Large Scale Computing Requirements for Basic Energy Sciences

(An BES / ASCR / NERSC Workshop)

Hilton Washington DC/Rockville Meeting Center, Rockville MD

3D Geophysical Modeling and Imaging

G. A. Newman

Lawrence Berkeley National Laboratory

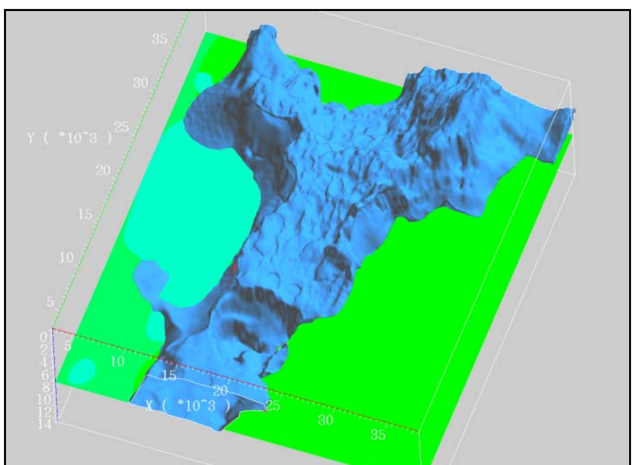
February 9 – 10 , 2010

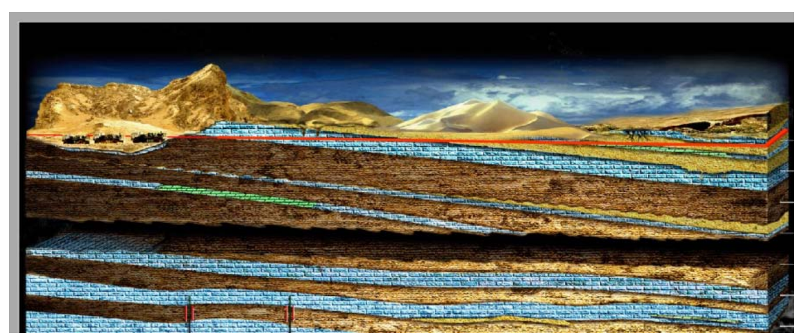
Talk Outline

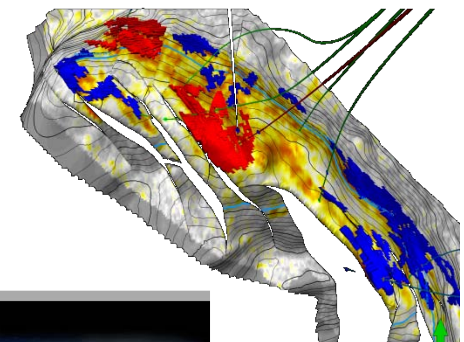
- SEAM Geophysical Modeling Project – Its Really Big!
- Geophysical Imaging (Seismic & EM) – Its 10 to 100x Bigger!
  - Reverse Time Migration
  - Full Waveform Inversion
  - 3D Imaging & Large Scale Considerations
  - Offshore Brazil Imaging Example (EM Data Set)
- Computational Bottlenecks
- Computing Alternatives
  - GPU's & FPGA's
  - Issues

SEAM
SEG Advanced Modeling

Why ? So that the resource industry can tackle grand geophysical challenges (Subsalt imaging, land acquisition, 4-D, CO_2, carbonates ……)







BERKELEY LAB

SEAM Mission

Advance the science and technology of applied geophysics through a cooperative industry effort focused on subsurface model construction and generation of synthetic data sets for geophysical problems of importance to the resource extraction industry.

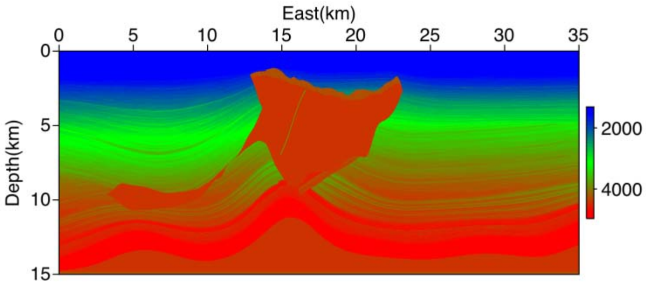
SEAM
SEG Advanced Modeling

Berkeley Lab

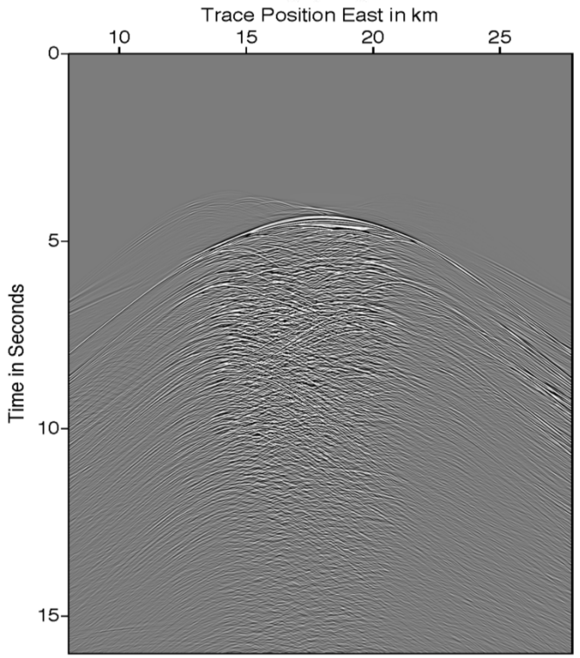
SEG Post-Convention Workshop
SEAM Phase 1: Initial results

Technical and Operational Overview of the SEAM Phase I Project

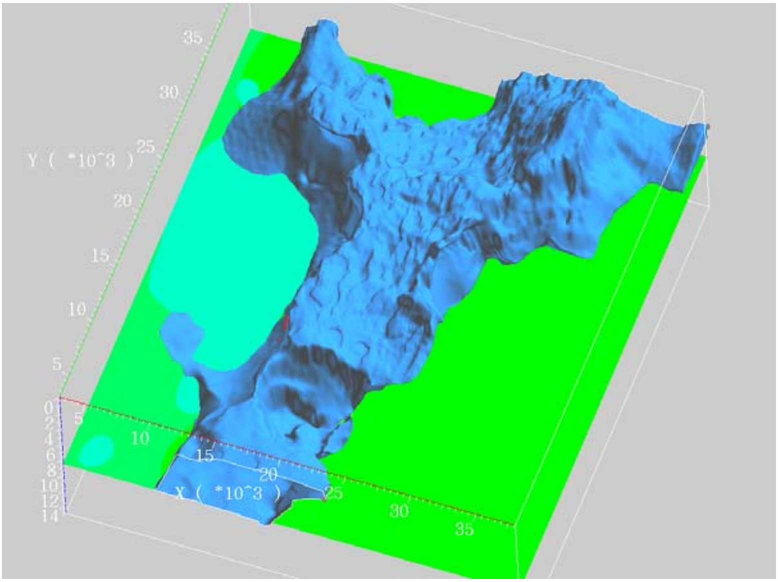
Michael Fehler

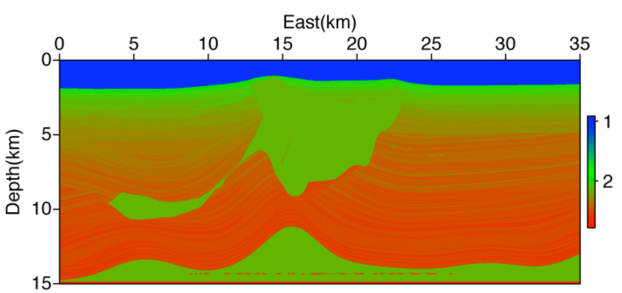


Final Model Velocity (m/s) for North=16 km

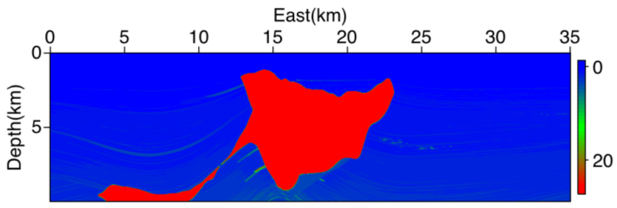


EW Shot Gather at N=16.010 km, Shot at E,N = 18.025,21.950 km

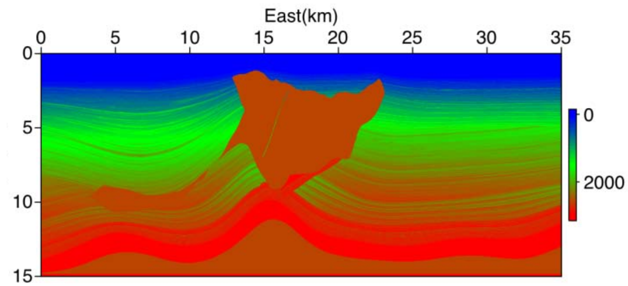




Final Model Density for North=16 km



Horizontal Resistivity in Ohm-m for North=16 km



Shear-Wave Velocity (m/s) for North=16 km

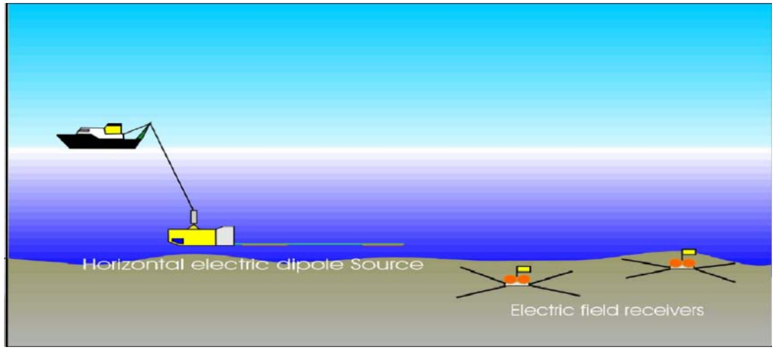
SEAM PROJECT
Seismic Modeling Considerations

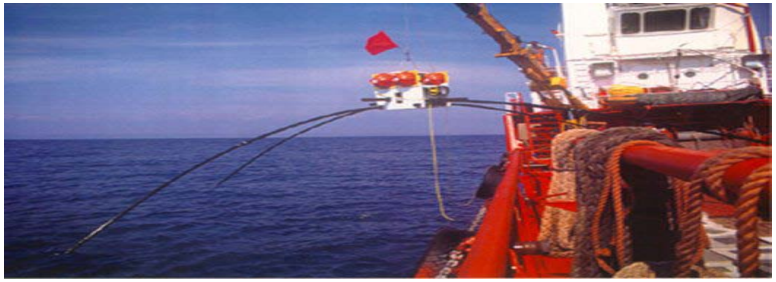
- 65,000 shots
- 450,000 traces per shot
- Traces 16 seconds length samples at 8 ms
- Data volume per shot: 3.5GB
- 228 TB Disk Space required for all shots & traces

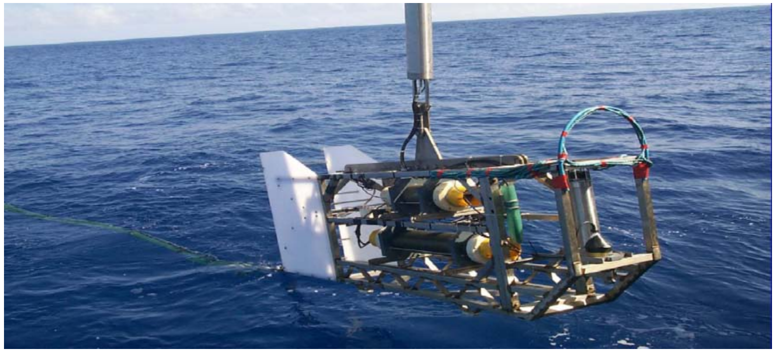
BERKELEY LAB

THE MARINE CSEM/MT METHOD

- Deep-towed Electric Dipole transmitter
  - ~ 100 Amps
  - Water depth to 5 to 7 km
  - Alternating current 0.01 to 3 Hz
  - 'Flies' 50m above sea floor'
- Seafloor MT receivers
  - Measure orthogonal E & H
- The Geophysical Signature
  - Oil & gas reservoirs electrically resistive than background media
  - A non seismic indicator
  - Still requires seismic data to constrain interpretation
- Data Volumes ~ 1% to 10% seismic
  - Still ~ 2TB







GEOPHYSICIAL IMAGING

- Seismic
  - 3D Reverse Time Migration
    - Large Scale Computations: 1,000s Cores, Weeks of Processing
  - 3D Full Waveform Inversion
    - Iterative reverse time migration
    - Promises Much Greater Image Fidelity
    - Formidable Numerical Issues – Local Minima, Very Good Starting Models Required
    - Frontier Research Area
    - Enormous Computation: 10,000's Cores, Months of Processing
- Electromagnetic (CSEM & MT)
  - 3D Full Waveform Inversion
    - Provides information on non-seismic attributes
    - Complements seismic imaging – through lower resolution
    - Constrained by seismic imaging
    - Computational demands also big: 1,000s to 10,000s cores
- Joint Seismic-Electromagnetic Imaging
  - The Holy Grail ?
    - Frontier Research Area
    - Grand Challenge Problem

BERKELEY LAB

Wave Equations for Geophysical Imaging

Acoustic Waves

Time Domain

$$\left[\frac{1}{v^2}\frac{\partial^2}{\partial t^2}-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\right]p(x,y,z,t)=s(t).$$

Frequency Domain

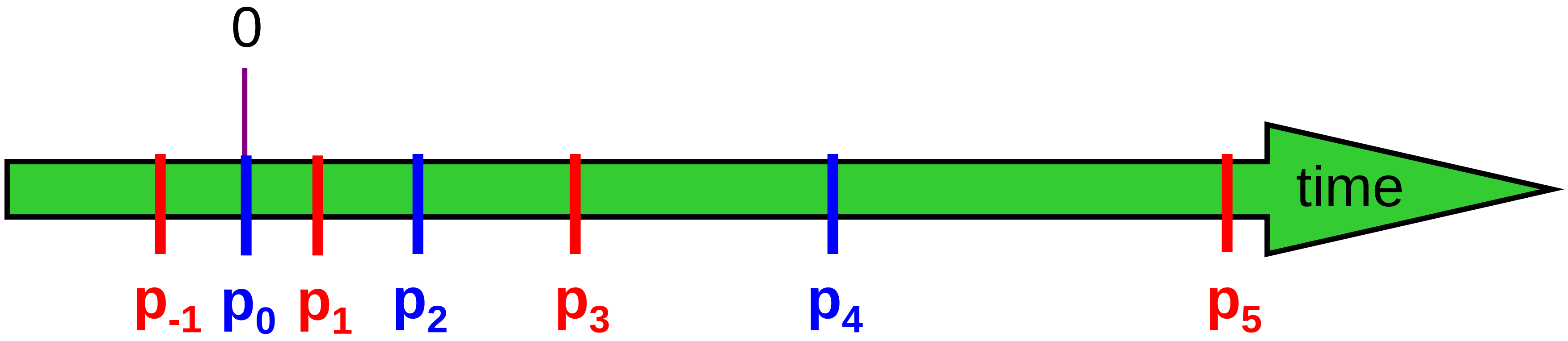
$$\left[\frac{\omega^2}{v^2}-\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)\right]p(x,y,z,\omega)=s(\omega).$$

Electromagnetic Waves

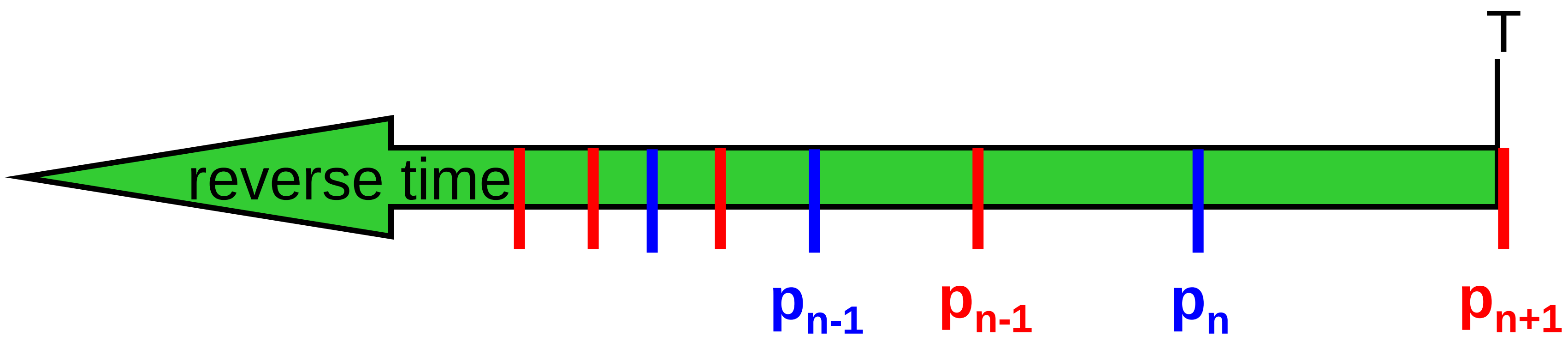
$$\nabla \times \nabla \times \mathbf{E_s} + i\omega\mu\underline{\sigma}\,\mathbf{E_s} = \mathbf{S}.$$

Discretization Methods: Finite Differences, Finite Elements
Solution Methods: Explicit, Implicit – iterative Krylov solvers for 3D problems

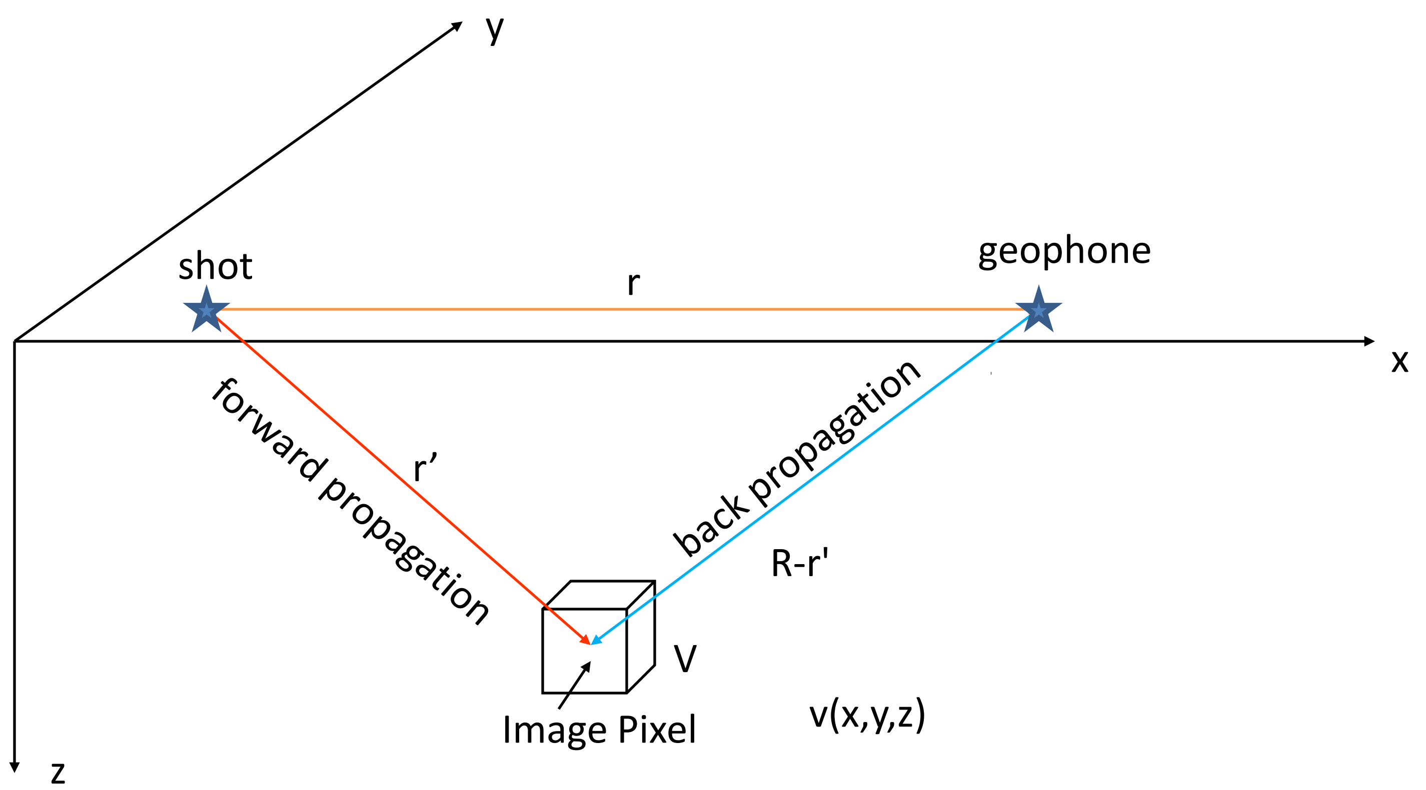
Forward Propagation



Back Propagation



Geophysical Imaging



Imaging Condition : cross correlation of the forward and back propagated wave fields
all shots, all geophones

Reverse Time Migration: main algorithm

For each shot ($s(r_i)$,i=1,..n)

1/ Solve Forward Propagation Problem and Store Wavefield

For $t=t_o,\ldots,t_{max}$ in time or $\omega=\omega_o, \ldots,\omega_{max}$ in frequency
    Compute source wavefield at time t or frequency ω
    Store wavefield
End

2/ Solve Back Propagation Problem and Apply Imaging Condition

For $t=t_{max},\ldots,t_o$ in time or $\omega=\omega_o, \ldots,\omega_{max}$ in frequency
    Compute receiver wavefield at time t or frequency ω
    Read forward wavefield at time t or frequency ω
    Compute imaging condition
End

3/ Update Image

End

Full Waveform Inversion: main algorithm

For Model Update(k); k=1 to k_{max}, or Until Convergence

For each shot ($s(r_i)$,i=1,..n)

1/ Solve Forward Propagation Problem and Store Wavefield

For $t=t_o,\ldots,t_{max}$ in time or $\omega=\omega_o, \ldots, \omega_{max}$ in frequency

Compute source wavefield at time t or frequency ω

Store wavefield

End

100's Iterations

Computing Cost 100x RTM

2/ Solve Back Propagation Problem and *Compute Gradient of the Error Functional*

For $t=t_{max},\ldots,t_o$ in time or $\omega=\omega_o, \ldots, \omega_{max}$ in frequency

Compute receiver wavefield at time t or frequency ω

Read forward sweep wavefield at time t or frequency ω

Compute gradient

End

3/ Update Attrubutes Using Simple Line Search (2 forward solves per shot)

4/ Compute Data Misfit; if < tol Stop, Otherwise Cycle Model Update Loop

End

3D GEOPHYSICIAL IMAGING

- Why 3D?
  - Data acquisition is 3D
    - 2D interpretation often not appropriate
  - Prospective Oil & Gas Reservoirs & Targets Inherently 3D
    - complex geology

- Philosophy on 3D Modeling & Inversion Methods
  - Interpretation must be as accurate as possible
    - high stakes; offshore platforms & drilling - 100's millions of dollars
  - Treat large-scale nature of the interpretation problem
    - High density Seismic, CSEM & MT data sets; millions of data points
    - Large-scale imaging volumes; millions of image pixels
  - Avoid approximations
    - Methods must be as accurate as possible & robust & reliable

LARGE-SCALE CONSIDERATIONS

- Require Large-Scale 3D Modeling and Imaging Solutions
  - 200 million field unknowns - forward (fwd) problem
  - Imaging grids 400 nodes on a side

- Parallel Implementation
  - Multiple levels of parallelization
    - Model Space (simulation and inversion mesh)
    - Data Space (each transmitter-receiver set fwd calculation independent)
    - Installed & tested on multiple distributed computing systems; 10 – 10,000s processors/cores

- Above procedure satisfactory except for very largest problems
  - To treat such problems requires a higher level of efficiency

- Optimal Grids
  - Separate inversion grid from the simulation/modeling grid
  - Potential for significant solution acceleration ~ order of magnitude

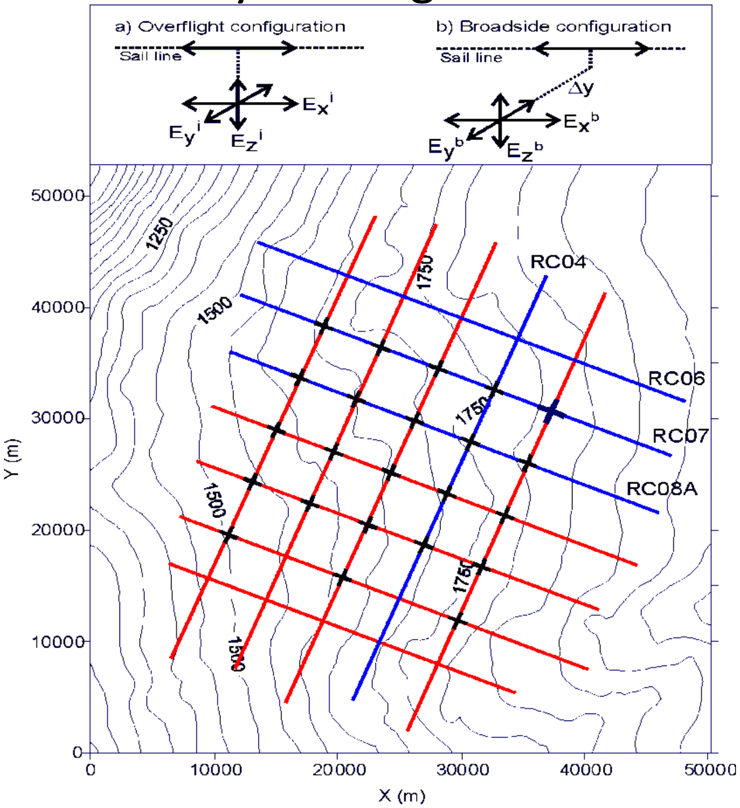
OFF SHORE BRAZIL CSEM DATA
3D Image Processing Requirements

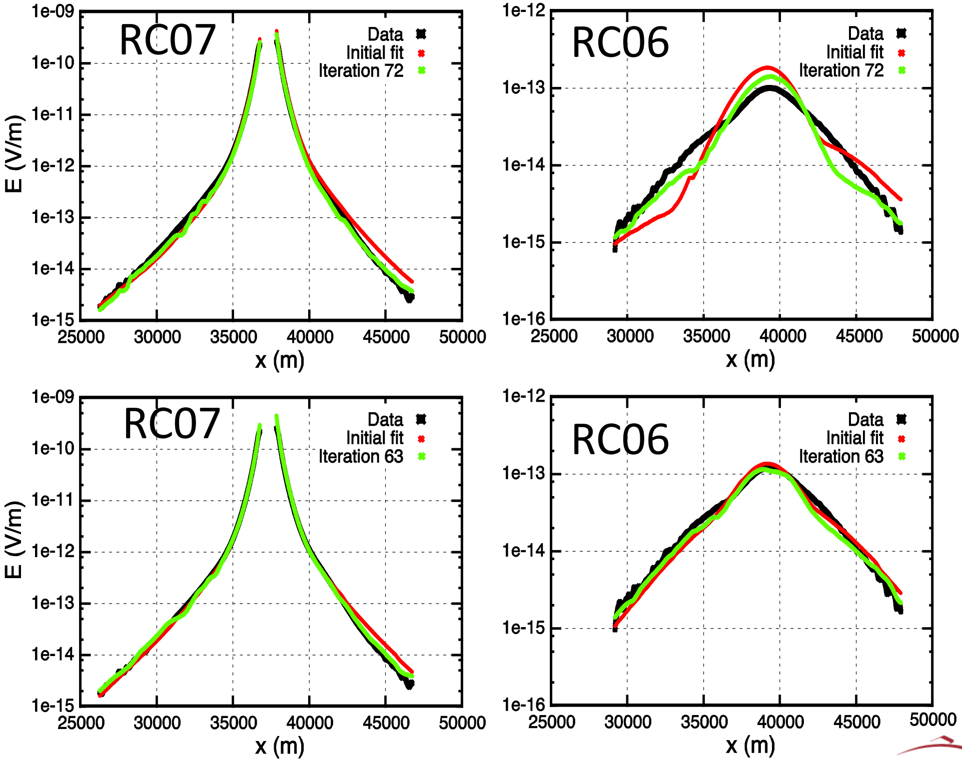
- 3D Data and Imaging Volumes
  - nearly 1 million data points, 207 effective transmitters (reciprocity processing significantly reduces number of transmitters)
  - more than 27 million modeling cells (a large subset to be updated within the inversion process)
- Image Processing Linux Clusters
  - 1024 tasks with Infinband fabric => several months of processing time
- Use Blue Gene (L) Super Computer for Faster Time to Solution
  - 32 766 processors/tasks used to image the data
  - each task has only 250 Mbytes memory (requires fine grained model decomposition over 512 tasks)
  - 64 data planes employed in the image processing
  - delivers imaging results in 24 hours compared to several months

OFF SHORE BRAZIL
3D CSEM IMAGING EXPERIMENT

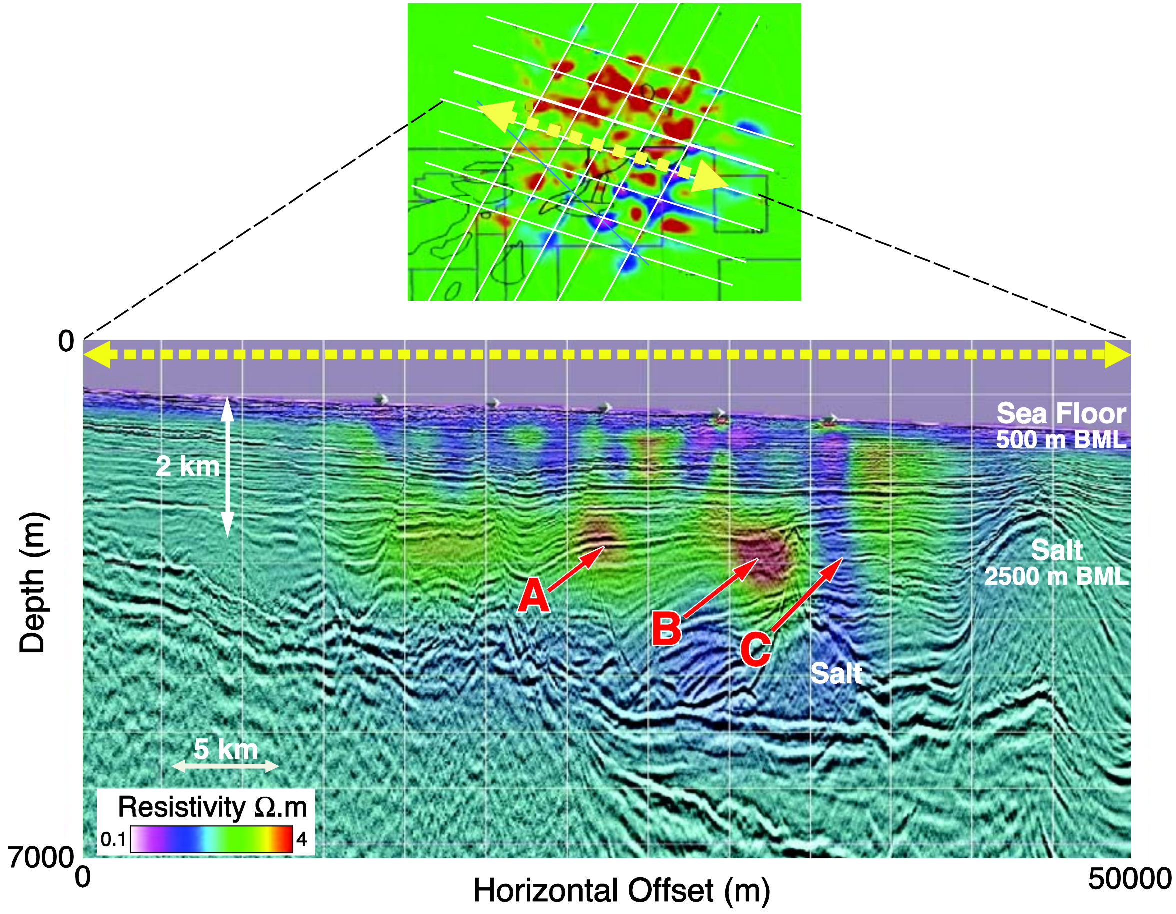
23 sea bottom detectors
10 sail lines
3 transmitting frequencies
1.25, 0.75 and 0.25 Hz
Survey coverage ~900 km^2

Isotropic Conductivity Model:
-- can not fit broadside data
Anisotropic Model Required:
-- horizontal & vertical conductivities





3D Electrical Conductivity Imaging



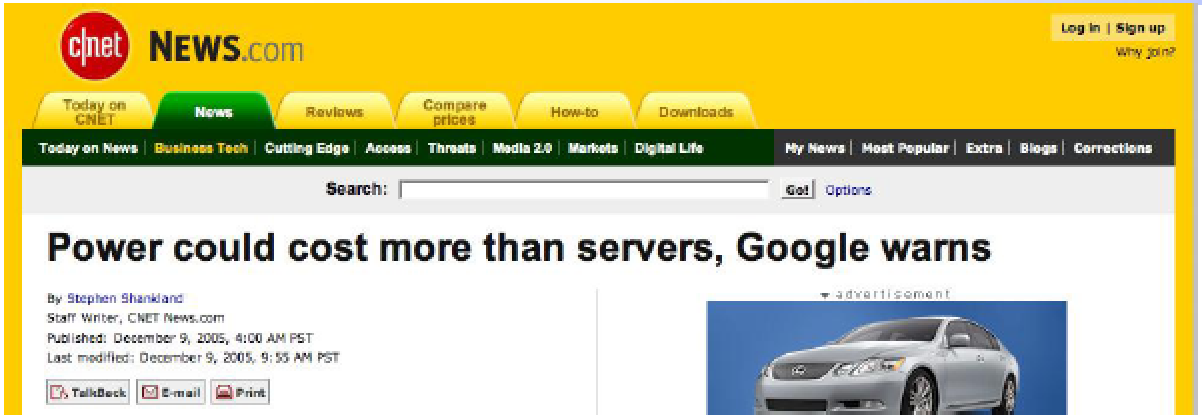
Computational Bottlenecks

- Forward and Backward Solves
- Data IO & Memory (seismic)
- Time to Solution
- Multiple Imaging Experiments Required
  - Assess Model Uncertainties
  - Test Different Starting Models
  - Test Different Noise Assumptions
- Scale Problem Up to Ever More Cores
  - Impractical; power demands and cost

Power is an Industry Wide Problem

(2% of US power consumption and growing)

NATIONAL ENERGY RESEARCH SCIENTIFIC COMPUTING CENTER



The New York Times "Hiding in Plain Sight, Google Seeks More Power", by John Markoff, June 14, 2006

New Google Plant in The Dulles, Oregon, from NYT, June 14, 2006

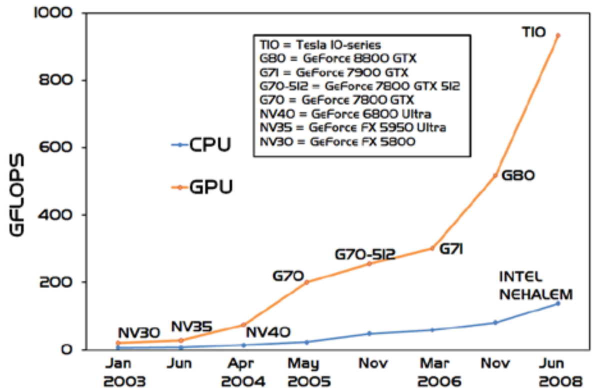
Relocate to Iceland?

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB

Computing Alternatives

- GPU's and FPGA's – Big Opportunities in Seismic & EM Imaging
  - 10x performance
  - Keeps Cooling and Power Cost Manageable



- Issues
  - GPU's and FPGA's Can Be Difficult to Program
  - Peak Performance Can Be Illusive (jungle programming)
  - IO Constrained ?
  - Double Precision ?