Present and Future Computing Requirements for **Direct Numerical Simulation of the Poisson-Nernst-Planck Equation in Clay**

Carl I. Steefel[1]
David Trebotich[2]
1.Earth Sciences Division, Lawrence Berkeley National Lab
2. Computational Research Division, Lawrence Berkeley National Lab

NERSC BES Requirements for 2017
October 8-9, 2013
Gaithersburg, MD

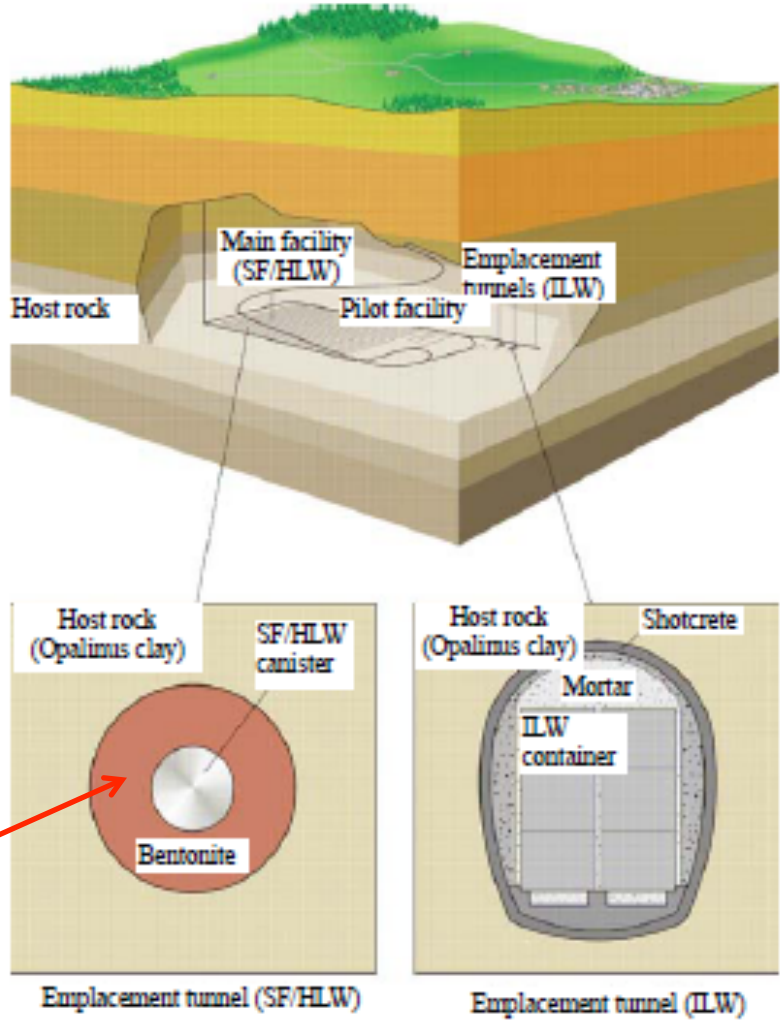
Computational Problem

- Nuclear waste repositories use clay backfill and clay-rich rock for radionuclide attenuation
- Simulate the transport of charged species through compacted clay and clay-rich rock
  - Resolution of nanoscale electrical double layers adjacent to charged clay particles
  - Centimeter to meter scale transport distances
  - Multicomponent system with consideration of inner sphere sorption and ion pairing
- Can it be done at the continuum scale??

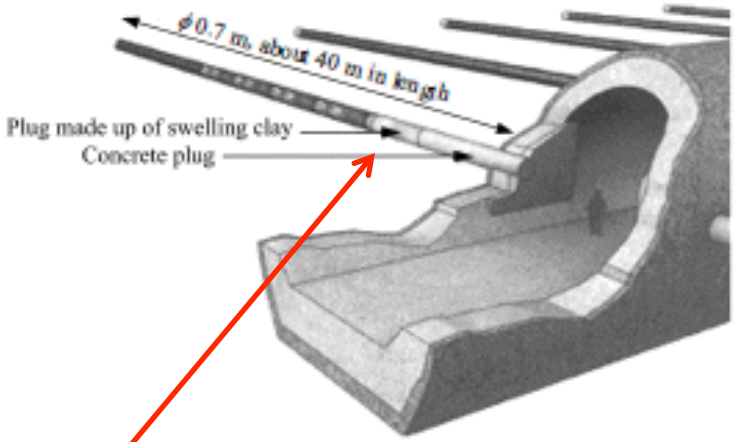
Schematics of Geological Nuclear Waste Repositories

Clay-rich rock as part of Natural Barrier System

Swiss Repository Schematic



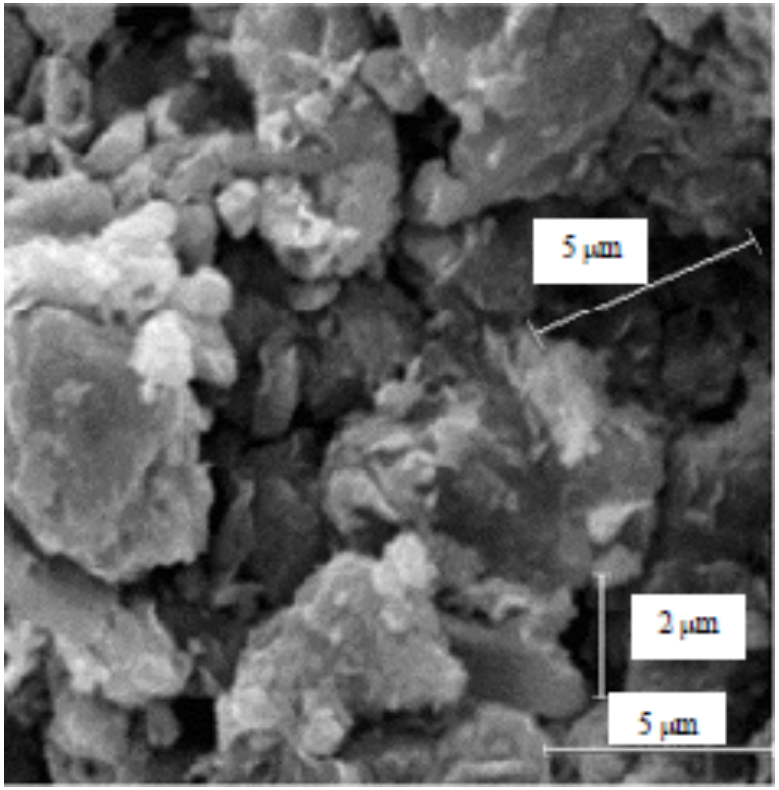
French Repository Schematic



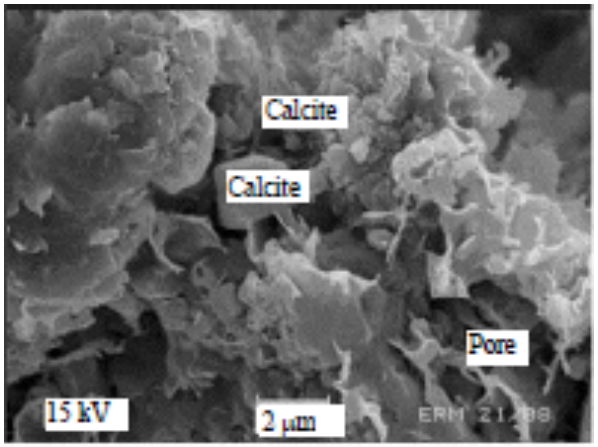
Swelling clay backfill

Clay and Clay-Rich Rocks

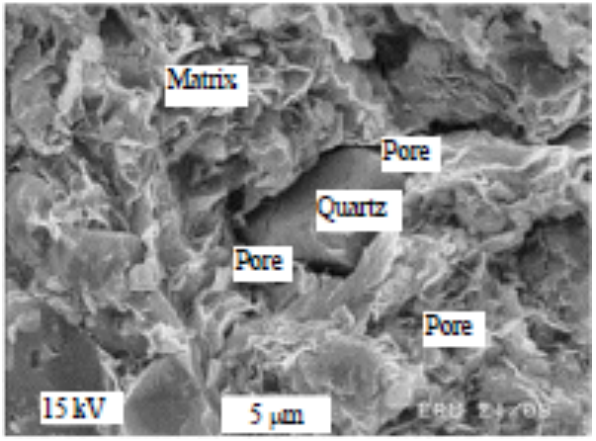
Compacted Clay (Bentonite) for Backfill



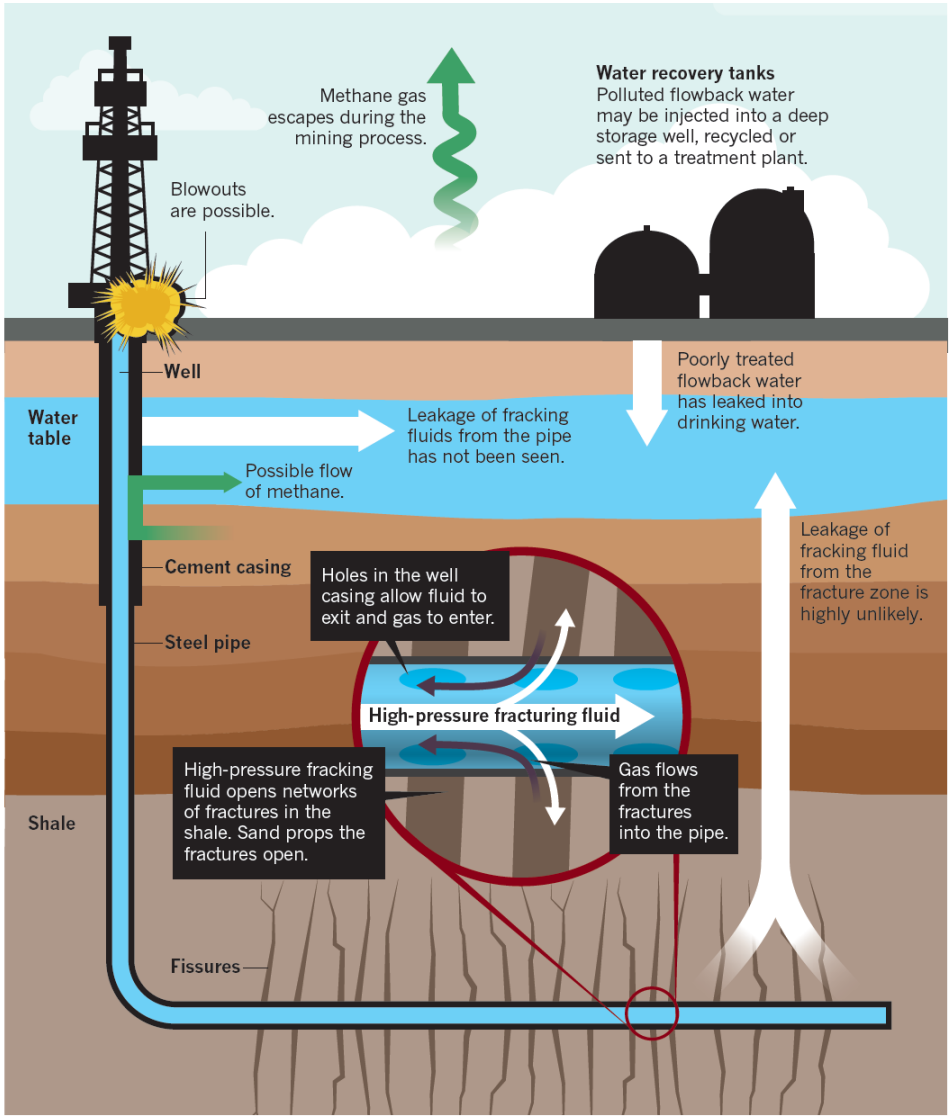
Clay-rich Rock at French Repository



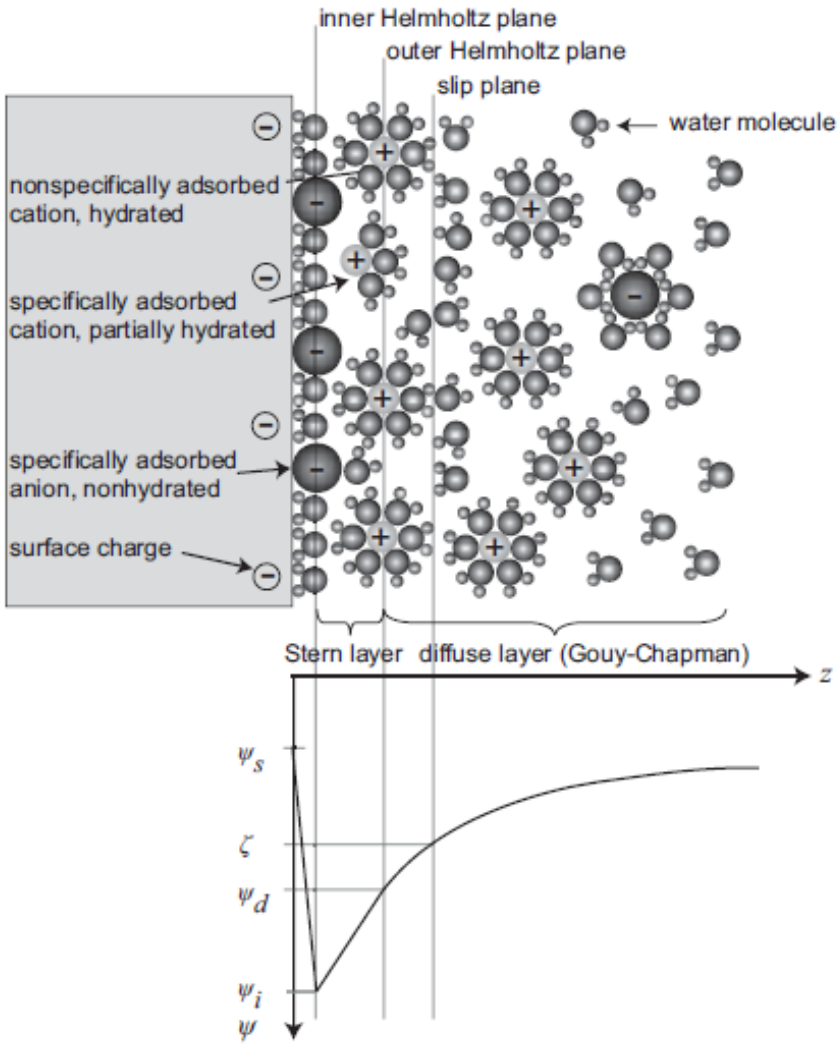
(a)



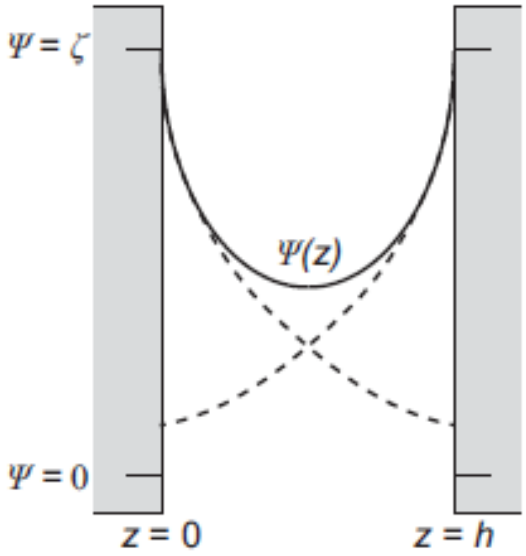
Application to Shale Gas



Schematics of Electrical Double Layer



Double Layers between Charged Surfaces



Solid line: Overlapping double layers
Dashed line: No opposite wall present

After Schoch et al, 2008

Mean Electrostatic (or Donnan) Potential Approach to Transport in Compacted Clay

Concentration in diffuse layer (DL) related to concentrations in bulk water (B) through the mean electrical potential, φ_m, of the diffuse layer

$$C_i^{DL} = C_i^B \exp\left(\frac{-z_i e \varphi_m}{k_B T}\right)$$

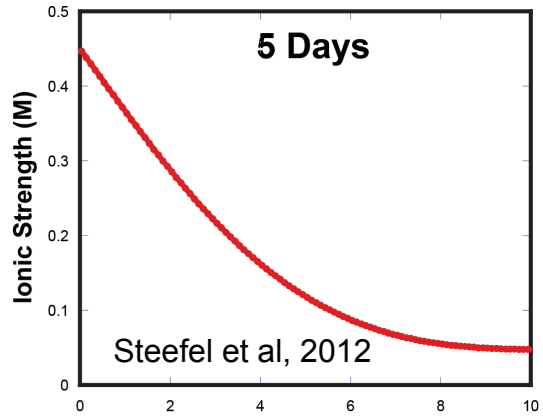
$$\phi^{DL} \sum_i z_i C_i^{DL} = Q^{SL} = \sum_k^{Ns} z_k \Gamma_k$$

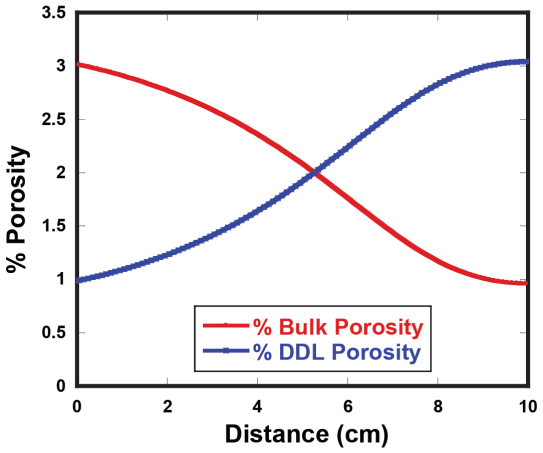
Charge Balance Equation

Dynamic model for electrical double layer thickness as a function of ionic strength

$$\frac{\partial\left[\phi^B C_i^B + \phi^{EDL} C_i^{EDL}\right]}{\partial t} = \frac{\partial\left[\phi^B C_i^B + \left(A_{clay}\lambda_{DL}\beta_{DL}/\sqrt{I}\right) C_i^{EDL}\right]}{\partial t}$$

Higher ionic strength front diffuses through clay, reducing EDL porosity, increasing bulk porosity





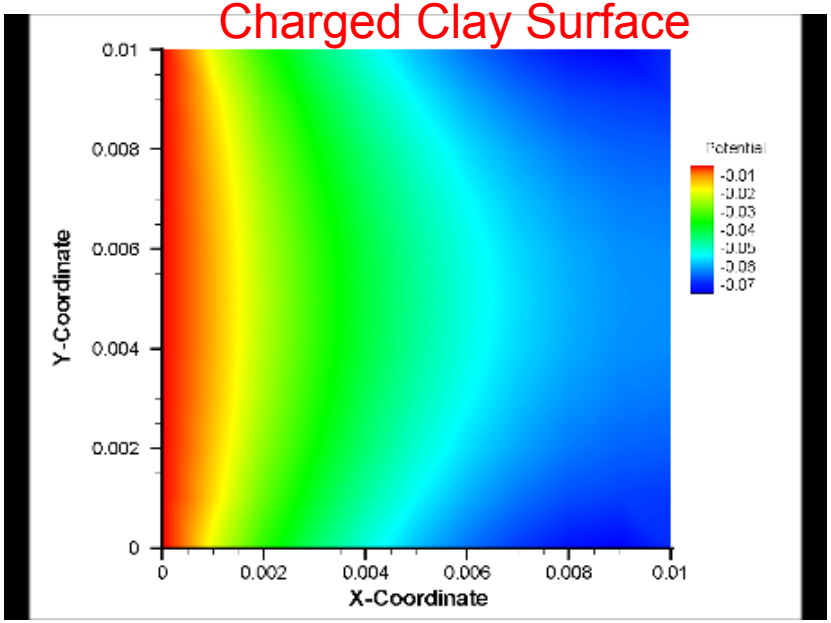
Poisson-Nernst-Planck Simulation of Transport through Compacted Clay

Improved description of transport and electrical double layer is offered by solving the full Poisson-Boltzmann equation

$$\frac{\partial^2 \varphi}{\partial x^2} = \frac{-e}{\varepsilon} \sum_i z_i C_i \exp\left(\frac{-z_i e \varphi(x)}{k_B T} \right)$$

Two-dimensional cation transport between charged clay surfaces simulated with Poisson-Boltzmann Equation, with longitudinal transport using the Nernst-Planck Equation

Galindez & Steefel, in prep.



Electric Potential Poisson Solver for Electrokinetic Flow

Prototype solver in Chombo for electric potential to model electrokinetic effects in microfluidic devices

$$\begin{aligned} \vec{u}_t + (\vec{u} \bullet \nabla)\vec{u} &= -\frac{1}{\rho}\nabla p + \nu \triangle \vec{u} \\ \nabla \bullet \vec{u} &= 0 \\ \triangle \phi &= 0. \end{aligned}$$

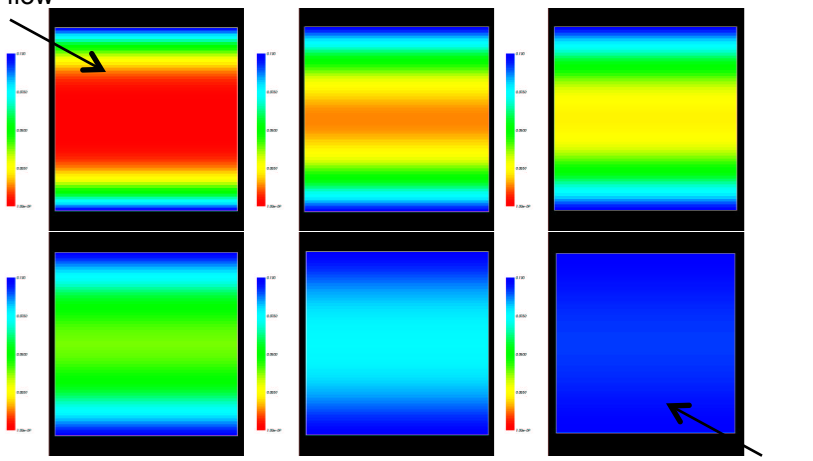
Smulochowski formula for slip velocity boundary condition

$$\vec{u}_{slip} = b\vec{E}_{||}$$

➔Slip boundary condition formulated for embedded boundary methods

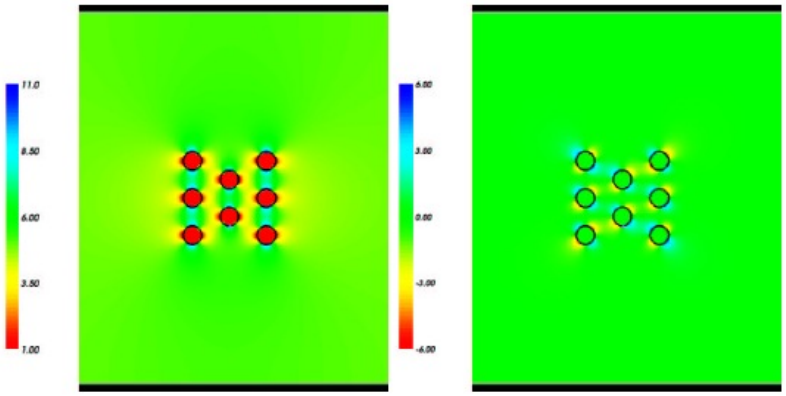
Viscous flow driven electrokinetically to steady-state plug flow by dropping instantaneous voltage across channel

Poiseuille flow



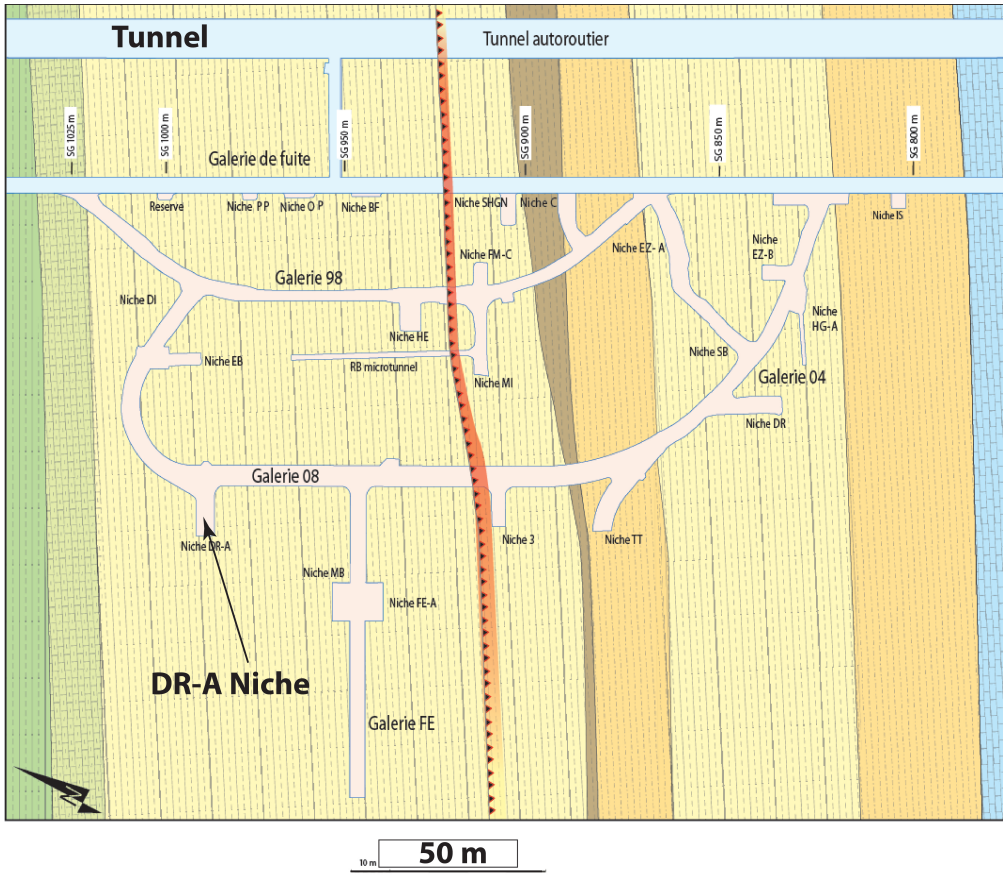
Plug flow

Components of electric field in electrokinetic flow past microarray



Chu and Trebotich, 2006

Mont Terri DR-A Test
Opalinus Clay, Switzerland



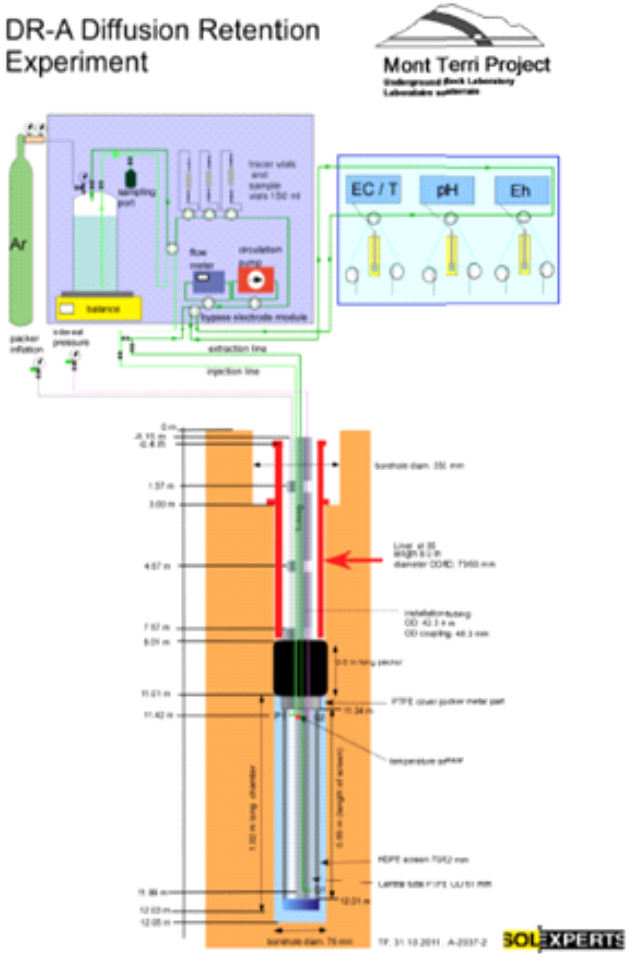
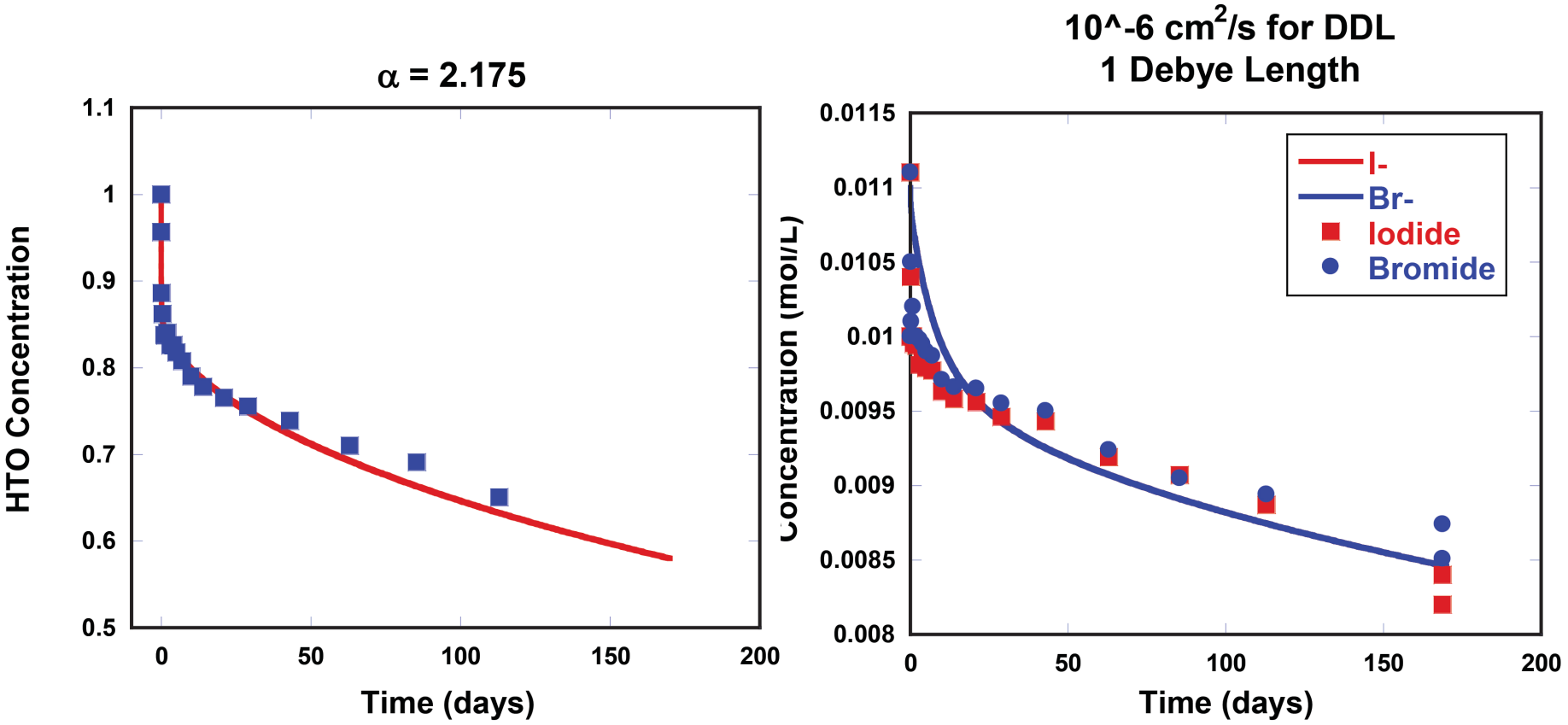


Figure 7: Schematic drawing of the entire test setup

Simulation of Anion Exclusion

Uncharged Tracer (HTO)

Anions (Negative Charge)

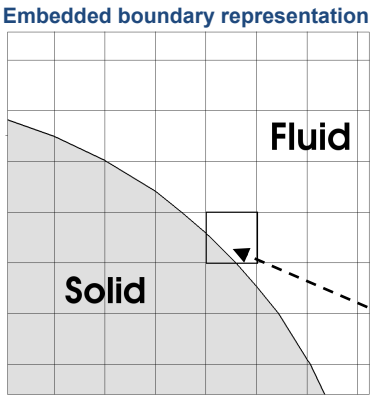


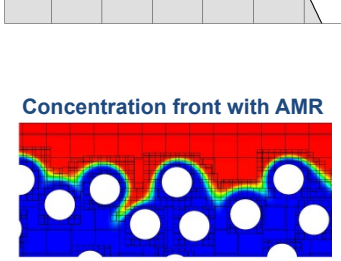
Chombo-Crunch is a high performance microscale flow and reactive transport solver

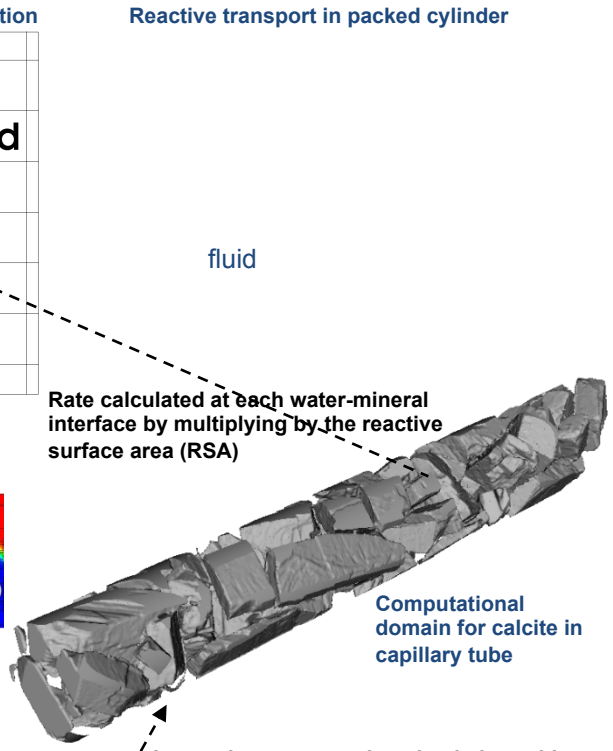
Chombo-Crunch

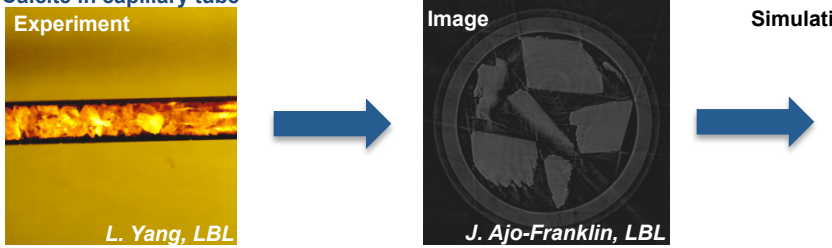
CFD + multi-component geochemical reactive transport in very complex pore (micro) scale geometries

- Adaptive, finite volume methods for advection-diffusion in Chombo
  - Accurate reactive surface area using embedded boundaries
  - Dynamic local refinement (AMR)
  - Scalable (100K processors)
  - Direct simulation of image data
- CrunchFlow geochemistry
  - SNIA operator splitting
  - Point-by-point calculation
  - CFL-limited timestep

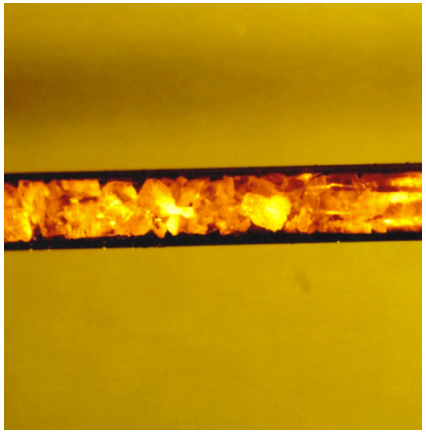






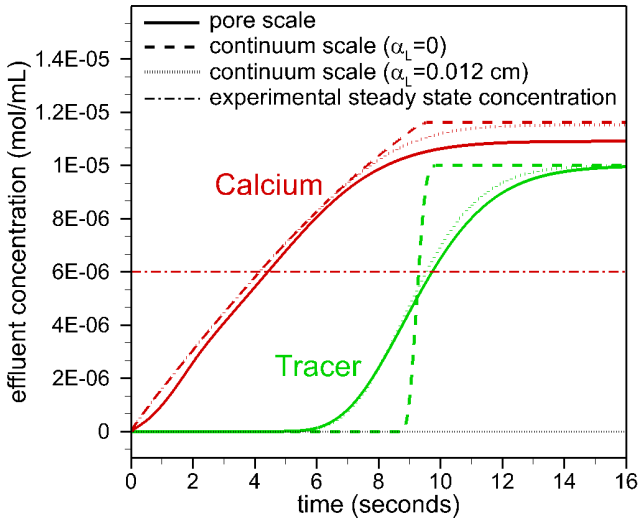


Science at scale for CO_2 sequestration 2012-13 NISE award



EFRC-NCGC validation experiment

- Reactive transport in capillary tube packed with calcite
  - 7 mm long x 500 μm diameter
  - 0.899 μm image resolution
  - 9.3 seconds residence time (steady-state)
- Chombo-Crunch simulation
  - 1.6B simulation grid points
  - 1.19 μm resolution
  - 700GB plot files, 600GB checkpoint files
  - 49,152 cores on Hopper Cray XE6
  - 100 time steps per hour
  - 1 time step = 0.000182 seconds real time



➔ Current rate law in reactive transport model overshoots steady-state effluent concentration

➔ Run coarser simulation (2μm) with new rate law, cf. Edison…

Preliminary Chombo-Crunch runs on Cray XC30 outperform XE6 by 2.3x overall run time

Edison

2 μm simulation resolution

6144 cores for 6 hours

1366 time steps

1 time step = 0.000265 seconds of simulated time

XC30 Performance

- 15.3 seconds per time step
  - 3.92 time steps per minute
  - 230.77 time steps per hour
- Lustre IO = 2.3% of total run time
  - 1.8% checkpoint files
  - 0.5% plot files
- 1% initialization overhead

Hopper

2 μm simulation resolution

6144 cores for 36 hours

3567 time steps

1 time step = 0.000265 seconds of simulated time

XE6 Performance

- 35.5 seconds per time step
  - 1.69 time steps per minute
  - 100 time steps per hour
- Lustre IO = 1.8% of total run time
  - 1.1% checkpoint files
  - 0.7% plot files
- < 1% initialization overhead

➔ Chombo-Crunch is memory bandwidth limited

Chombo Parallel I/O and Storage

- Current Chombo simulations run on up to 100,000 cores w/ O(1TB) timestep dumps
  - Striping across all OSTs on Lustre → I/O < 2% of run cycle
- Future simulations will utilize up to 500,000 cores w/ O(10TB) dumps
- Need I/O bandwidth and hardware resources to minimize time spent in writing data
- Need production stack (HDF5, MPI-IO, Lustre) to scale on NERSC platforms
  - Currently working w/ ExaHDF5 project

What is the effect of heterogeneous pore structure on average reaction rates?

Porosity = 0.762

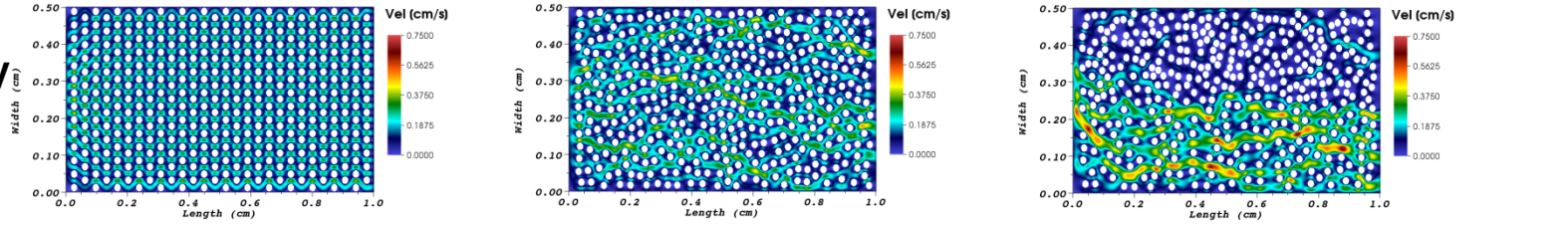
A^{tot} = 4750 m^2 m^{-3}

Regular packing ***Random packing*** ***2-Zone random packing***

Velocity

inlet

0.1 cm/s



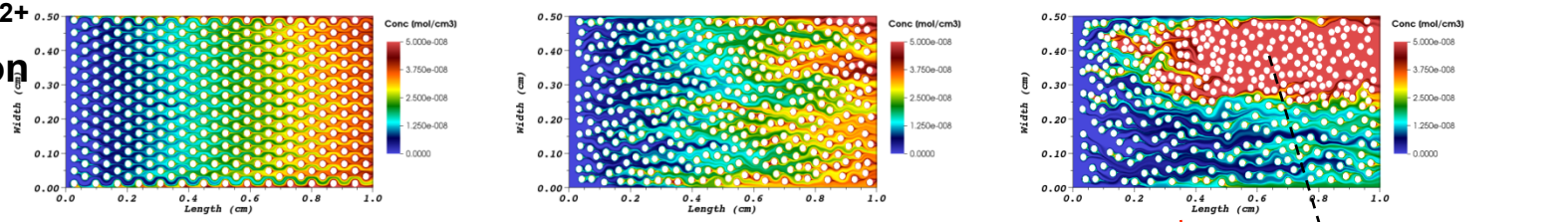
Ca^{2+} Concentration

inlet

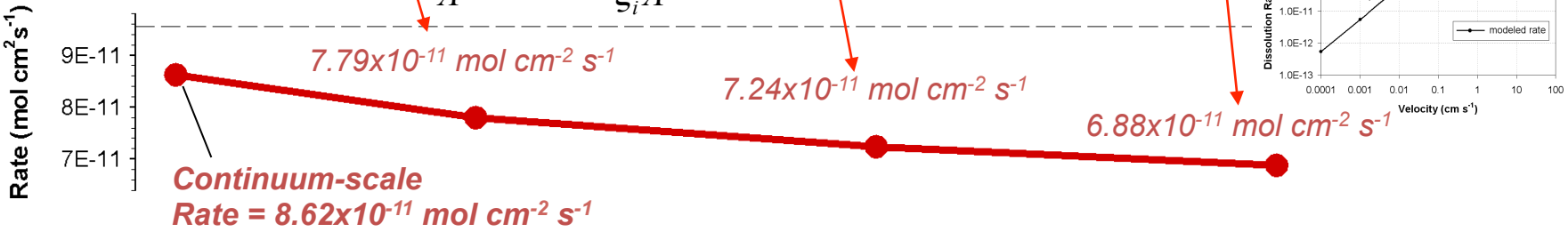
pH 5

0.01 NaCl

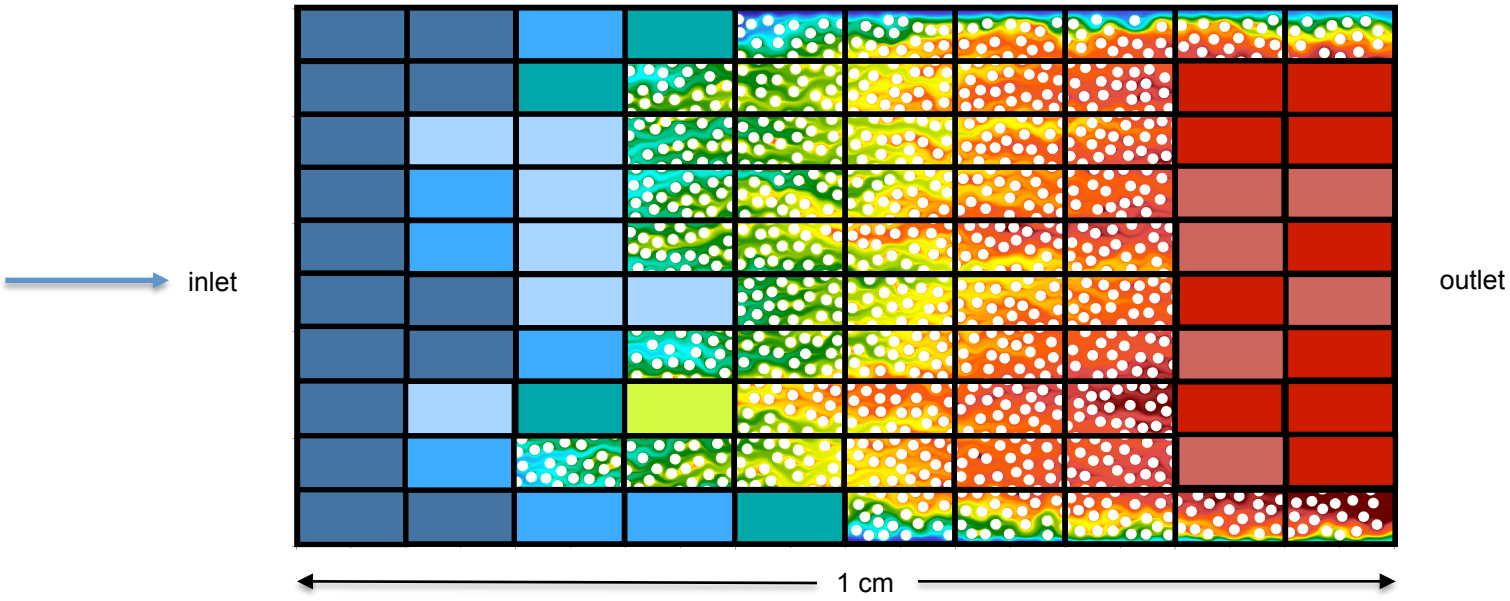
pCO_2 = 3.15 10^{-4} bar



$$\frac{R}{A^{tot}} = \frac{\left(C_i^{outlet} - C_i^{inlet}\right)Q^{tot}}{\xi_i A^{tot}}$$



Volume averaging alone of pore scale does not preserve heterogeneity



$$k = \frac{Q^{tot}\,\mu\,L}{A^{channel}\left(p^{outlet} - p^{inlet}\right)}$$

$$\frac{R}{A^{tot}} = \frac{\left(C_i^{outlet} - C_i^{inlet}\right)Q^{tot}}{\xi_i A^{tot}}$$

We are investigating multiscale methods that preserve heterogeneity of pore scale

Multiscale approach: Use first-of-its-kind highly resolved pore scale simulation data

(i) to inform better parameterizations of permeability, reaction rates and dispersion at continuum Darcy scale, and

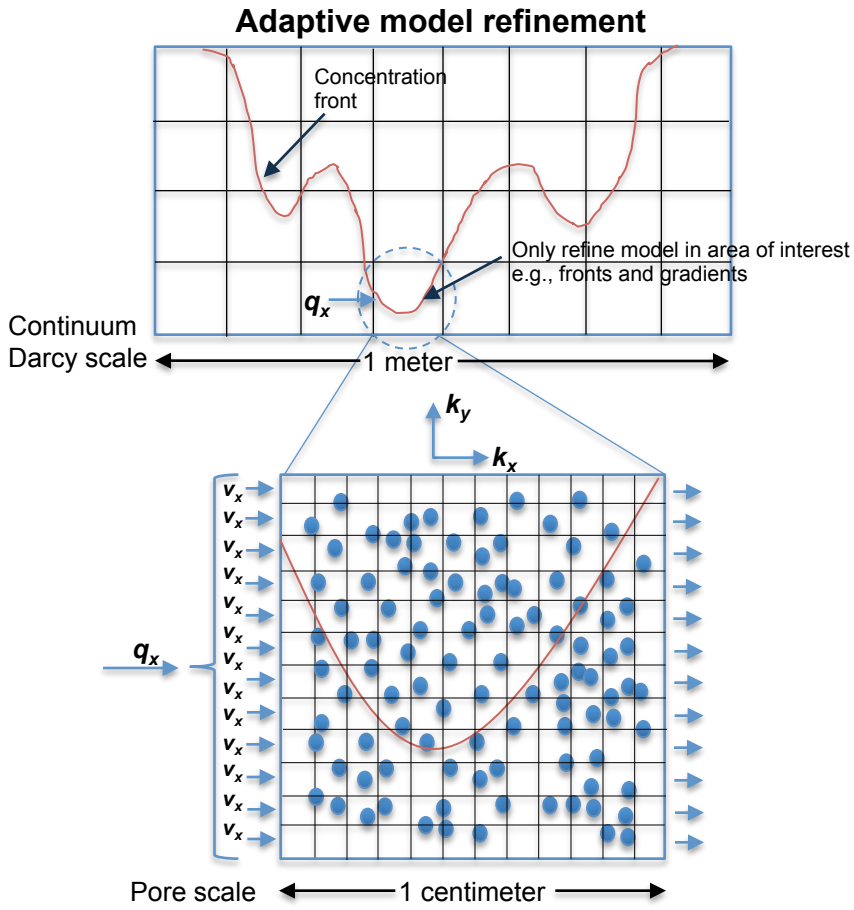
(ii) to verify multiscale approaches:

Deterministic method:

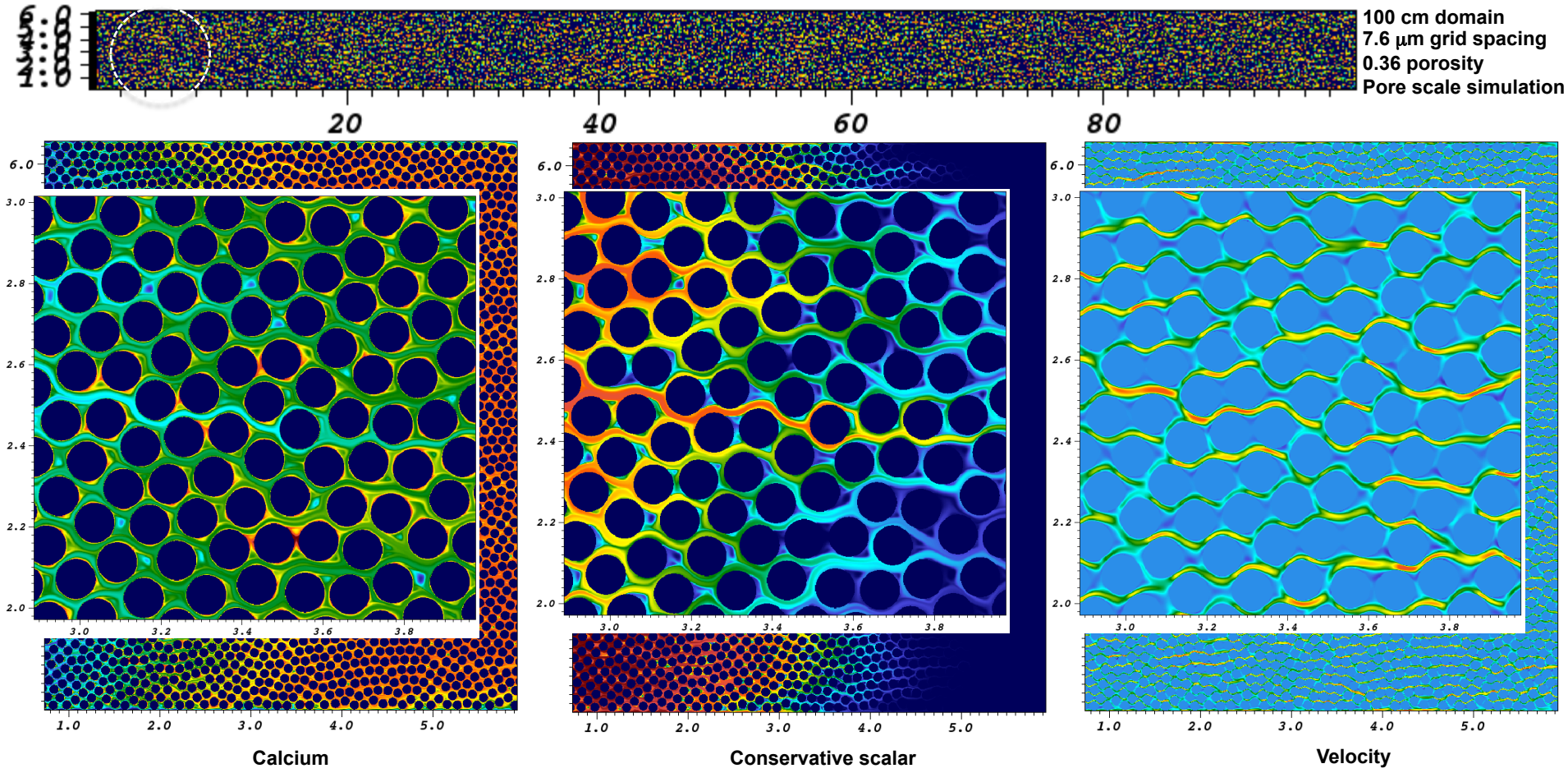
- Adaptive model refinement: directly upscale pore scale data to continuum scale *locally* in areas of interest in domain

Stochastic method:

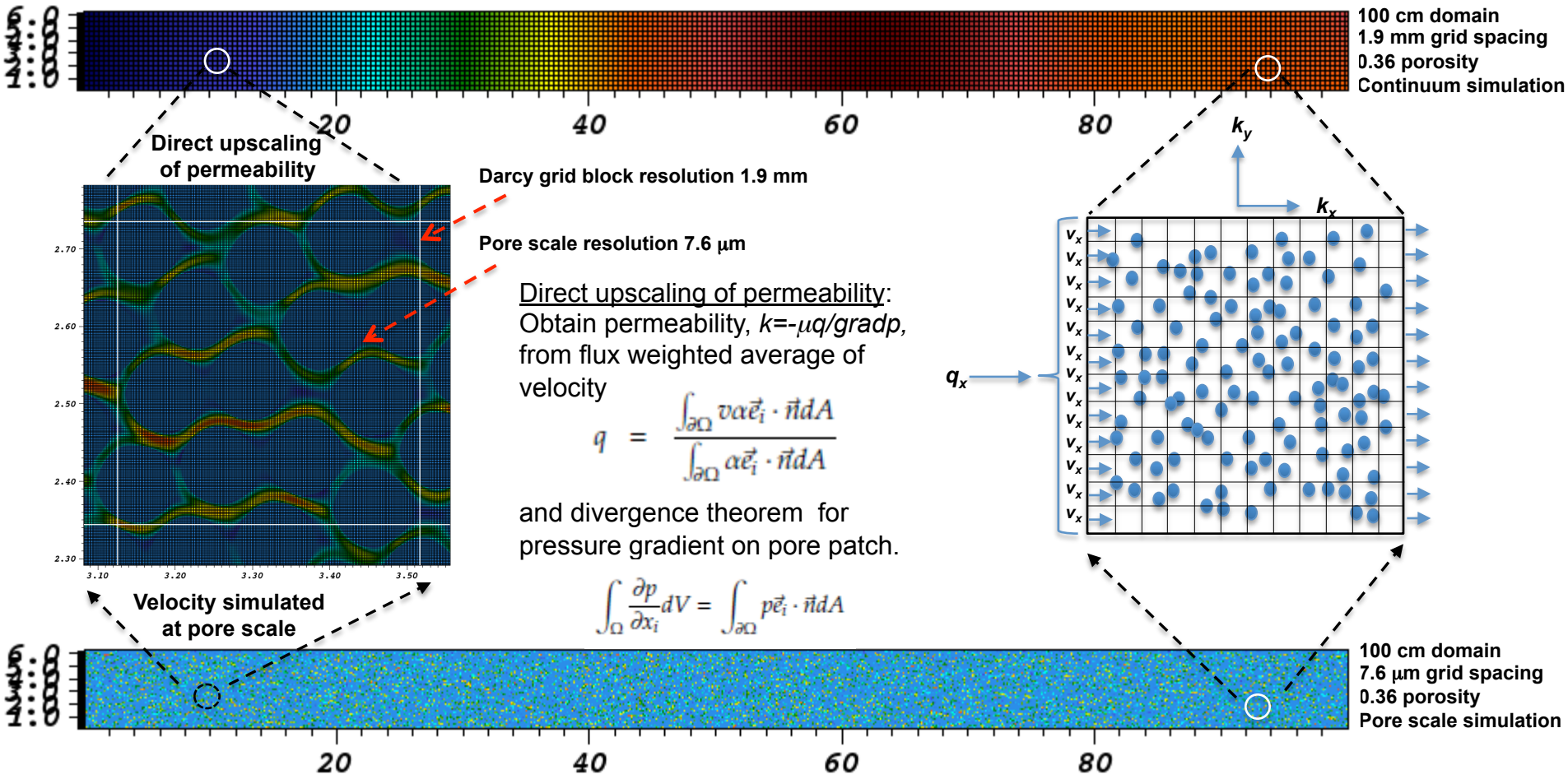
- Intermediate pdf: characterize pore space with a probability density function that is fitted to the pore scale data and graph connectivity



We have established basis for upscaling with high resolution pore scale models



We have constructed algorithms to do direct upscaling from pore to continuum



Direct upscaling of permeability: Obtain permeability, $k=-\mu q/gradp$, from flux weighted average of velocity

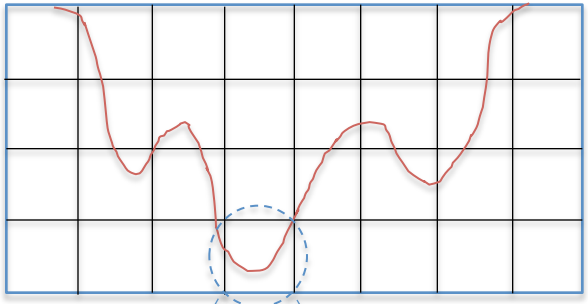
$$q = \frac{\int_{\partial\Omega} v\alpha\vec{e}_i \cdot \vec{n}dA}{\int_{\partial\Omega} \alpha\vec{e}_i \cdot \vec{n}dA}$$

and divergence theorem for pressure gradient on pore patch.

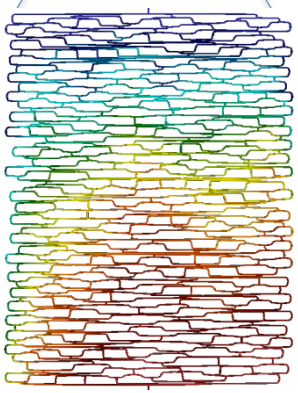
$$\int_{\Omega} \frac{\partial p}{\partial x_i} dV = \int_{\partial\Omega} p\vec{e}_i \cdot \vec{n}dA$$

We can use graph theory to statistically characterize heterogeneous pore space

Continuum Darcy scale



Convert to pore network model (ΔP=RQ) and recouple via Darcy

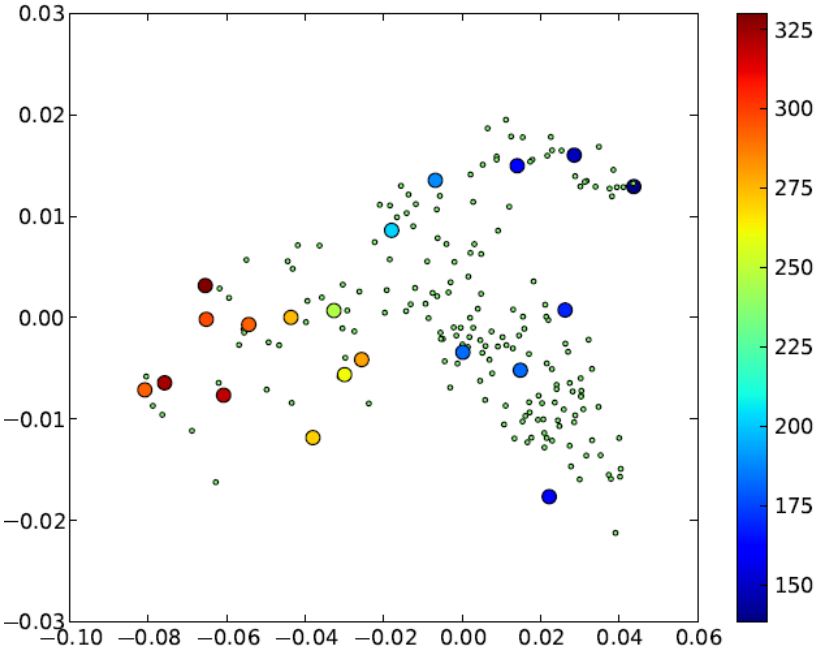


Pore scale simulation with direct upscaling via Darcy

Construct fast pore network solver by using intermediate probability density function to sample characterization of pore space obtained from graph theory and ensemble of pore data

k=-μq/gradp

Single metric derived from graph theory to determine permeability



Principle axes of graph metric using max flow algorithm. 200 packings statistically represented by the metric, tested and compared to 20 simulations (large circles) with dimensionless permeability in color.

1. Project Description

Carl I. Steefel, David Trebotich, Ian Bourg, Ben Gilbert
(Lawrence Berkeley National Laboratory)

- Primary objective is to develop a nano-continuum Poisson-Boltzmann (PNP) description of the electrical double layer (EDL) in clay (used as backfill for geological nuclear waste repositories) and clay-rich rock (shale) based on a general multicomponent chemistry framework.
- Present focus is to compare continuum representations of PNP and Mean Electrostatic Model for clays with molecular dynamics simulations and experiments.
- By 2017 we expect to be able to solve for lateral migration of ions at the cm scale through charge, compacted clay in 2D and 3D, while incorporating Stern Layer sorption and ion pairing effects.

2a. Computational Strategies

- Computational approach is for Direct Numerical Simulation of CFD of compacted clays and clay-rich rock at the nano- to pore scale, with treatment of the full Poisson-Boltzmann and Nernst-Planck equations.
- The codes we use are Chombo and CrunchEDL.
- Chombo uses Embedded Boundary Methods (EB) and Adaptive Mesh Refinement (AMR), and will use a new elliptic solver for electrical potential. CrunchEDL is a Jacobian-based (bio)geochemical module coupled to electrical potential.

2b. Computational Strategies

- Biggest computational challenge is to resolve the nanoscale of the EDL bordering charged clay particles while considering longer space and time scales for ion transport.
- Our parallel scaling is limited by current algorithms for CFD that resolve nanoscale to microscale fluid-solid (mineral) interfaces.
- We expect our computational approach to change by 2017 so as to capture the multi-scale character of the clay and clay-rich rock pore structure. Approach based on either multiple length scales in Chombo, or a hybrid multiscale approach coupling to molecular dynamics.

3. Current HPC Usage (see slide notes)

- Machines currently used: Edison, Hopper
- Hours used in 2012-2013:
  - 75M at NERSC in 2013 (charged for 32M)
  - 80M awarded at ALCF (Mira) as part of ALCC award
- Typical parallel concurrency and run time, number of runs per year
  - 2 micron resolution of calcite: 6K cores, 36 hour production runs, 16 runs
  - 1 micron resolution of calcite: 48K cores, 36 hour production runs, 32 runs
  - 4 micron resolution of packed spheres: 64K cores, 36 hour production runs, 16 runs
- Data read/written per run: 80 TB (2 micron), 200 TB (1 micron), 200 TB(4 micron)
- Memory used per (node | core | globally): 0.5 GB per core
- Necessary software, services or infrastructure: Chombo, PETSc, VisIt, Lustre filesystem
- Data resources used (/scratch,HPSS, NERSC Global File System, etc.) and amount of data stored: /scratch, HPSS == 500 TB

4. HPC Requirements for 2017

(Key point is to directly link NERSC requirements to science goals)

- Compute hours needed (in units of Hopper hours): 500M to 1B
- Changes to parallel concurrency, run time, number of runs per year: factor of 2 in resolution for any of our production runs is going to require 8x the cores and twice the time to access experimental time scales
- Changes to data read/written: 10 TB timestep dumps
- Changes to memory needed per (core | node | globally): 1 GB per core
- Changes to necessary software, services or infrastructure: hybrid parallel model--Fat MPI nodes and dynamic OpenMP threading. Thread safe Chombo4 alerady makes use of the hybrid model.

Computational Challenge

Year	Problem dimension	Domain size	Grid points	Spatial resolution	Problem time scale	HPC resources	Problem DOF	Data storage per plot file
2013	2D	1m	2B	O(1 mm)	O(hours)	50M hrs	25	1TB
	3D	1cm	2B	O(1 mm)	O(minutes)	100M hrs	25	1TB
2017	2D	10cm	20B	O(1 nm)	O(hours)	500M hrs	5	10TB
	3D	1cm	20B	O(1 nm)	O(minutes)	1B hrs	5	10TB

5. Summary

- What new science results might be afforded by improvements in NERSC computing hardware, software and services?
  - We have been the first to resolve microscale behavior in reactive transport using NERSC and BES resources. We would be able to resolve mesoscale behavior of reactive transport processes in the subsurface, and in particular, charged clays and shales.
- Recommendations on NERSC architecture, system configuration and the associated service requirements needed for your science
  - We would like a NERSC machine configured like Edison but with 8x the cores.

- NERSC generally refreshes systems to provide on average a 2X performance increase every year. What significant scientific progress could you achieve over the next 5 years with access to 32X your current NERSC allocation?
  - We would be able to perform direct numerical upscaling of mesoscale and microscale processes to larger continuum scale for more accurate field scale models.
- What "expanded HPC resources" are important for your project?
  - More cores, more time, priority use, more storage
- General discussion