Coupling of Models for Energy and Environment
(NPACI Alpha Project)
Review Meeting

Participants

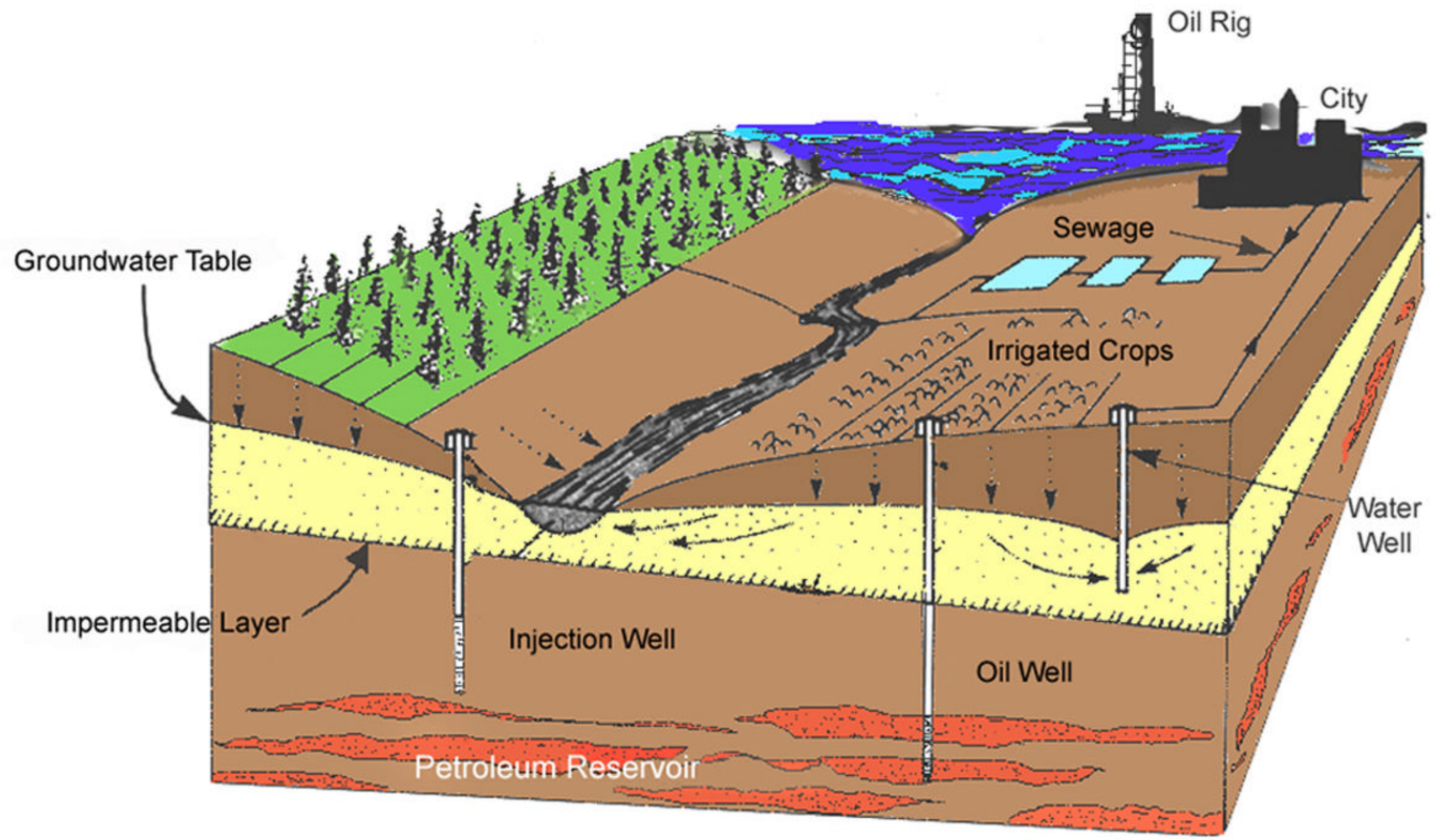
- Project Leaders: Joel Saltz, Mary Wheeler
- Project Manager: Joel Saltz
- University of Maryland: Joel Saltz, Alan Sussman, Tahsin Kurc
  - ADR, Meta-Chaos
- University of Texas: Mary Wheeler, Malgo Peszynska, Jichun Li, Shuyu Sun
  - Surface, ground water, and petroleum reservoir models
- University of California at San Diego: Scott Baden
  - KeLP
- USC/Information Sciences Institute: Carl Kesselman
  - Globus

Motivation and Challenges

- Cost-effective contaminant remediation, minimizing cost of cleanup
  - Defining a strategy as soon as possible
  - Efficient coordination and management of possible strategies
- Implementing effective oil and gas production
  - Optimizing well placement
  - Efficient exploration of possible production strategies

- Environments with poorly known properties
- Complex chemical and physical interactions
- Multiple physics, Multiple scales

Multiscale Fluid Flow for Energy and Environment Applications



Goals of the Alpha Project

- **Improving our ability to model the flow of contaminants through the ecosystem**
  - Development of multi-physics, multi-scale models
  - Coupling of ground water and surface water simulations
- **Enabling on-demand simulation, exploration, and comparison of multiple scenarios**
  - Integration of a robust, Grid-based computational and data handling infrastructure

Significance of Goals

- Establish prototypes for coupling multi-physics, multi-scale, multi-phase models
- Develop tools that allow users to integrate information obtained from ensembles of multi-physics, multi-resolution simulations
- Enhance the ability of users to develop parallel applications
- Allow users to run application software anywhere and anytime

Database Tools

- Develop database tools that support users' need to synthesize data obtained from selected sets of realizations.
  - Tools will make it possible to drill down
  - Combine data from different realizations
  - Carry out spatial subsetting and processing of data from a given realization
  - Aggregate data to answer scenarios about what would happen in a given portion of a domain under different conditions
  - Identify instances of a given physical phenomena in each realization and aggregate instances of the physical phenomena across realizations

Distributed Computing Tools

- Develop tools to support development of distributed multi-physics applications
  - Tools allow programs to flexibly interoperate.
  - Programs will be directly coupled, or will share very large, disk or memory based multi-resolution datastructures.
  - Each program can invoke procedures that act on shared datastructures to carry out user-specified patterns of spatial query and data aggregation
  - Selected portions of a program's distributed or disk based datastructures are identified, processing is carried out, and the resulting data is copied into another program's distributed or disk based datastructures.
  - Data descriptors based on common standards (e.g. XML DTDs) will be used to describe datatypes and to describe distribution of datasets across memories and/or disks.

Meta Computing Tools

- Develop tools that support efficient integration of information from distributed data collections
  - Widely dispersed collections of datasets.
  - Efficiently identify, query, subset and aggregate data from geographically dispersed data sources
  - Build on top of Grid computing infrastructure such as Globus

Specific Projects to meet the Goals

- Coupled Flow and Reactive Transport Simulation with Parssim using Globus and ADR
- Coupling of Surface Water Codes and Parallelization of selected Ground Water/Projection Codes using KeLP/MetaChaos
- History Matching and Uncertainty with IPARS and ADR

Project 1

Coupled Flow and Reactive Transport Simulation with Parssim using Globus and ADR

- Interactive exploration of various scenarios.
- Transformation of Parssim into a set of computationally intensive flow part and the data intensive part: reactive transport, post-processing.
- These parts will be run at different locations, Globus will provide resource allocation and data exchange between flow and transport codes.
- ADR will be used to store and visualize Parssim output

Parssim
Parallel Subsurface Simulator

- Multicomponent, Multiphase
  - logically rectangular 3D
  - higher order Godunov
  - characteristics-mixed finite element

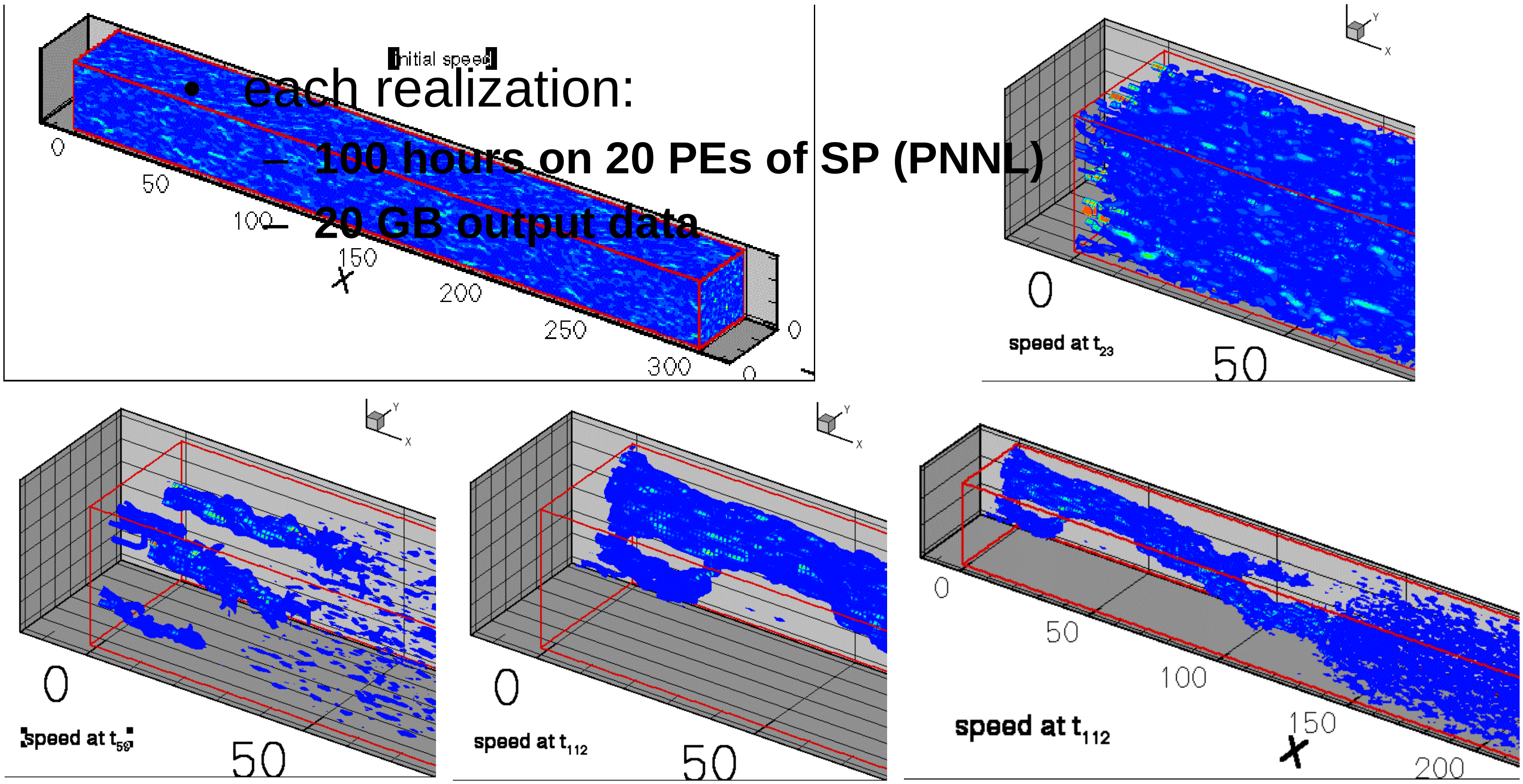
1 flowing phase,

N stationary phases

- General biogeochemistry
  - interior point minimization of free energy
  - explicit integration of kinetics ODEs
- Scalable Parallel
  - domain decomposition (MPI)
  - SP2, cluster of PCs, T3E, Workstations
  - dynamic load balancing

Typical Flow/Reaction Runs with Parssim

- each realization:
  - **100 hours on 20 PEs of SP (PNNL)**
  - **20 GB output data**

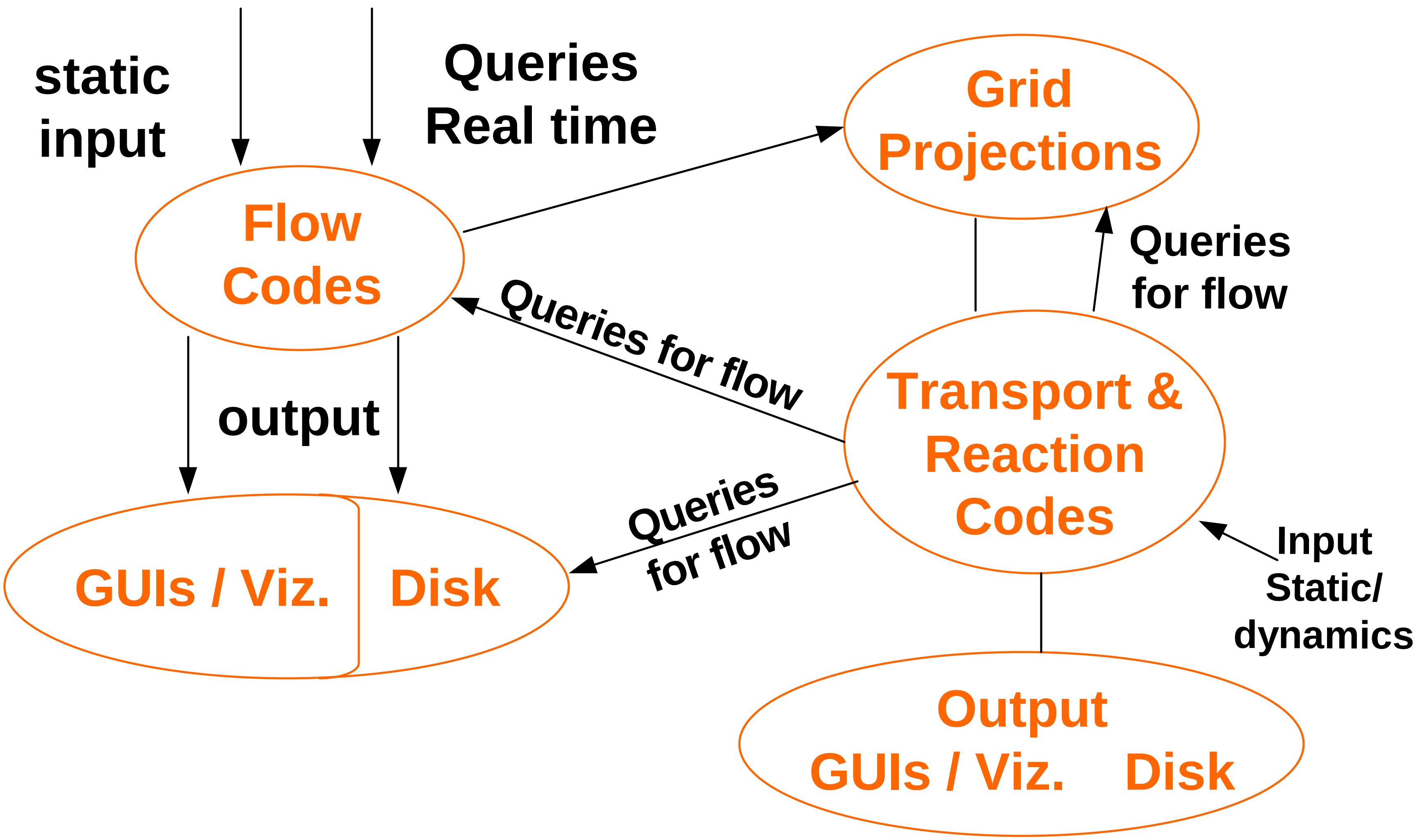


Project 2

Coupling of Surface Water Codes

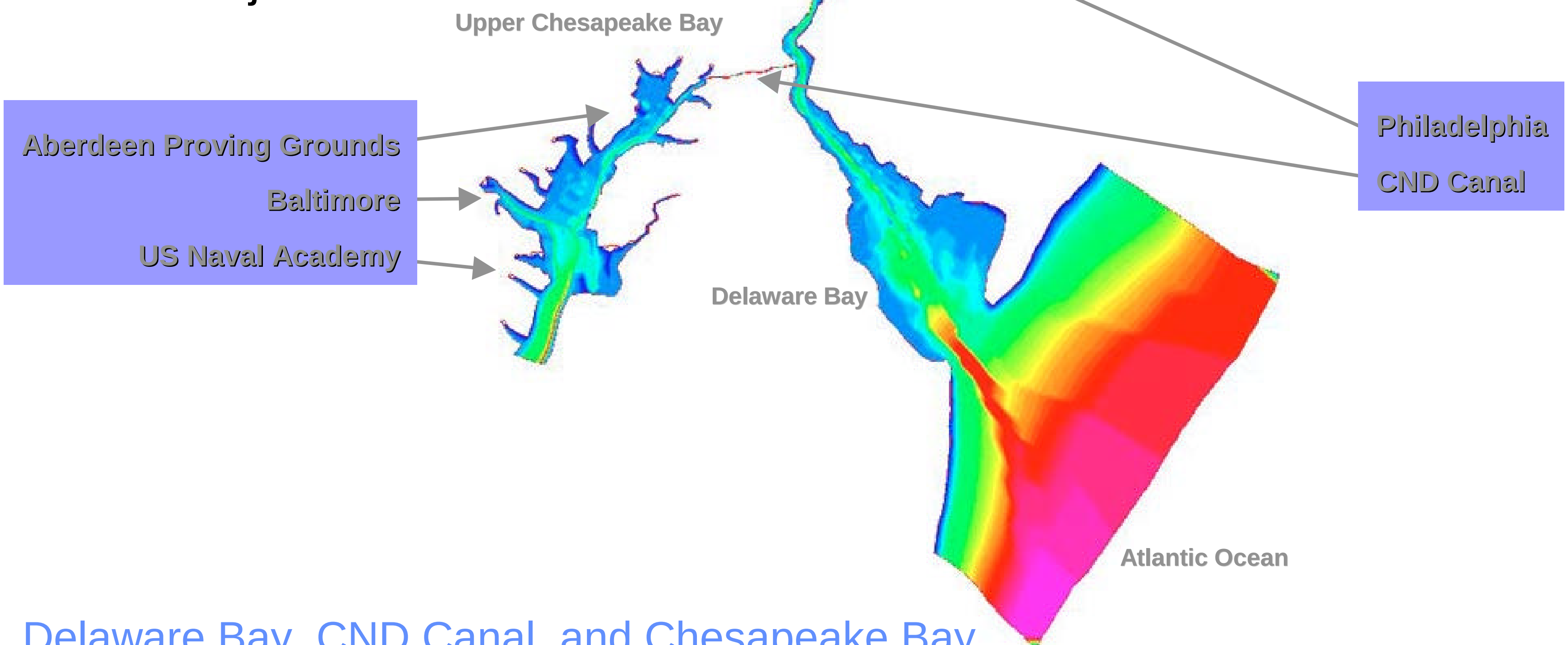
- Carry out a surface water pollution remediation using a chain of flow codes and reactive transport codes.
- Codes will be run on separate platforms and their results will be stored in ADR which will provide the coupling.
- Parallelization of Projection/Ground Water Code using KeLP

A data management framework for environmental modeling



Projection Code: UTPROJ

- Couples 3D surface water flow model to contaminant and salinity transport models, can be used as ground water code
- Implements conservative velocity projection method
- Improves local mass conservation
- Projection formulation based on mixed finite element method



Delaware Bay, CND Canal, and Chesapeake Bay

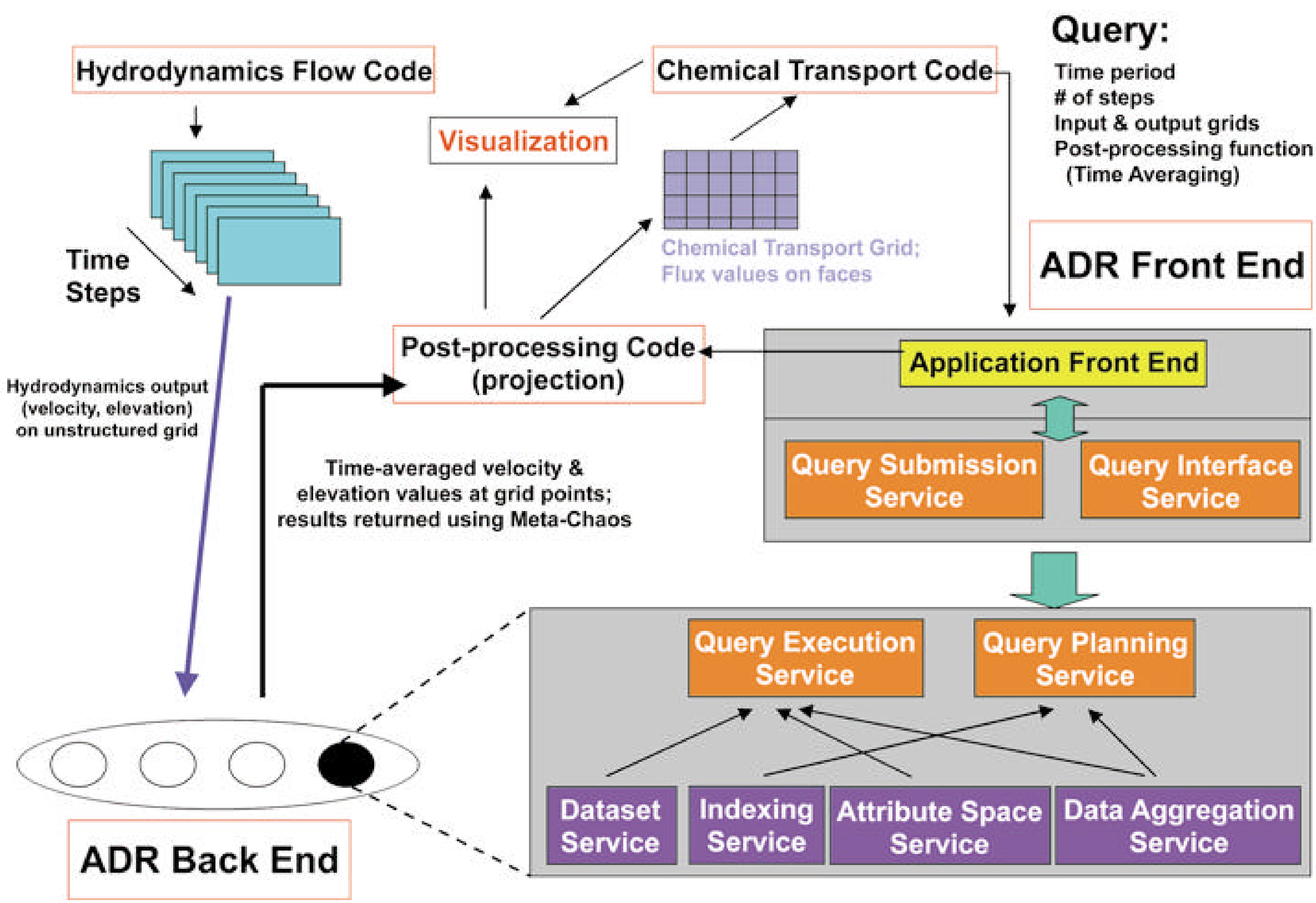
Active Data Repository (ADR)

- Enables integration of storage, retrieval, & processing of multidimensional multiple datasets on parallel machines
- Easy to customize for various applications
- Provides support for common operations (e.g., spatial queries, complex data aggregations, such as transferring data between irregular and multi-resolution datastructures)

- Two major components:
  - Parallel back end: data storage, retrieval, processing
  - Front end: interface between clients and back end nodes

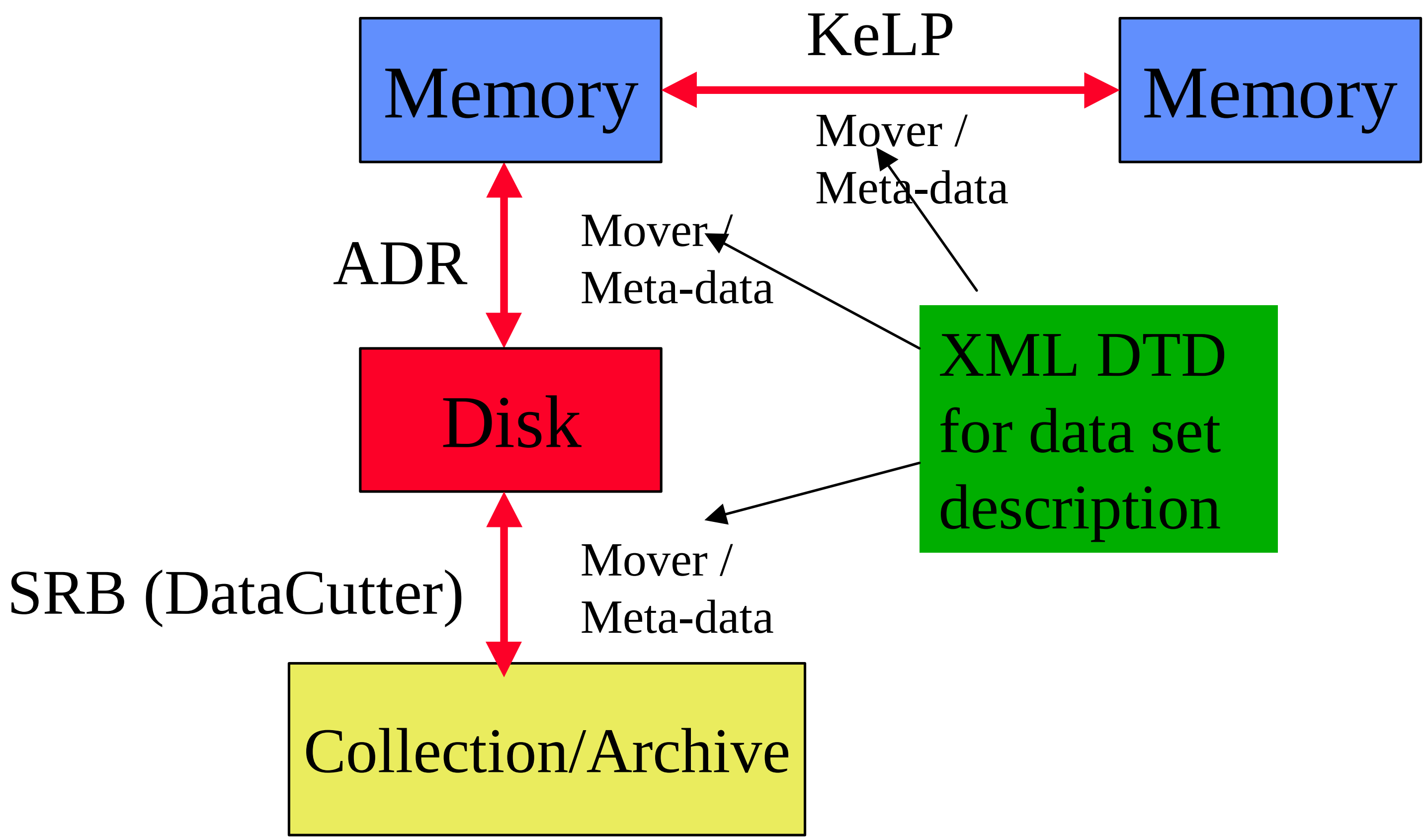
- Developed as set of modular services implemented in C++; customization via inheritance and virtual functions

Active Data Repository



Data Handling Integration

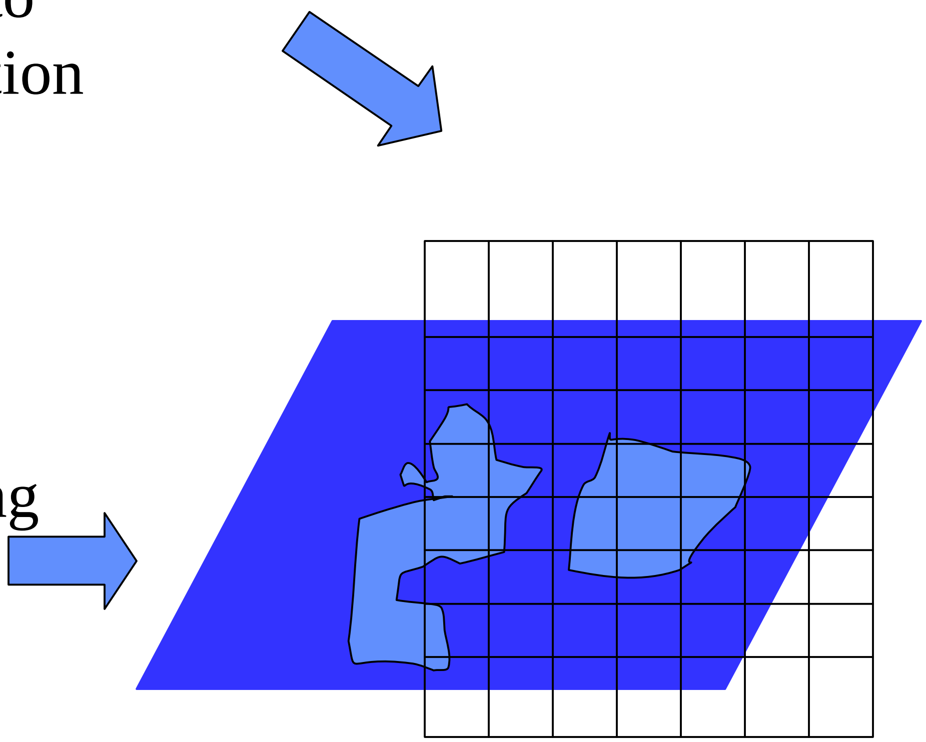
National Partnership for Advanced Computational Infrastructure (NPACI)



Typical Query

Output grid onto which a projection is carried out

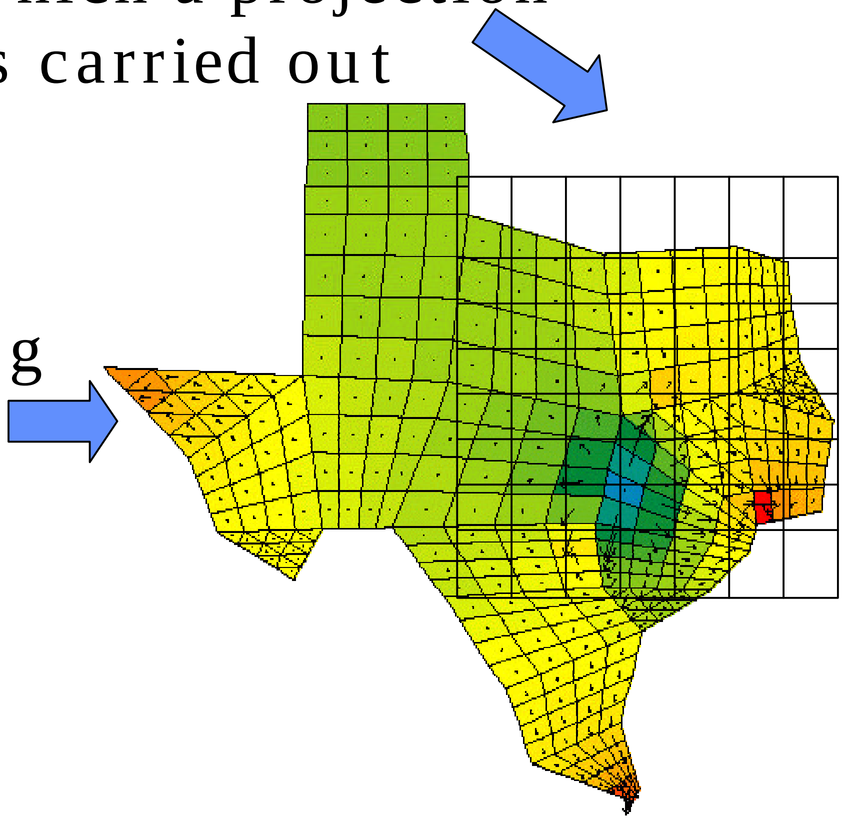
Specify portion of raw sensor data corresponding to some search criterion



Processing Irregular Datasets
Example -- Interpolation

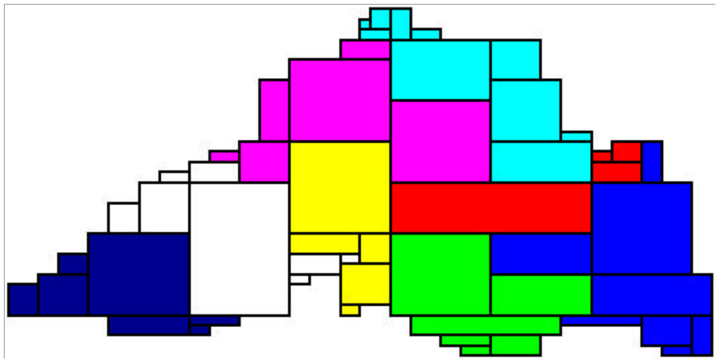
Output grid onto which a projection is carried out

Specify portion of raw sensor data corresponding to some search criterion

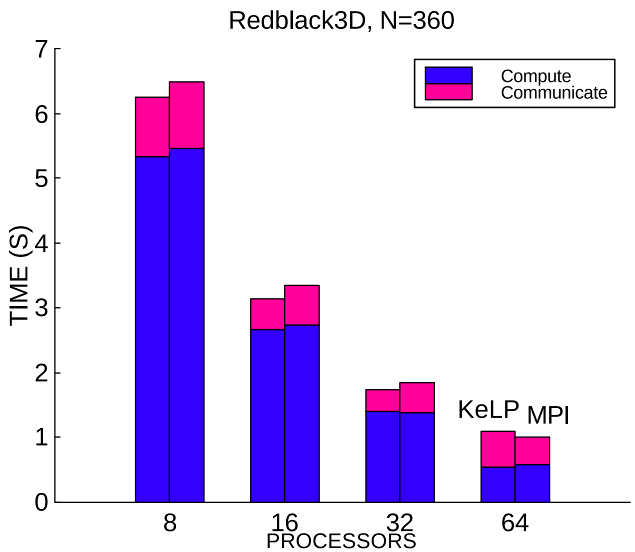


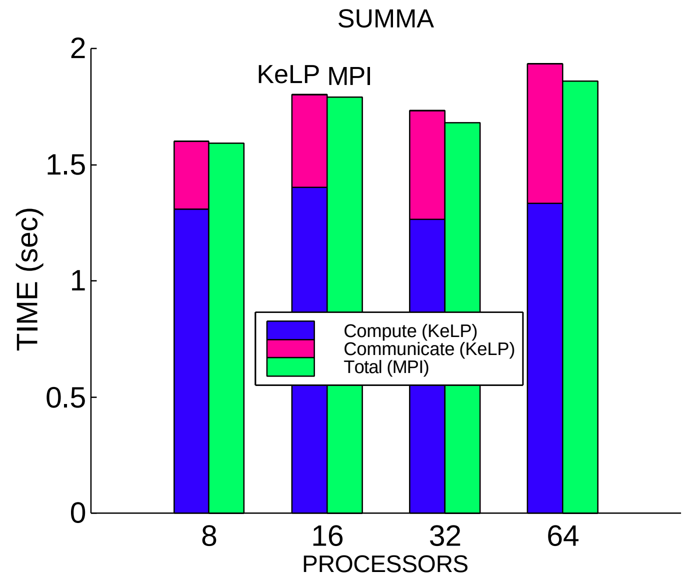
KeLP

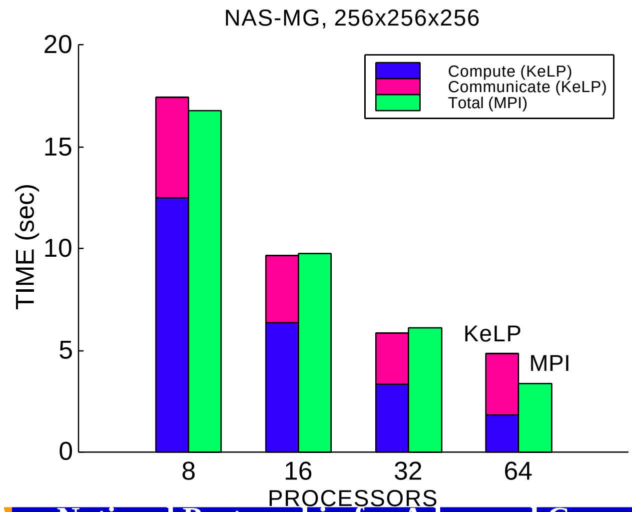
- C++ framework
- Meta-data types to express the structure of an application and the underlying data dependencies
- Spatial regions provide "structural abstraction"
- Employ global knowledge about communication to improve performance at run time
- Customized load balancers tuned to the application

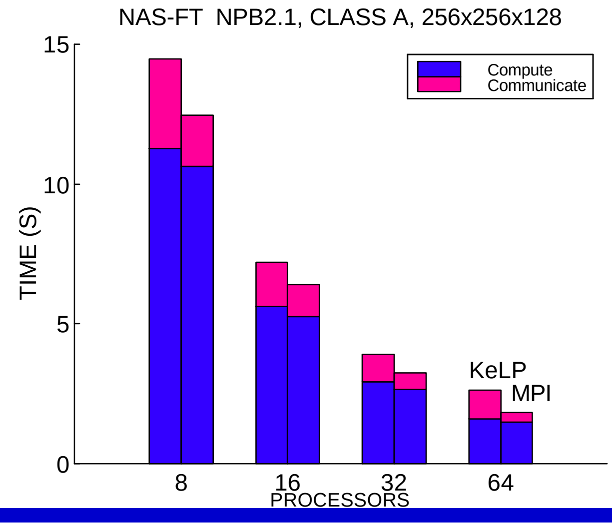


KeLP does not sacrifice performance







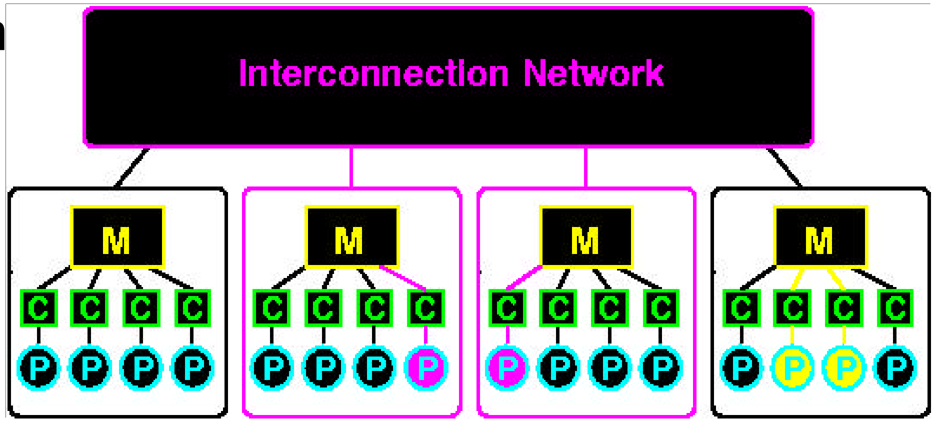


Managing Irregular Problems

- Heterogeneities in the computational domain introduce the need for multiblock domains
- Multiblock domains require load balancing because the computational domain is irregular
- UTPROJ3D employs non-matching unstructured grids, and uses "mortars" to glue the grids together
- KeLP hides significant software overheads that handle the bookkeeping
- User-defined load balancers may be customized to the specific geometry used by the application

Multi-tier KeLP

- Multiprocessor at each node amplifies the cost of communication
- MPI implementations require multithreading to mask communication
- Hybrid programming model is difficult to use
- A multi-tier prototype of KeLP provides a unified model for treating two levels of parallelism, hiding the details
- Adds an overlap capability to support latency-tolerant reformulation



Project 3

History Matching and Uncertainty with IPARS and ADR

- Exploration of history matching scenarios in view of uncertainty of geological data available
- Use of IPARS to generate multiple realizations
- Use ADR to store results of multiple realizations from IPARS simulator and to execute queries for history matching.

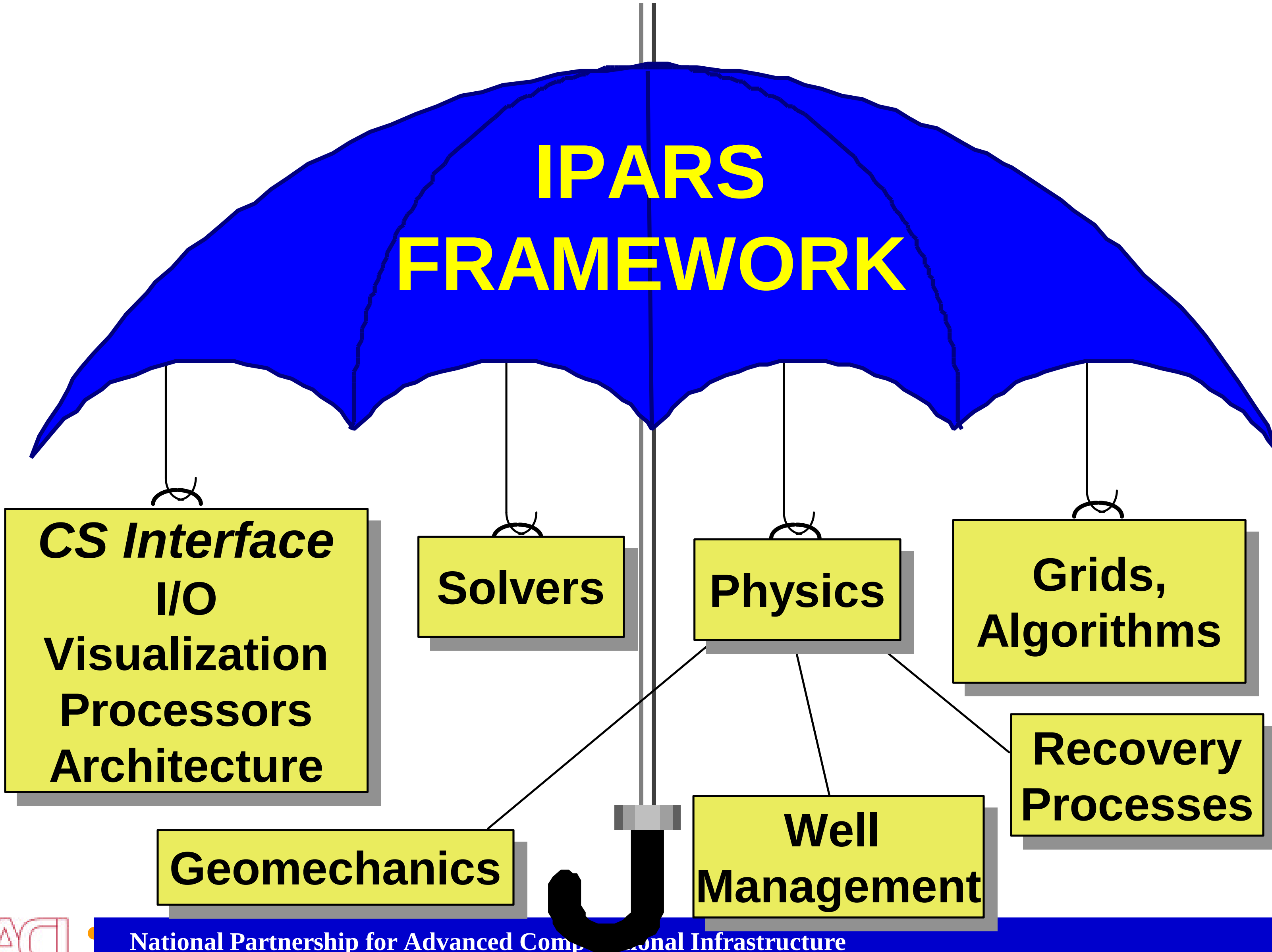
Implications of Uncertainty Hierarchy

- Different features of the problem are sensitive to different levels of the hierarchy
  - Recovery curves (waterflood, tracer) insensitive w.r.t. realization but highly sensitive to geological model
  - Local concentrations/saturations highly sensitive to realization
  - Single level of hierarchy insufficient to capture essential features

IPARS: Integrated Parallel Accurate Reservoir Simulator

- 8 individual physical models / algorithms for multiphase flow and transport
- Implemented in a common framework providing
  - memory management for general geometry grids
  - linear solvers with state-of-the-art preconditioners
  - portable parallel communication
  - keyword input and output with visualization
  - "hooks" for well management and other reservoir processes
- Code is portable across several serial and parallel platforms including Linux (clusters), SGI, RS6000, T3E, Windows (DOS)

IPARS


IPARS FRAMEWORK
CS Interface
I/O
Visualization
Processors
Architecture
Solvers
Physics
Grids, Algorithms
Geomechanics
Well Management
Recovery Processes

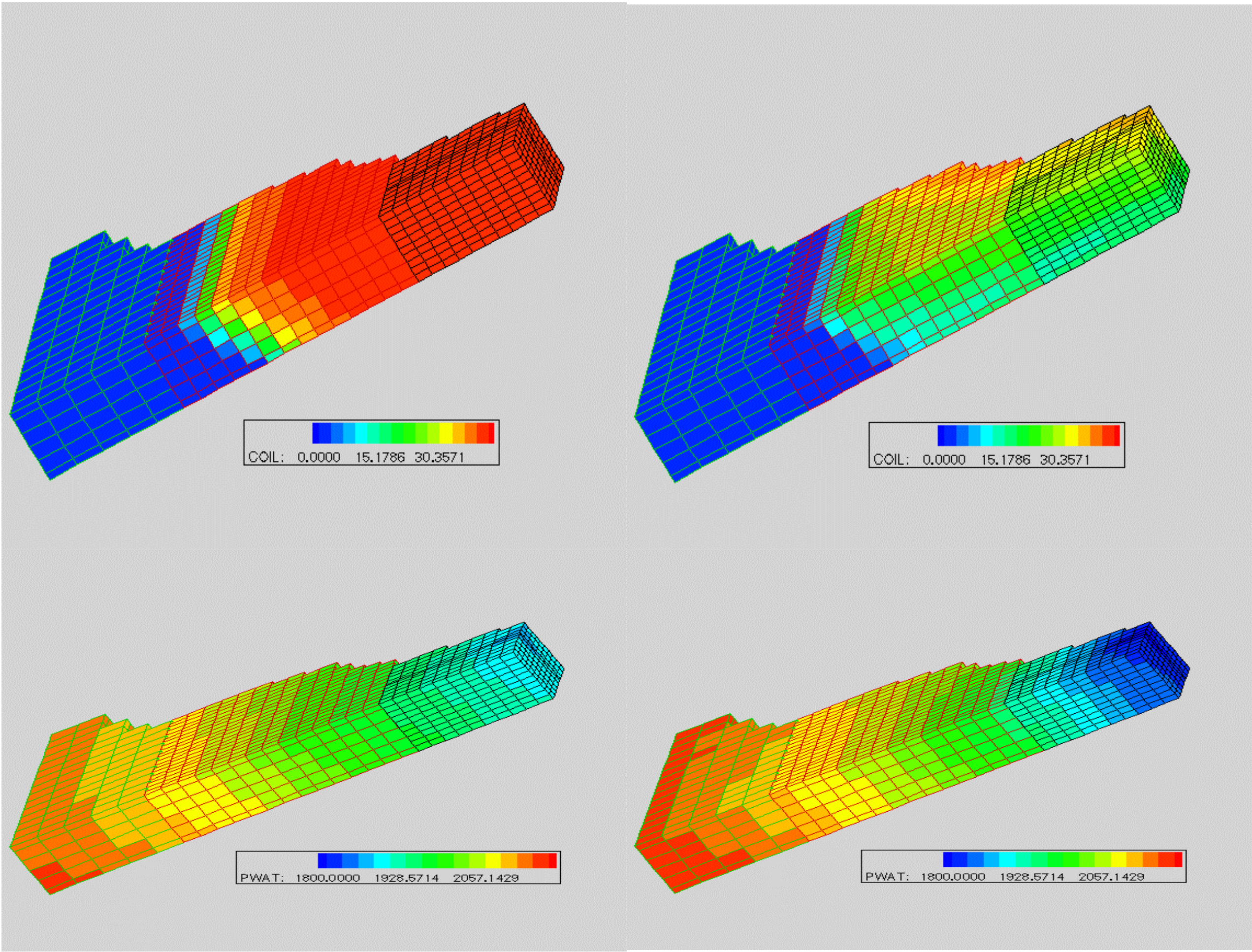
IPARS Multi-physics Results

Example:
(based on Borregos field) : dipping reservoir 1500 x 1300 x 40

•Simulation A: use Multiphysics Multiblock (black oil + two phase).

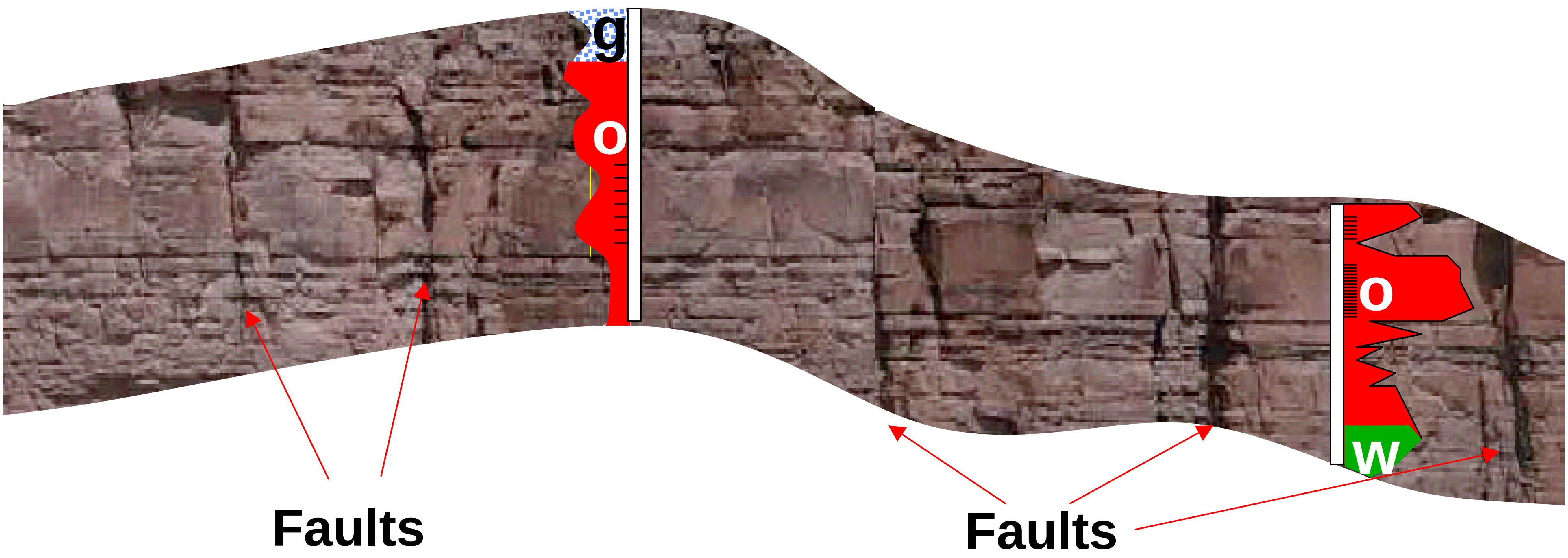
•Simulation B: use black oil single block code in the whole field.

(A) Day 1 **(A) Day 1000**

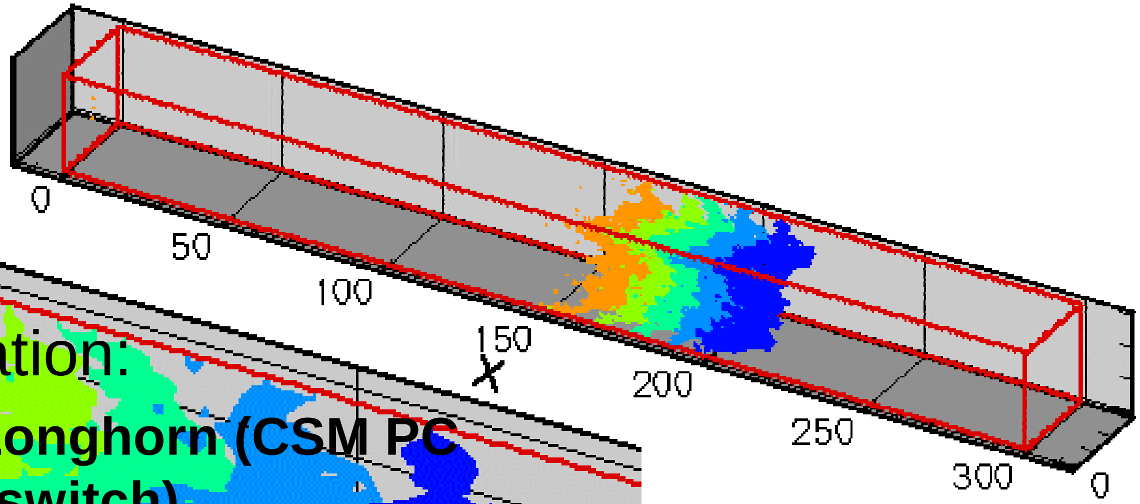


Flow in the Subsurface

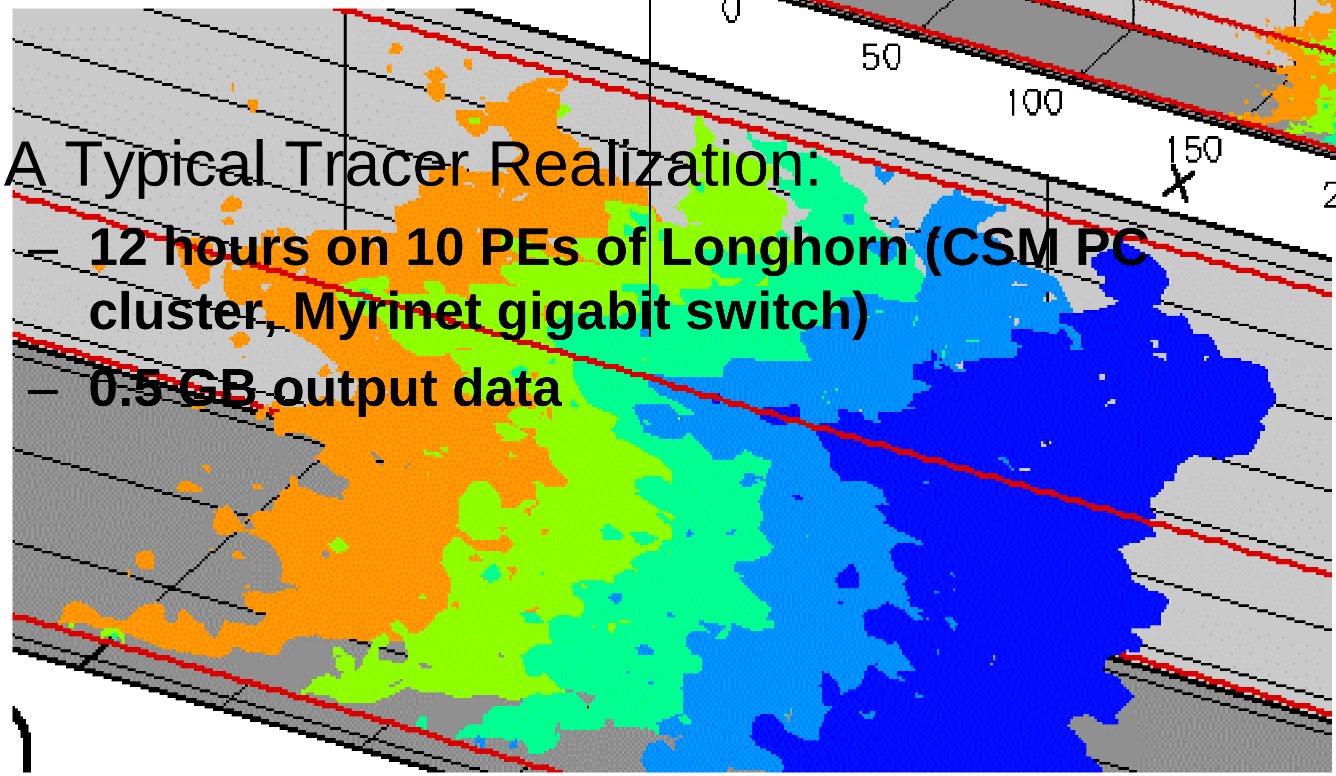
- Heterogeneities on multiple scales -- limited sampling
- Geostatistics is essential
- Hierarchy of uncertainty; Geological models, and Geostatistical models



Why ADR is important for IPARS

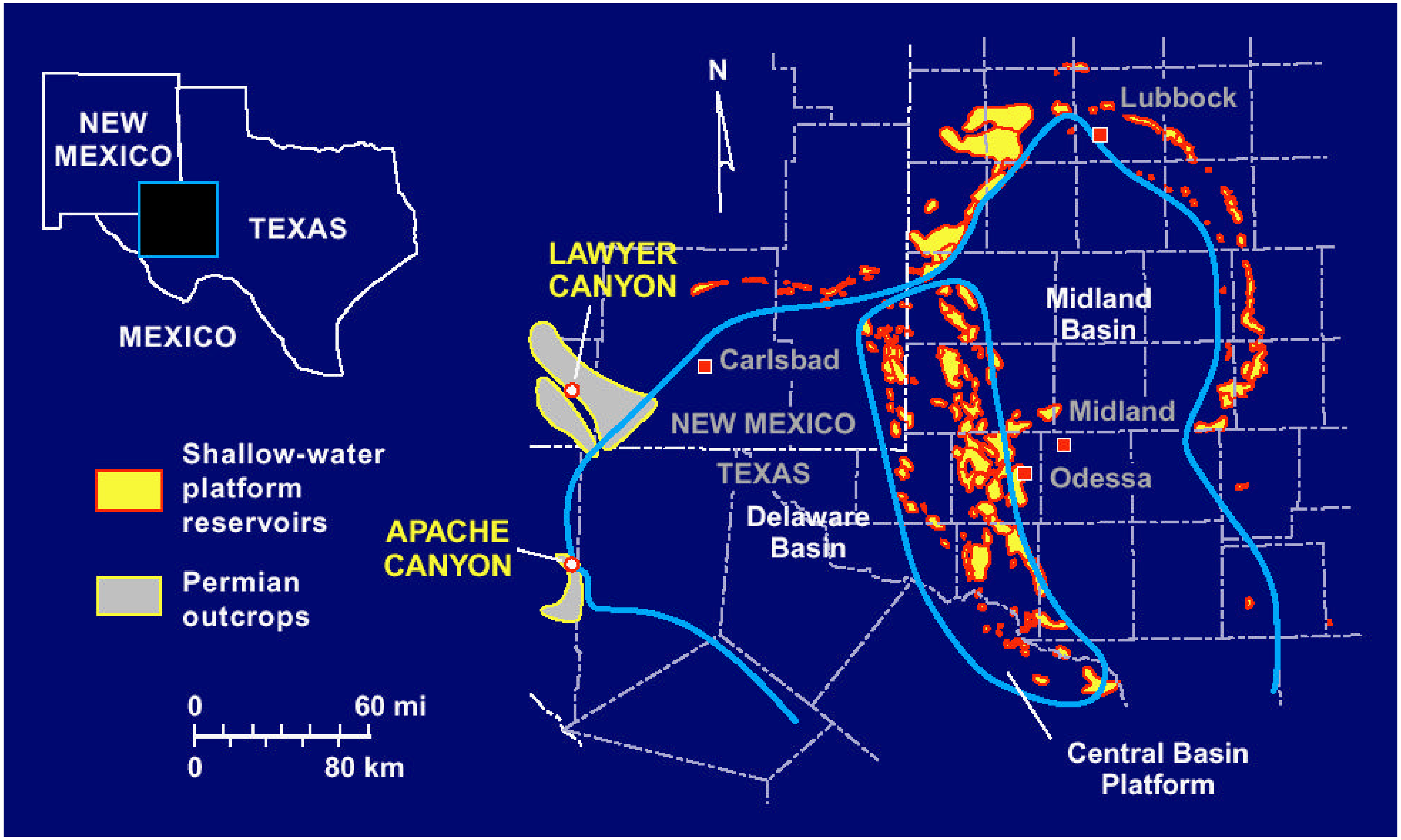


- A Typical Tracer Realization:
  - **12 hours on 10 PEs of Longhorn (CSM PC cluster, Myrinet gigabit switch)**
  - **0.5 GB output data**

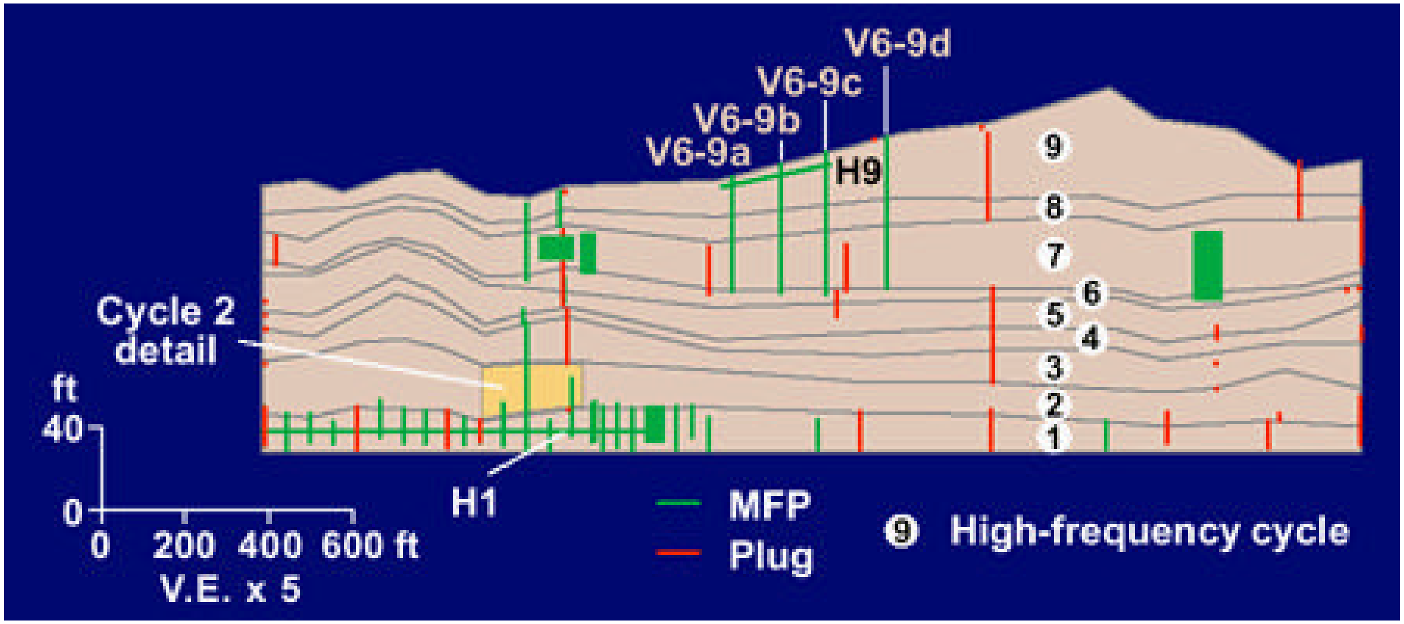


An Example: Guadalupe Mountains

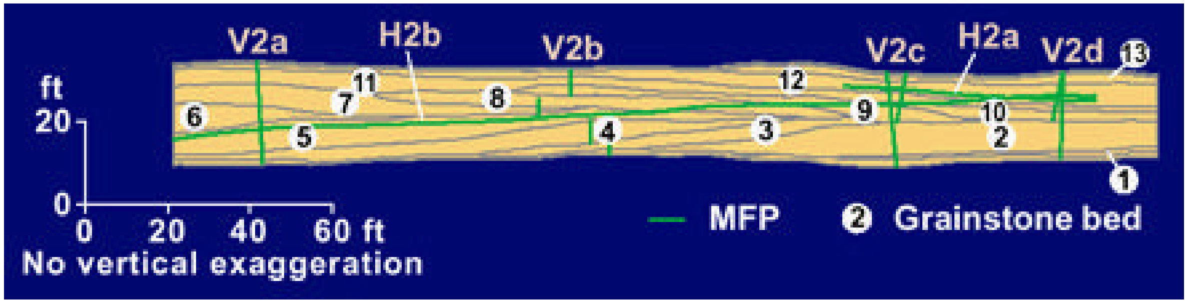
Outcrop Locations and Permian Paleogeography



Lawyer Canyon Ramp Crest Window

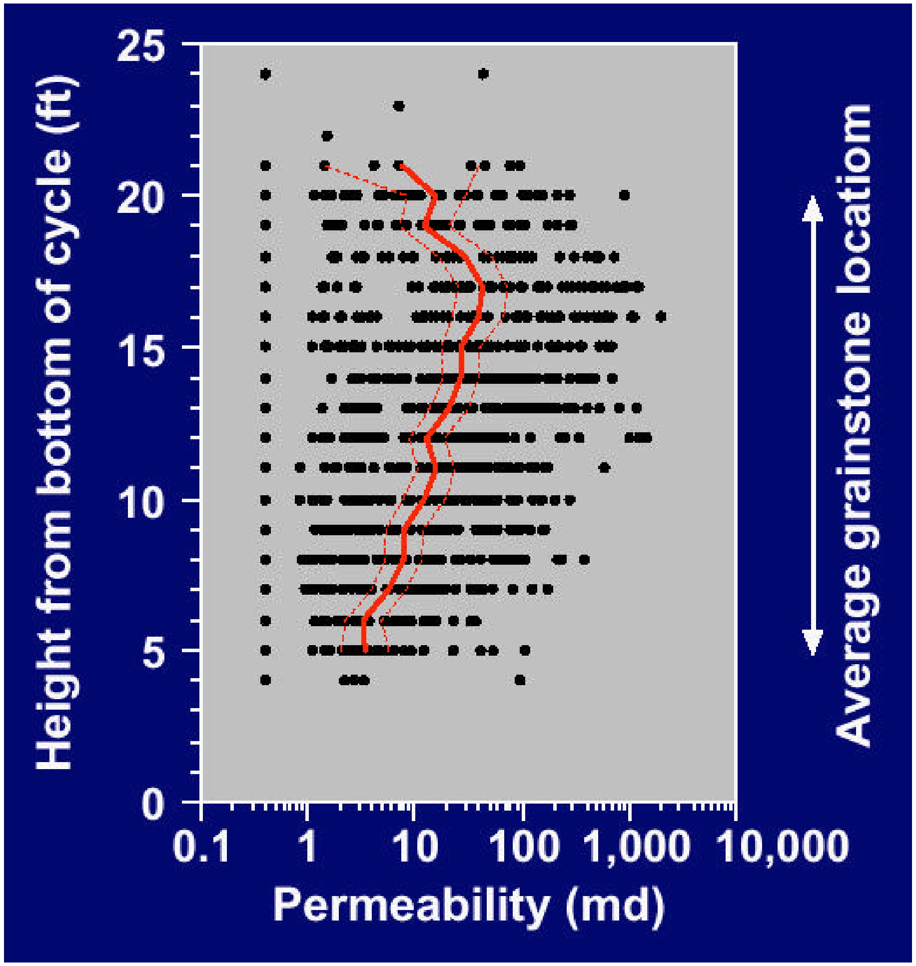


Cycle 2 detail



Grainstone Permeability

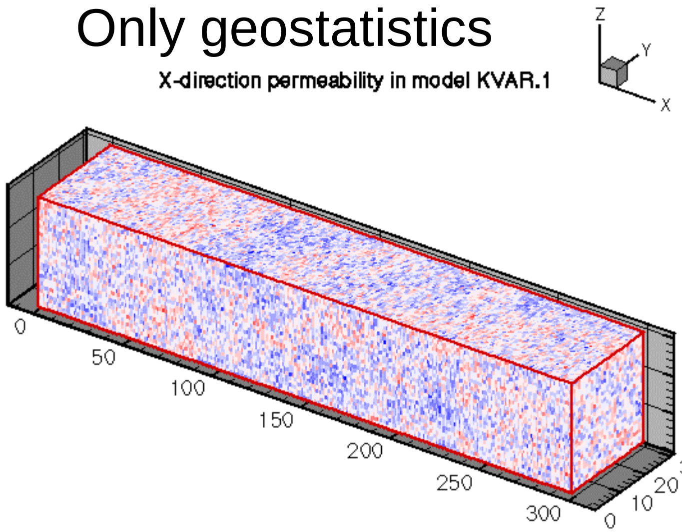
All Vertical Transects

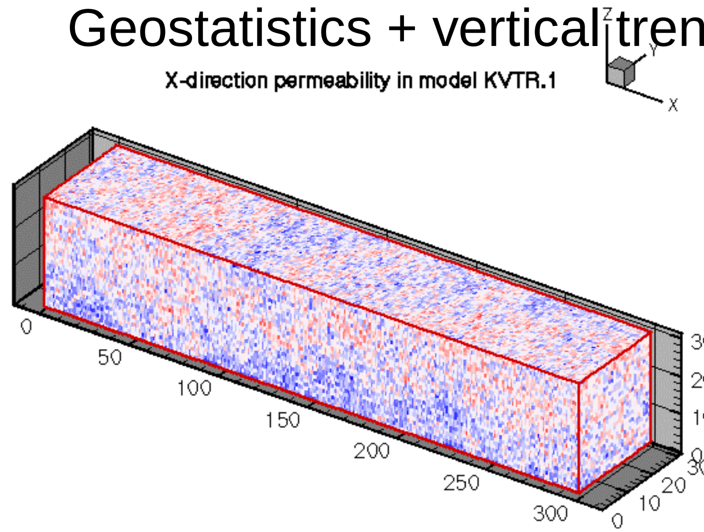


Higher level of hierarchy: different geological models

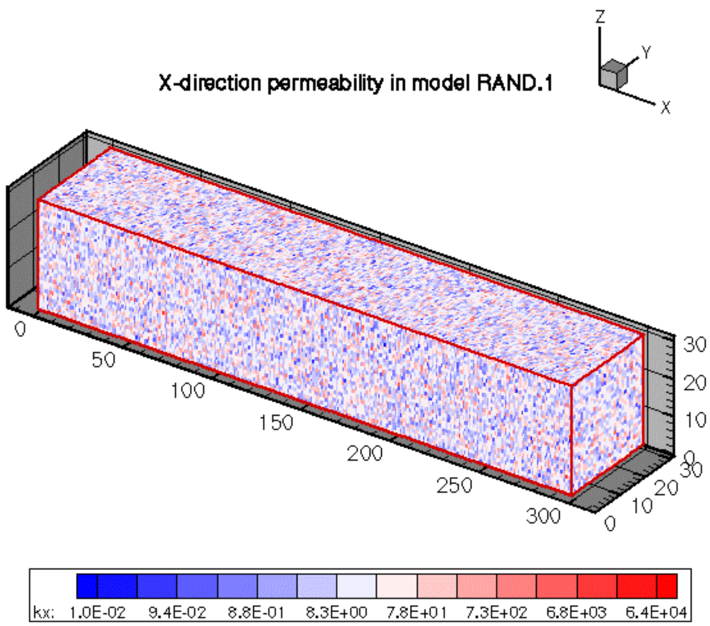
Only geostatistics | Geostatistics + vertical trend

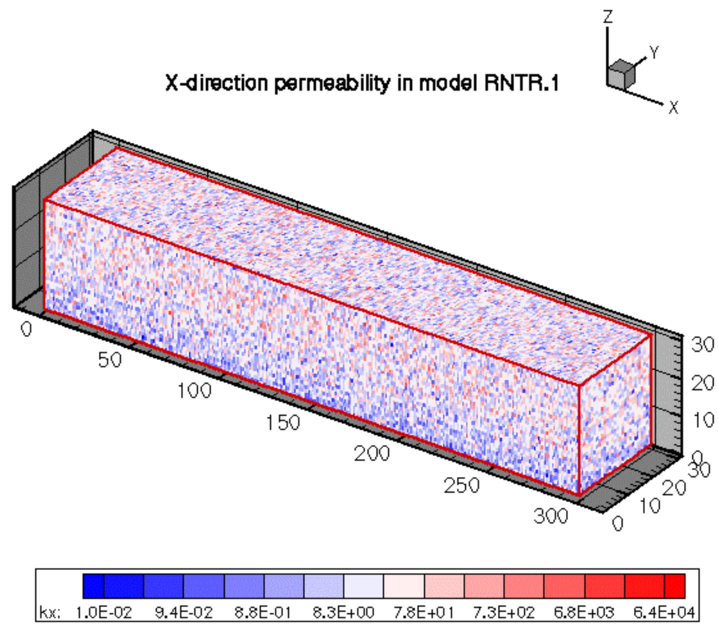
p=1



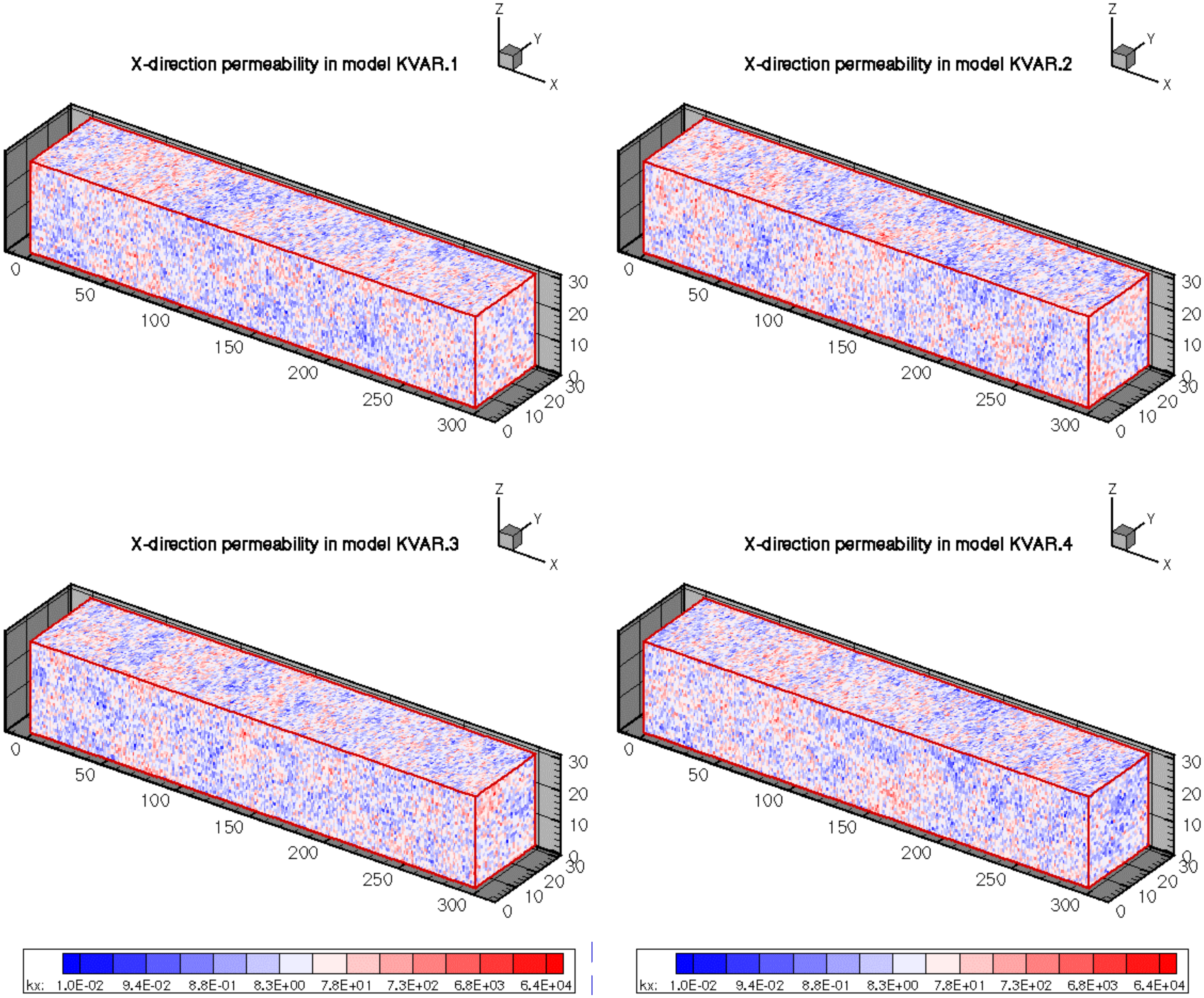


p=0

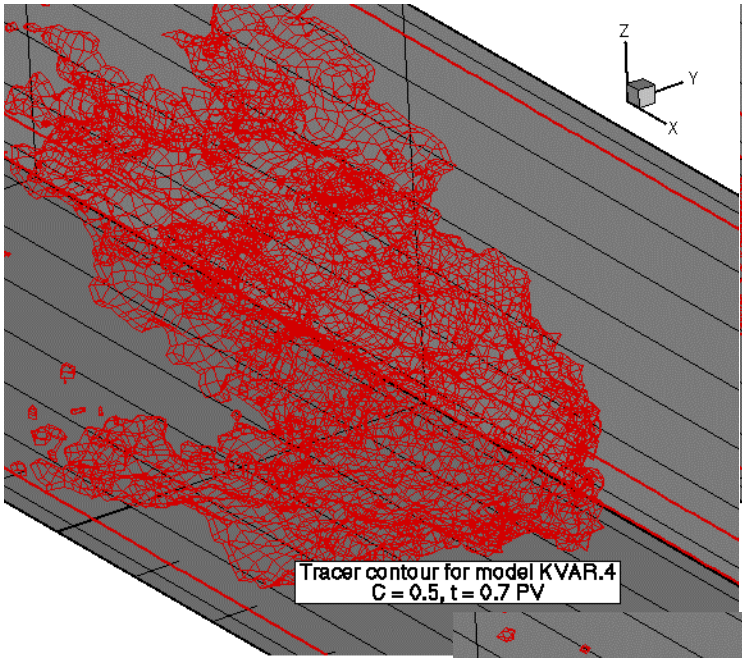


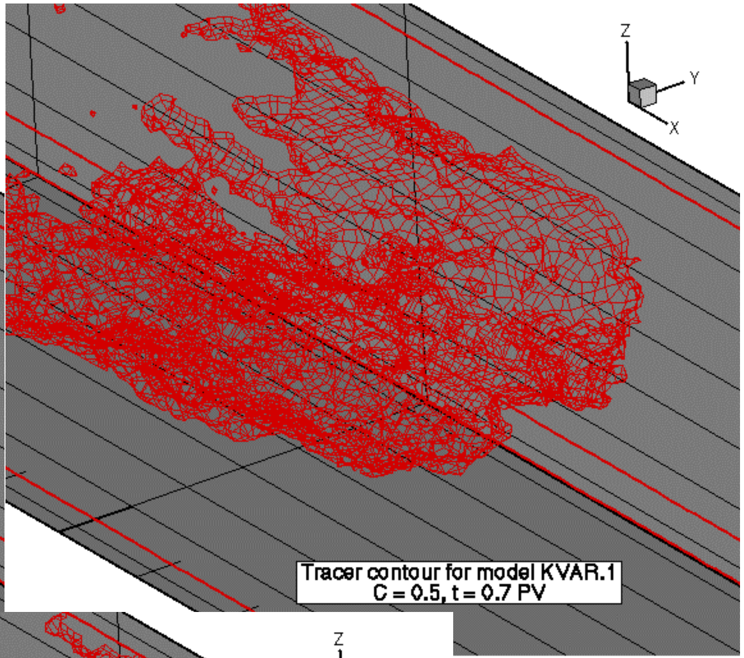


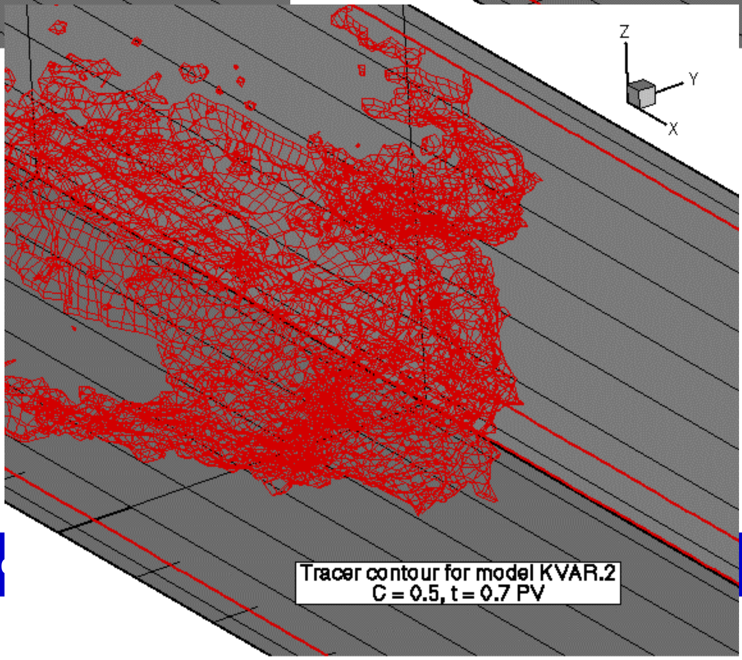
Lower level of hierarchy: different realizations of same model



Local behavior varies with realization





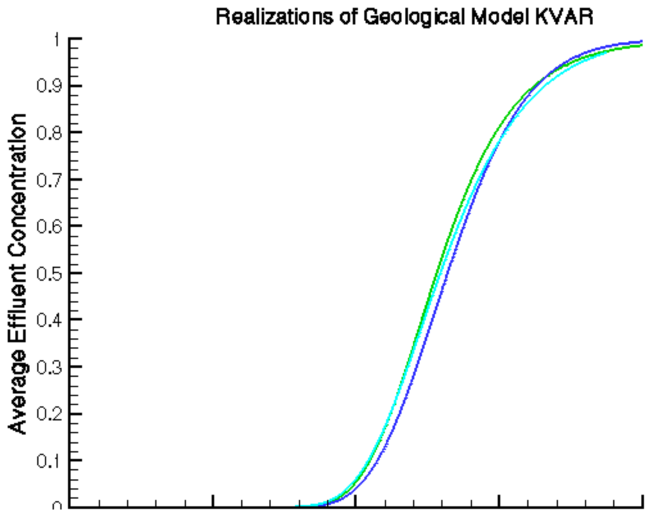


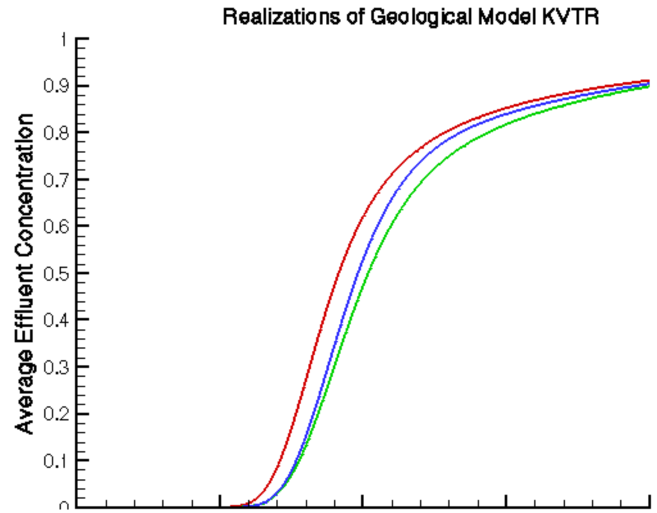
Global behavior varies with model

Only geostatistics

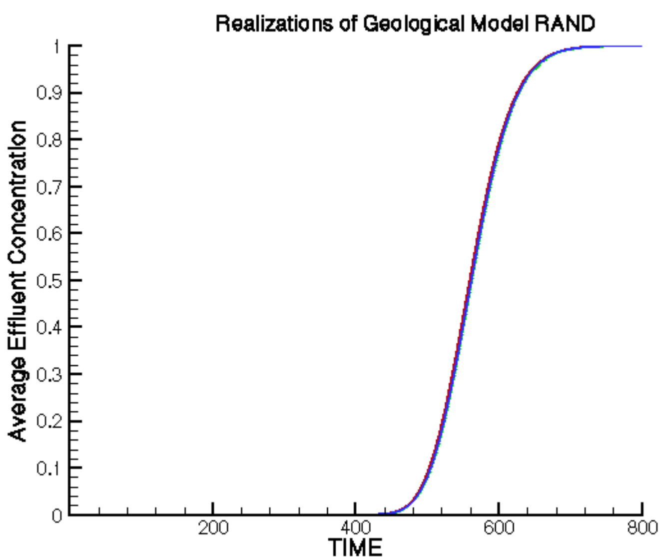
Geostatistics + vertical trend

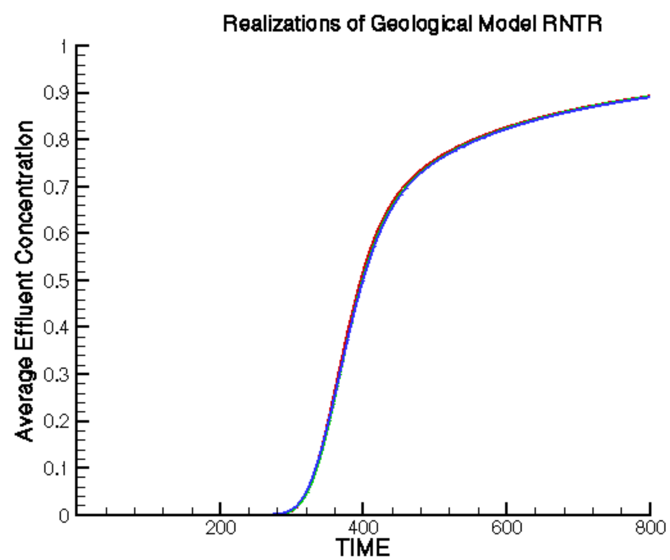
p=1





p=0





Implications of Uncertainty Hierarchy

- Different features of the problem are sensitive to different levels of the hierarchy
  - Recovery curves (waterflood, tracer) insensitive w.r.t. realization but highly sensitive to geological model
  - Local concentrations/saturations highly sensitive to realization
  - Single level of hierarchy insufficient to capture essential features

Goals and Deliverables for FY00

- Enhancement of Tools
  - Integrate KeLP datatypes into ADR, layer KeLP on top of Globus.
  - Initial porting of ADR and KeLP-based codes to Tflops machine
  - Generalize KeLP to handle parallel static unstructured grids
- Coupling of Surface Water (SW) and Ground Water (GW) Models
  - Customize ADR to process GW and SW simulation data
  - Multiple scenarios using IPARS and ADR
  - Couple surface and groundwater codes using KeLP
- Parallelization of Selected Ground Water Codes
  - Parallelize UTPROJ3D using the generalized KeLP

Goals and Deliverables for FY01

- Enhancement of Tools
  - Optimize port of ADR, KeLP, MetaChaos, Globus to Tflops
  - Integrate KeLP, ADR, Globus and MetaChaos
- Testing of Coupled Multi-Physics Models
  - Demonstrate KeLP coupled subsurface/surface simulators
  - Use ADR for examination of coupling of realistic surface and subsurface water scenarios and carry out history matching
  - Use of ADR supported analysis of IPARS to examine petroleum production scenarios and history matching

Accomplishements since 4/99

- Definition of specific projects to meet the goals
- Initiation of the work on simulators and infrastructure
  - Parssim has been decomposed into flow and transport parts: code is being tested.
  - Design of use of KeLP for UTPROJ3D. Single block version is running.
  - Identification of input/output for history matching scenarios with IPARS and ADR.
  - Hardening of ADR
  - Porting of ADR to Linux clusters

Vision Beyond FY01

- Coupling of compositional multi-component and multiphase flow with energy balance, geomechanical and geochemical models
- Pilot Scale Modeling of Regional C02 Sequestration
  - Collect CO_2 and inject into subsurface for enhanced oil recovery
  - Mix CO_2 with mineral salts and inject into subsurface to form solid mineral compounds
- Coupling of Geomechanical with Flow Simulation

Remaining Questions

- Are you dependent on anything out of your control?
  - Continuity of staff
  - Stability and availability of hardware
- What has changed?
  - No major changes in focus or direction
- Other thrust support needed?
  - Production Visualization