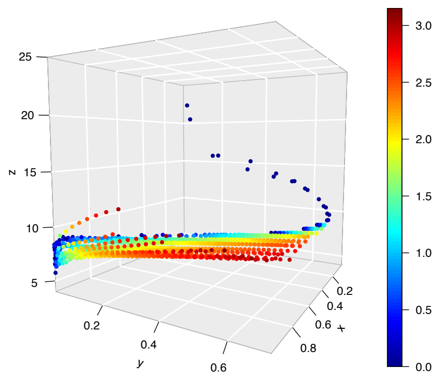
Raw data fields from simulation saved to disk

Data Exploration



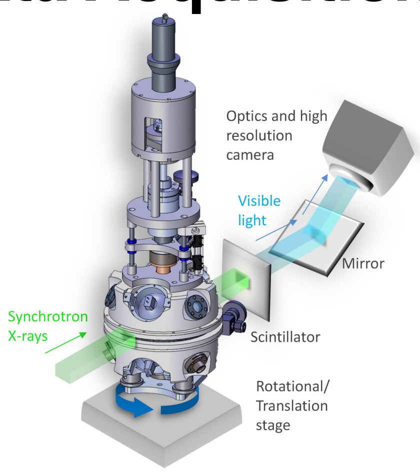




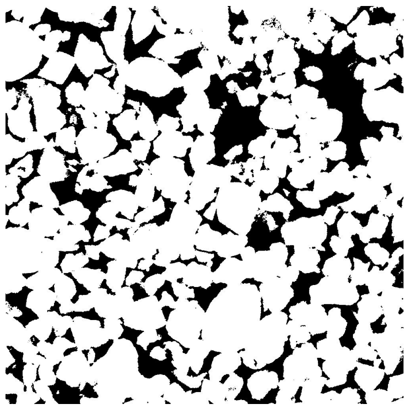


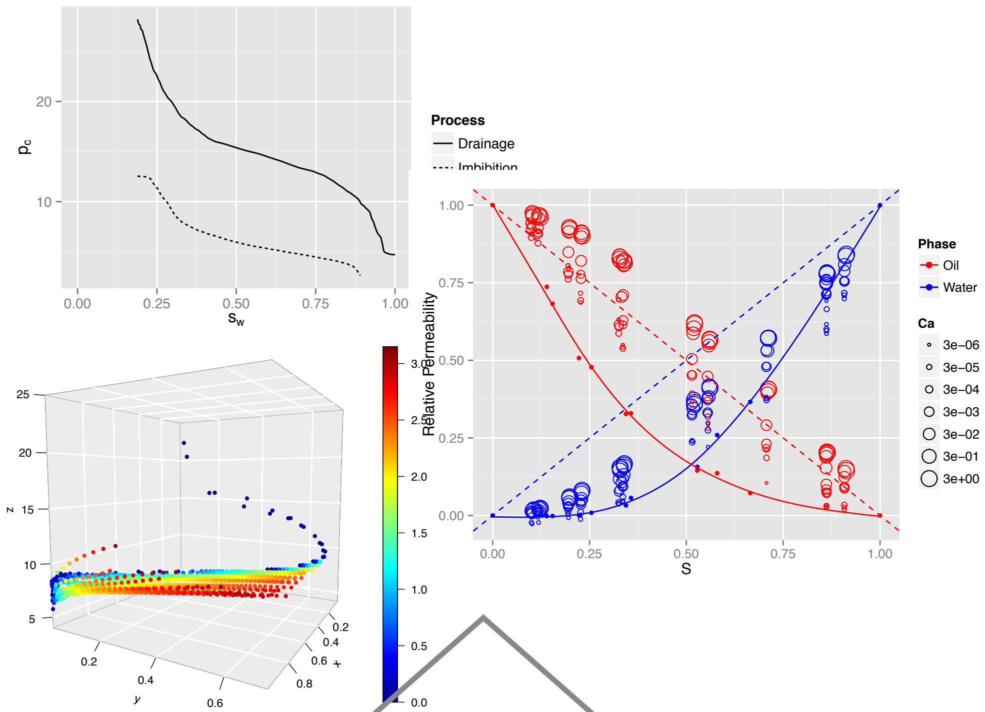
Workflow Summary

Data Acquisition



Pre-processing (inverse modeling)

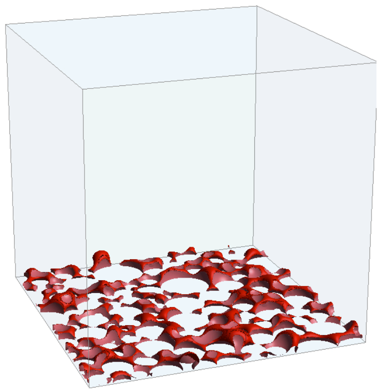




Simulator

Physics Modeling

Multiscale Analysis



Reduced Representation of Simulation Data

Order of magnitude smaller than raw data fields

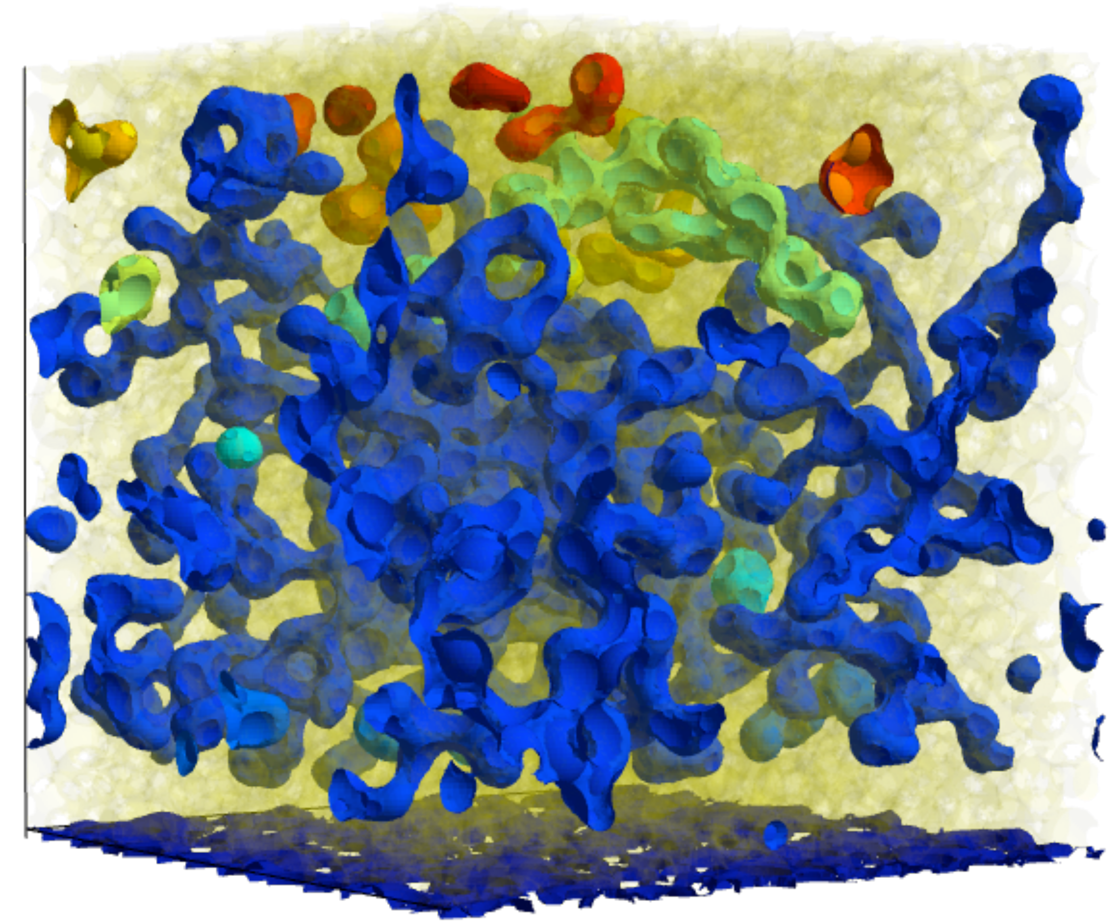
Workflow Summary

- Homogenization theory: develop multiscale relationships for partial differential equations that describe transport phenomena
- Integral geometry: approaches to statistically characterize the geometry / topology
- Percolation theory: understand the role of connectivity for transport in complex systems

Data Reduction

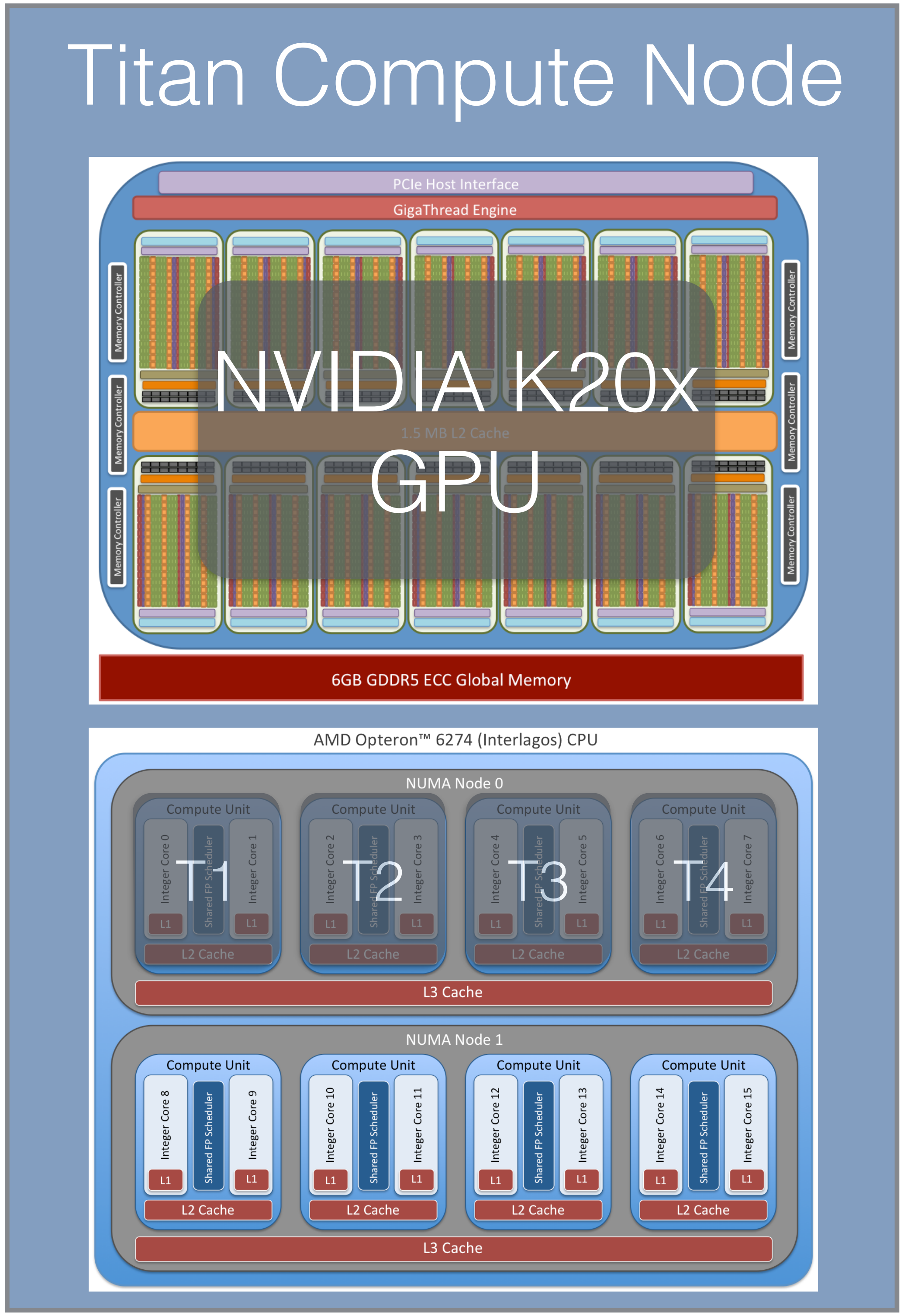
Image Processing

Graph Problems



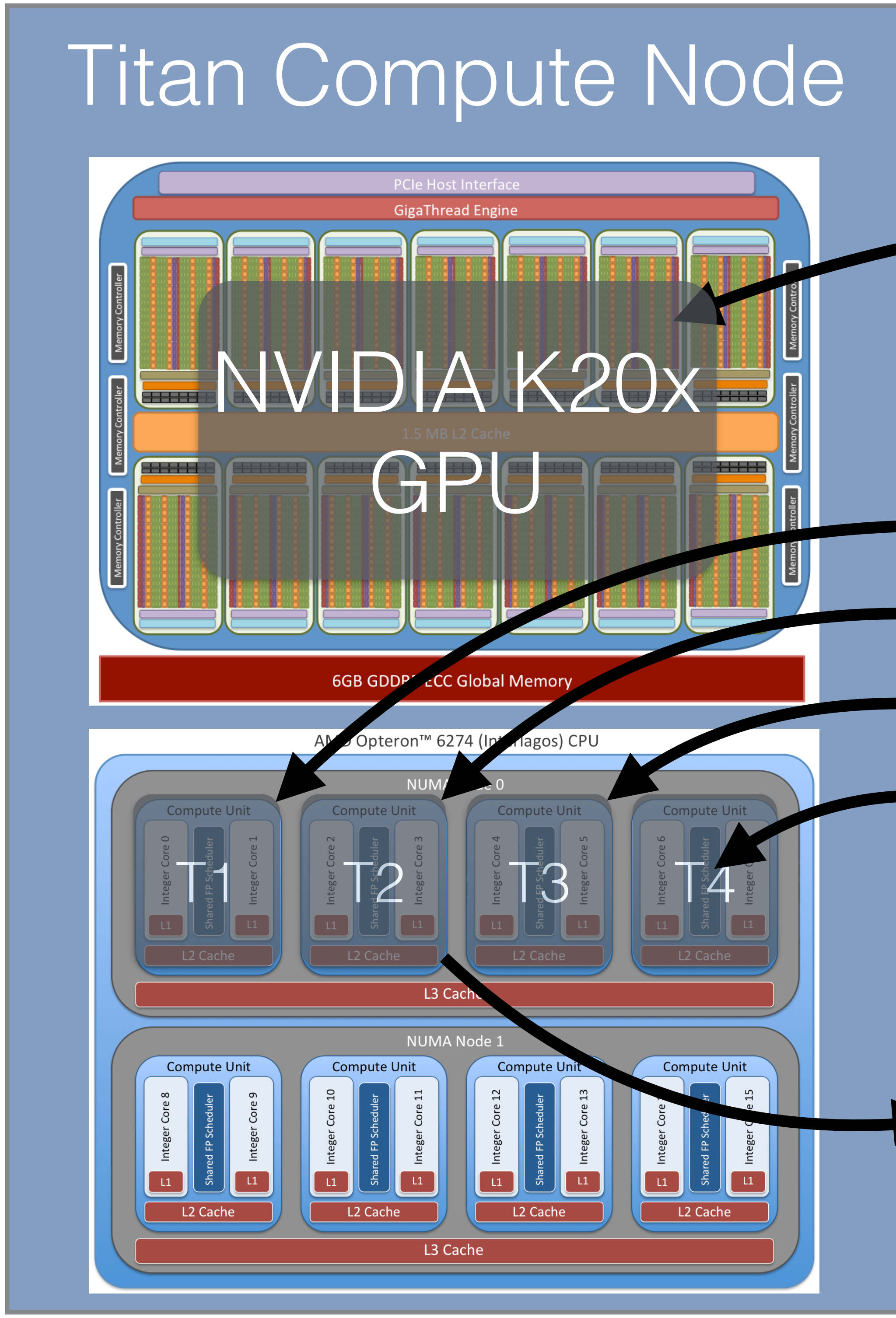
Connected components analysis applied to identify connected portions of a phase within simulation

Intranode Task Management



- Every task performed within a dedicated MPI communicator
- Data movement between CPU and GPU controlled explicitly
- C++11 threads used to spawn analysis tasks
- Task dependencies are incorporated into threadpool class

Intranode Parallelism



Multiphase simulation

- Lattice Boltzmann

Analysis

- Connected components
- Extract interfaces
- Topological analysis
- Solve $\frac{\partial \phi}{\partial t} = \mathrm{sign}(\phi)\,(1 - |\nabla \phi|)$

Data reduction

- Average quantities

Conclusions

- Analyzing the simulation state *in situ* allows us to extract an order of magnitude more information tracking the system behavior

- Heterogeneous compute node is advantageous for our complex workload which can be decomposed using task parallelism

- Even data-driven workloads can be compute bound! It depends on the questions we ask and how smart we are with the data

Acknowledgements

This work was supported by National Science Foundation Grant 0941235,Department of Energy Grant DE-SC0002163, and Army Research Office Grant W911NF-14-1-0287. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. This research used resources of the Oak Ridge Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC05-00OR22725.