High-Performance Computing (HPC)（高性能计算) in the Geosciences

David A. Yuen 阮维基

University of Minnesota ,
U.S.A. and
China University of
Geosciences, Wuhan ,China

My 40 year experience in HPC

- I always like to use the biggest computer available ,since I encountered CDC 7600 in 1971 at Livermore. I have started to use CRAY machines at Boulder in 1984 and moved to Minnesota Supercomputing Institute in 1985 and since have used CRAY-2, CRAY-YMP, ETA, CM-5 , SGI Clusters, IBM and now focussing GPU on MIC(迈克) (many integrated core) from Intel. I have also been working on **interactive visualization（交互式可视化）** and **collaborative computing（协同计算）**. I am a member of AGU (Fellow), A.P.S. and S.I.A.M.

Outline

- INTRODUCTION
- Why do we need High-performance Computing (HPC) in geosciences ?
- Joe Scientist's efforts to do scientific computing
- Need to carry out visualization and collaboration in HPC for big problems
- Conclusions and **Future Expectations （前景）**

Why do we need HPC in the geosciences？

- (1) in order to do problems with greater degree of realism and complexity
- (2) Multi-physics, multi-chemistry and multi-scales（多尺度） because of nonlinearities（非线性）
- : Earthquake physics, 3D seismic wave propagation ,volcanic eruptions , hydrology, mantle and crustal dynamics, mineral physics （矿物物理学）(DFT and MD), surface processes(basin formation), tsunami and floods, geodynamic and many more other examples.

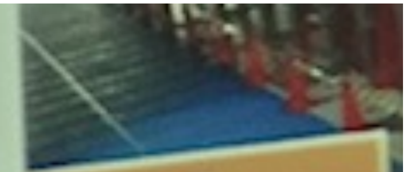
Desk-top（桌面计算） or office server computing is limited in scope .

- Linux workstation with 8 cores and 64 Gbytes server cannot do state-of-the-art work. Only good for code-development.It has 2 Tbytes of internal storage and 2 Tbytes
- outside
- Simulations of 300x300x300 grid points or 20 million tracers. That was what people were doing on high-end supercomputers
- 25 years ago
- Workstation-type computing is not competitive at all. No funding

Progress in HPC

- 25 years ago in 1987 CRAY-2 and ETA were vector machines , cruising at O(100) Mflops. Gigaflop barrier 10**9
- was broken around that time, in early ‘90s.
- Teraflop 10**12 barrier was broken in the late 20th century
- Petaflop, 10**15, barrier was broached in 2009 with massively parallel computing
- Today we are at around 16 Petaflops with the Sequoia, Livermore computer, No 1
- Exaflops hopefully will be reached in 2017 or 2018. This is EXASCALE COMPUTING
- In the last 25 years an increase of computing speed by a factor of **10,000**

accelerators.



Top5 of TOP500 List - November 2011

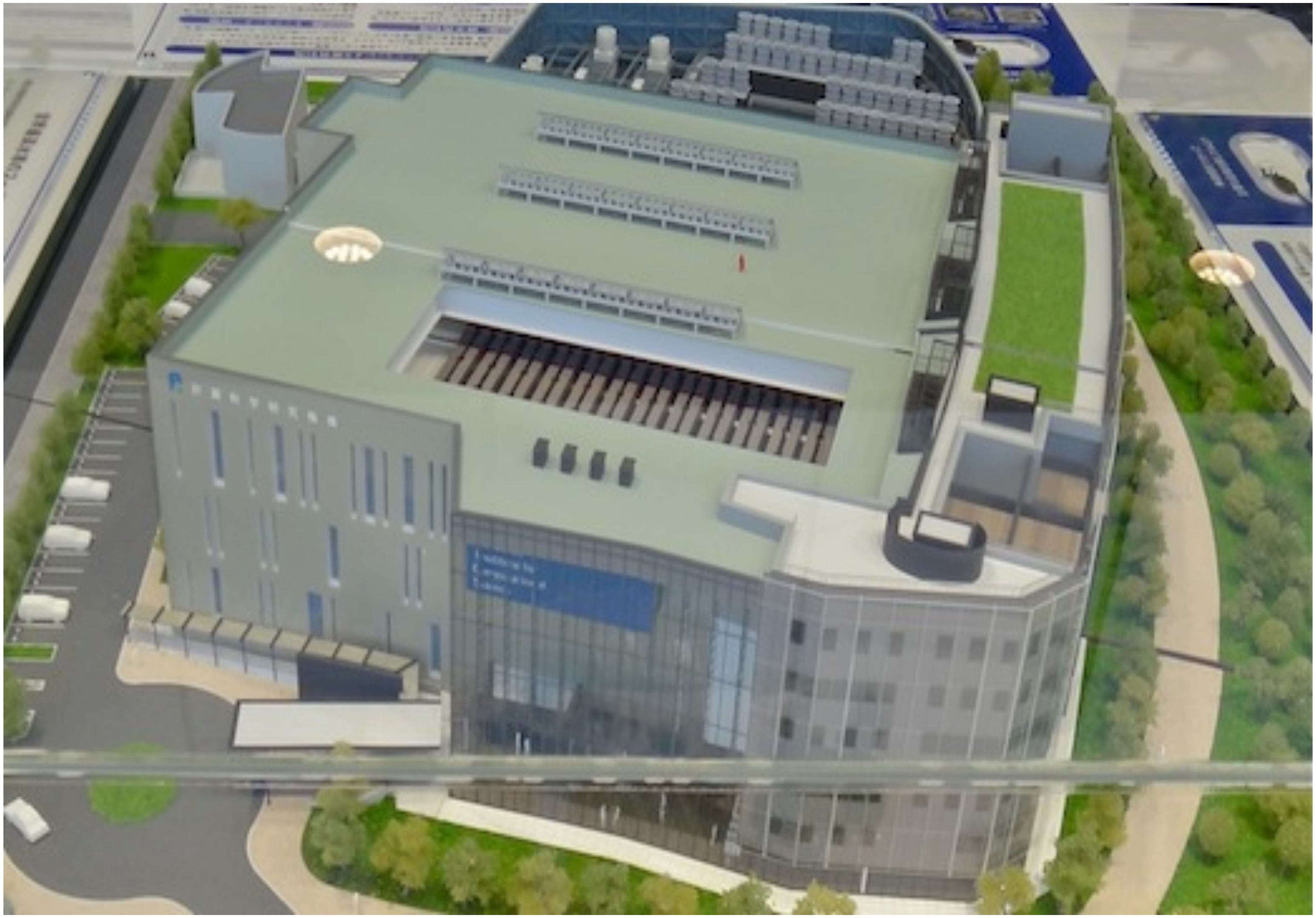
	Country	Institute	PF (Efficiency)
1	Japan	K. RIKEN Advanced Institute for Computational Science	10.51 (93.2%)
2	China	Tianhe-1A, National SuperComputer Center in Tianjin	2.566 (54.6%)
3	USA	Jaguar, Oak Ridge National Laboratory	1.759 (75.5%)
4	China	Nebulae, National Supercomputing Centre in Shenzhen	1.271 (42.6%)
5	Japan	TSUBAME2.0, GSIC Center, Tokyo Institute of Technology	1.192 (52.1%

FIRST to break the 10 Petaflop barrier in 2011

K Computer at Riken, Kobe is A Multicore machine

- 864 cabinets SPARC processors (around 700,000 of them)
- Annual running costs - $10 million
- Large energy consumption
- 820 Gflops/Kwatt
- Water cooling system minimizes failure rate and power consumption

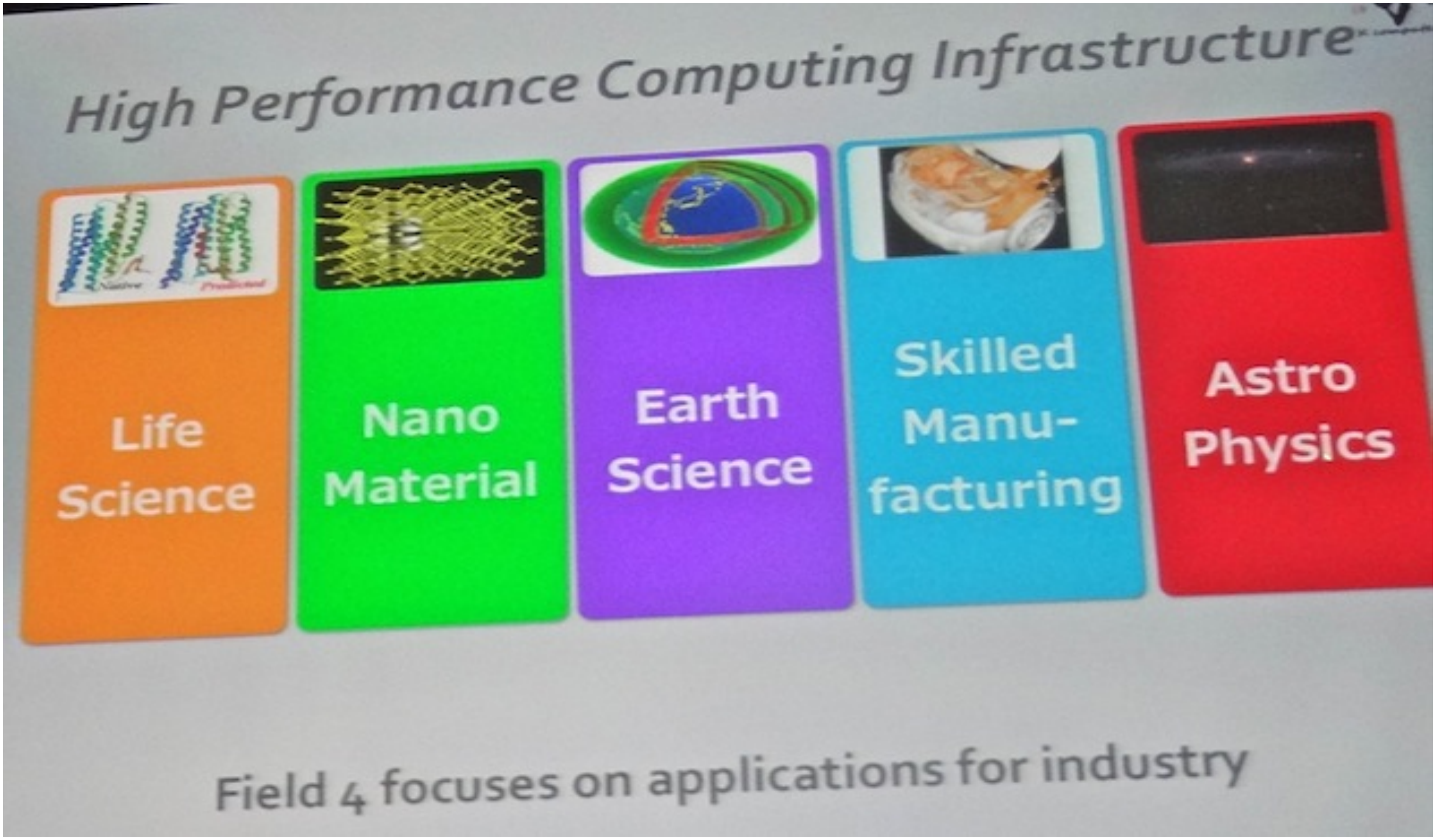
TORUS configuration network with over 700,000 cores



RIKEN COMPUTATIONAL SCIENCE BUILDING in Kobe, devoted to HPC for many disciplines

Advanced Institute for Computational Science（计算科学）, Kobe , Japan

- Operating the K Computer for Japan
- Leading-edge research in Computational science in many disciplines
- Promoting strong collaborations between computer scientists and scientists
- Plot and Plan Japan’s computational science strategy toward exascale computing

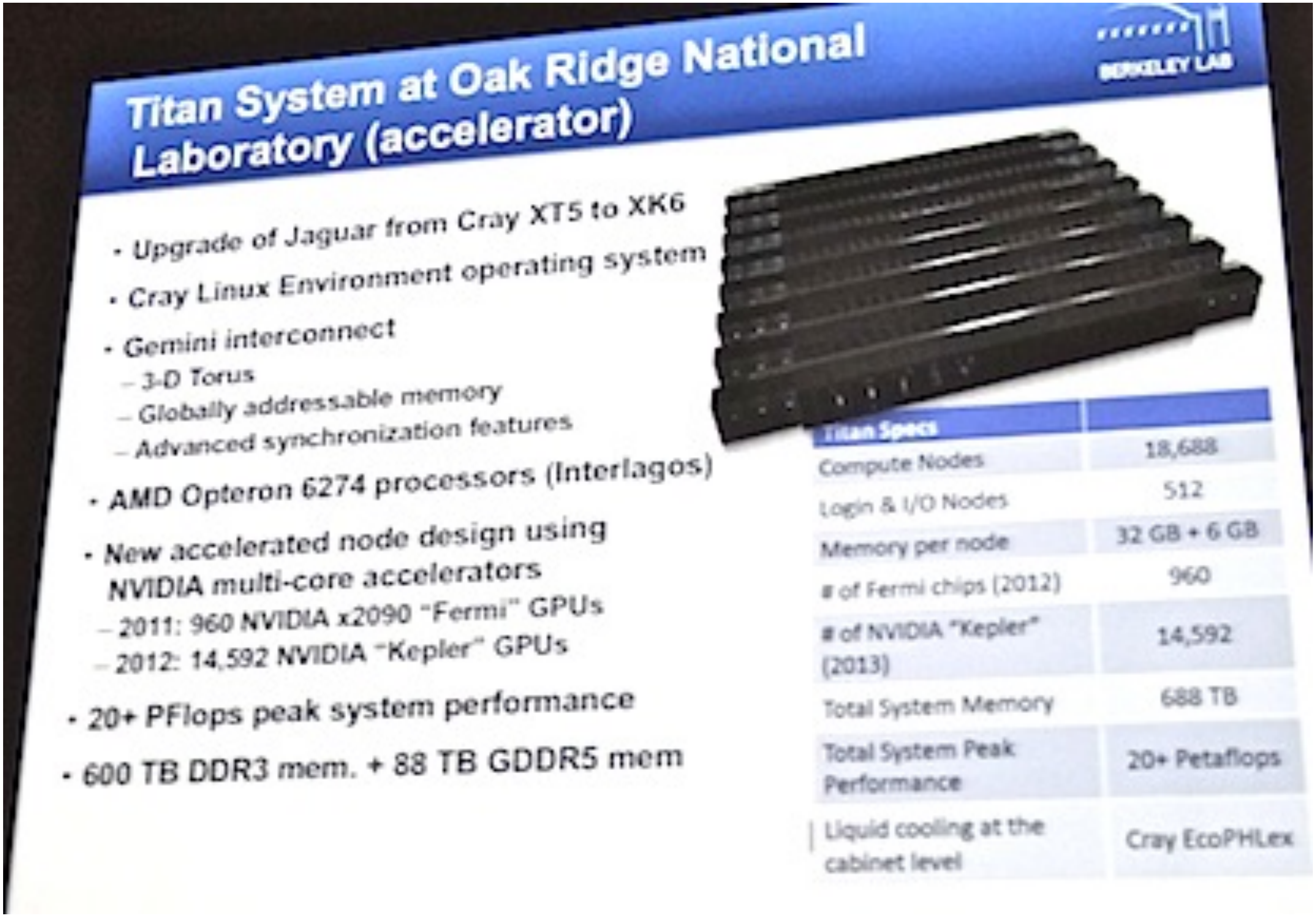


Major Aims of the Riken Institute in Kobe, Japan

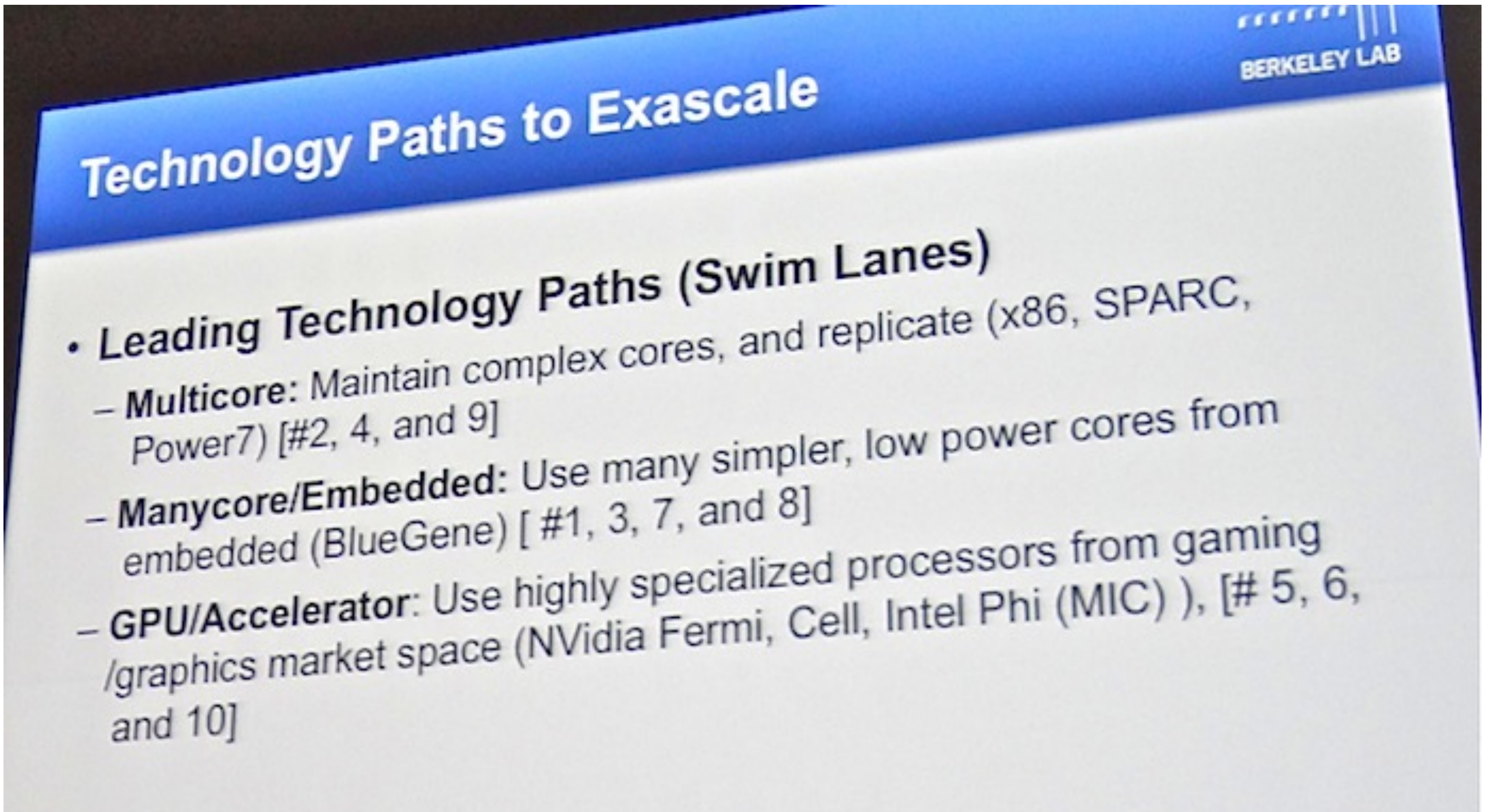
Horst Simon , from LBL, Berkeley is a leading proponent of **exascale** computing and has been organizing meetings on this topic since 2006.

There are many **exascale** projects started in different countries in the world, . Progress has been made, but many key **challenges（挑战）** remain, such as power consumption and stability of the system and benchmarking. How long you need to do this for getting a real speed measurement ? For the Sequoia at Livermore it took one day to get to 16.3 Petaflops ! This costs $18,000 of electric power consumption to get this result for the TOP 500 list.

Upgrade of Jaguar in 2013 to 20 plus Petaflops AMD + GPU



Different directions to reach Exascale 10** 18 Goal



HPC for Geosciences

- Non-linearity and Multi-scale Physics
- Supercomputing as a Third-Branch of Science
- Computing following Moore's Law since 1985.
- From 2005 on , GPU arrives on the scene
- Case study with AMR (Adaptive Mesh Refinement) but on a Desktop workstation (my friend Joe Scientist)
- Need for Higher Resolution Numerical Solutions and Visualization
- **Exascale** Computing is needed for breakthrough progress in realistic geophysical
- problems.

The notion of multi-scales and crossing of scales（跨尺度） has been around since the mid ‘90s, thanks to DOE.

- The laws of Nature are nonlinear and produce a multiplicity of scales in time and space for physical and biological processes. Quantum and classical mechanics and equations of general relativity by Einstein are intrinsically nonlinear, as well as biological equations
- Separation of scales（尺度分离） are sometimes possible analytically
- Scales are coupled over several decades:
- Much more challenging and are very difficult

What can HPC do for multiscale geoscience problems？

- Equally spaced（等间距）grids
- Unequally spaced grids, such as finite-elements, nodes of polynomials (Chebychev)
- Meshless（无网格） schemes, radial-basis functions,arbitrarily distributed（任意分布）. Wavelet based methods
- Curse of dimensions（维度祸根）
- In one dimension 1D, we can now do problems with spatial-scale difference of at least 10**10 points or more
- In three-dimensions, 3D ,we can go perhaps to 10**4.7 points in each direction, but really pushing it

But this is what Joe Scientist needs to see for new discoveries（新探索）

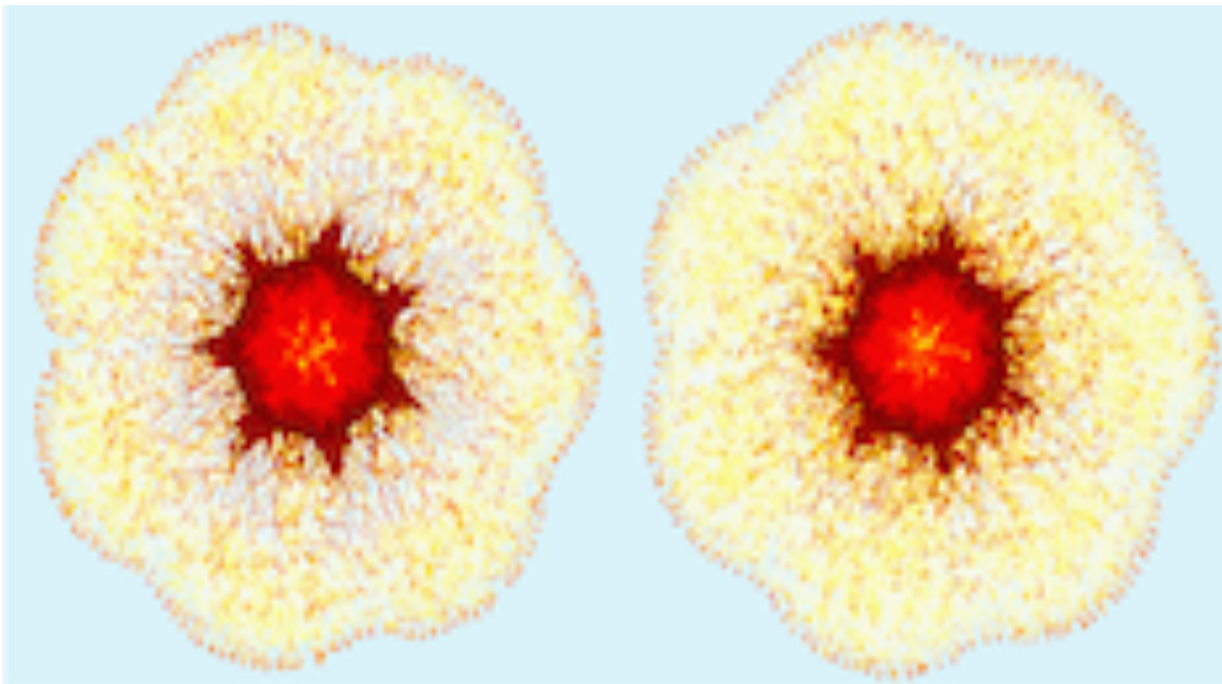


Figure 2. Density distributions in PPM multifluid simulations of a 2-D cylindrical implosion test problem, derived from that of Youngs [2], on grids of 6144^2 (left) and 12288^2 cells (right). A perfectly symmetrical boundary condition has driven the compression of a 20-times denser cylindrical shell with 1% and 0.5% amplitudes of sinusoidal modes 140 and 147. Mode 7 has arisen from the nonlinear amplification of the beat frequency, and has been followed through a 10-fold compression.

From P R Woodward

Right panel has a grid of 12400x12400,

Left panel has a gird 4 times less points

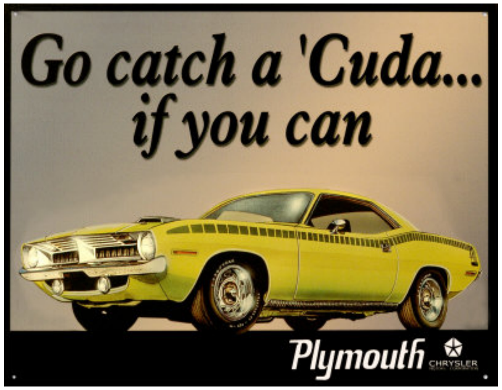
Moore’s Law

- The number of transistors that can be placed inexpensively on an integrated circuit doubles approximately every two years. This trend has continued until around early 2000’s, then it has flattened out since . In 2005 **GPU(图形处理器)** started to take hold. But Joe Scientist does not know this or does not want to know.

NVIDIA GPUs

CUDA: Compute Unified Device Architecture

- the most popular GPGPU parallel programming model today



- from notebooks and personal desktops to high performance computing
- the GPU with many cores runs the kernel concurrently by many threads
- two-level hardware parallelism on a GPU: SIMD and MIMD
- a programming model reflects the hardware parallelism: blocks and grids

NVIDIA GPUs
Architectures and compute capability

G80 (since 2006): **compute capability(计算能力)**1.0, 1.1
features 1.1: 8 cores/multiprocessor, single-precision real arithmetic

GT200 (since 2008): compute capability 1.2, 1.3
features 1.3: 8 cores/MP, max. 30 x 8 = 240 cores, first double precision

Fermi (since 2010): compute capability 2.0, 2.1
features 2.0: 32 cores/MP, max. 16 x 32 = 512 cores,
fast double precision, cache
features 2.1: 48 cores/MP, max. 8 x 48 = 384 cores,
slower double precision

Kepler (since 2012): compute capability 3.0
features 3.0: 192 cores/MP, max. 8 x 192 = 1536 cores
slower double precision

CUDA Developer Tools

Languages and directives by Portland Group

PGI (company has been around since 1987) compiler suite

PGI Workstation, Server, CUDA Fortran, CUDA-x86, Accelerator

current version 12.5 (5th minor version in 2012, 10 versions in 2011)

for Linux, Mac OS X and Microsoft Windows, 32/64 bit

with OpenMP, MPI, parallel debugger and profiler

with ACML and IMSL libraries, linkable with Intel MKL

with Eclipse IDE (Linux) or Microsoft Visual Studio (Windows)

http://www.pgroup.com

How much has Moore's Law helped Joe Scientist since ? 1990 ?

- In 1990 Joe could do a FEM（有限元法） calculation
- with 10**4.5 degrees of freedom in his office , if he was extremely lucky.
- Today thanks to Moore's law Joe can do a FEM calculation with 10**6.5 degrees of freedom in his room, which must be air-conditioned.
- So an increase by a factor of 100 in 20 years, Is this progress, Joe ?

For Joe Scientist relying just on a desktop workstation

- The scales he/she is able to cross are much more restricted by 4 orders in magnitude in 1-D.
- And around 3 orders in magnitude or less in 3D
- In other words,today he/she is set back at least 20 years by using just a desktop workstation to explore parameter space of the past of 1992 vintage

What is Joe going to do now in 2012?

- Joe can turn to GPU, or many-core computing , or now MIC from Intel
- Joe can go to meshless schemes,
- hierarchical schemes like AMR or wavelet-based methods, while still keeping his workstation because he wants to remain comfortable.
- Or Joe can use HPC , (petascale computing)

Joe Scientist is Taras V. Gerya from E.T.H., Zurich

Joe(Taras)'s aim in life is to look at variable viscosity problem in mantle convection

Problem has many scales from plates on the order

Of 10000 km to faults with 10-25 m width, a scale of

five and half orders in magnitude

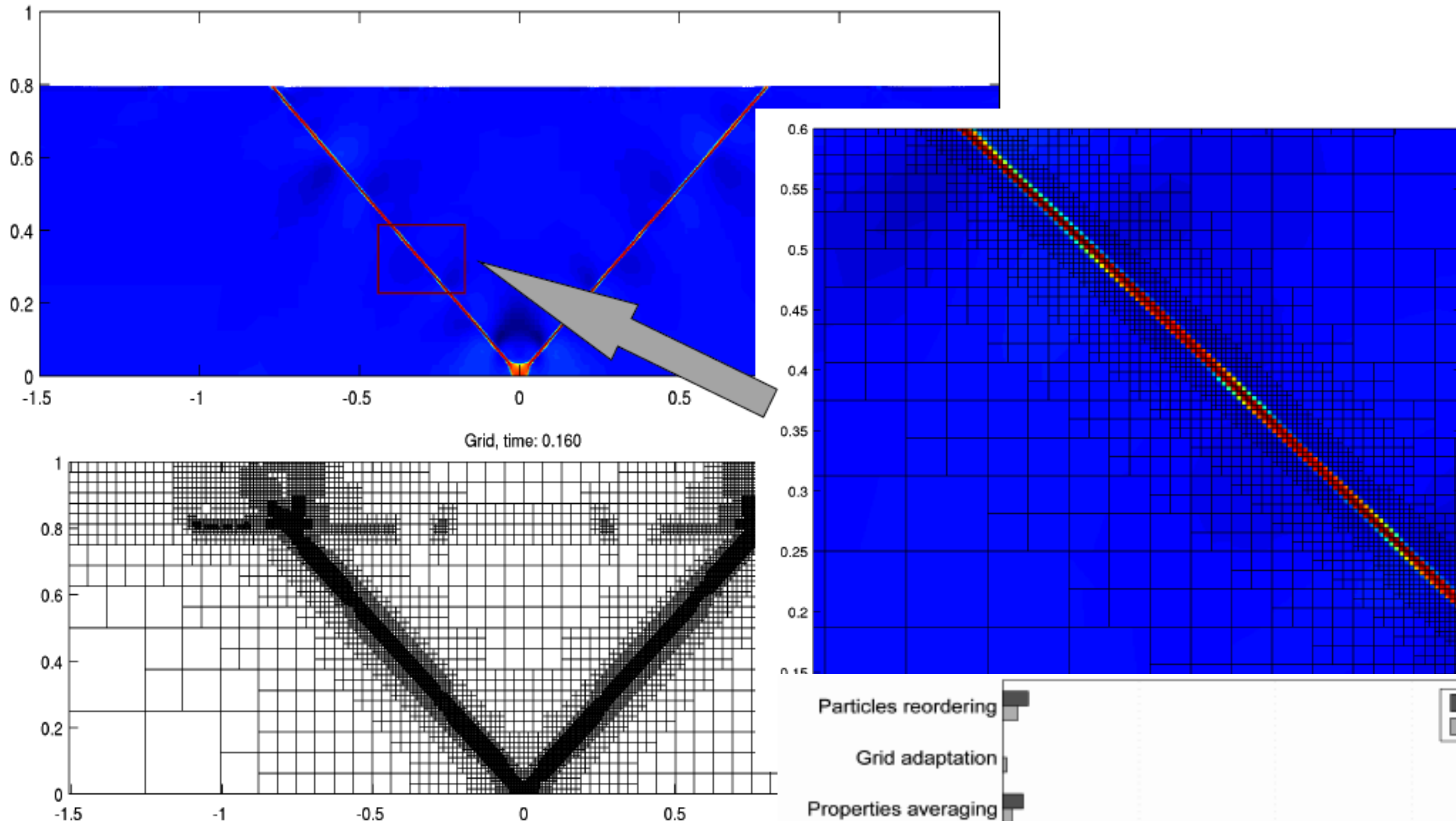
very similar to tsunami wave problem.

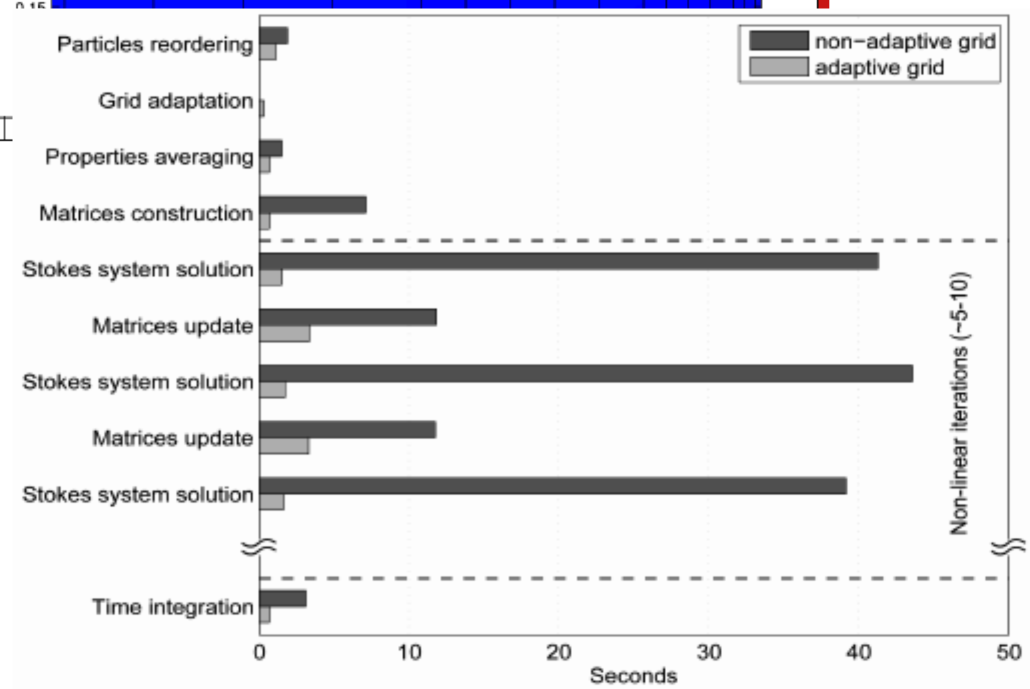
Now Joe has sought refuge in Adaptive Mesh Refinement,

first used by astrophysicists in the 1980's

Adaptive Mesh Refinement (AMR)

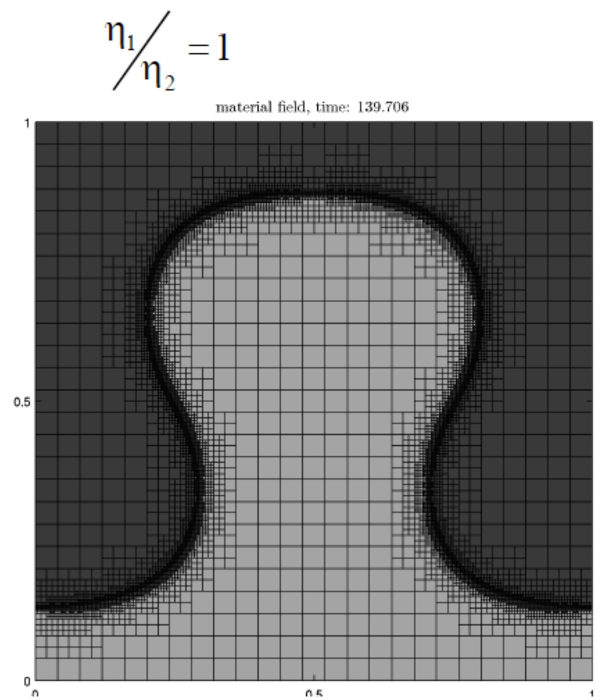
- Hierarchical Method（层级法）, (tree-structure)
- Uses information about the local behavior of the solution (gradient)
- And refine the mesh on the spot
- Time-stepping is constrained by the smallest mesh
- Used recently for steady-state situations in mantle dynamics and ice sheets (Omar Ghattas , Univ.Texas, Austin)

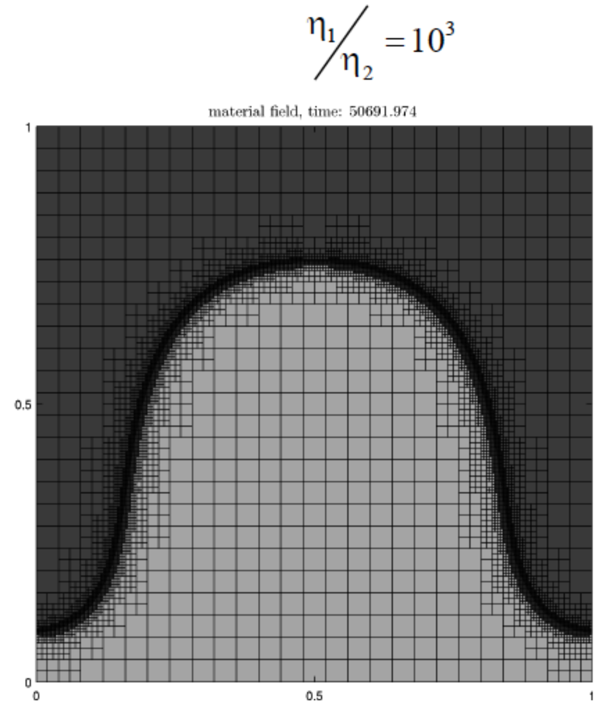




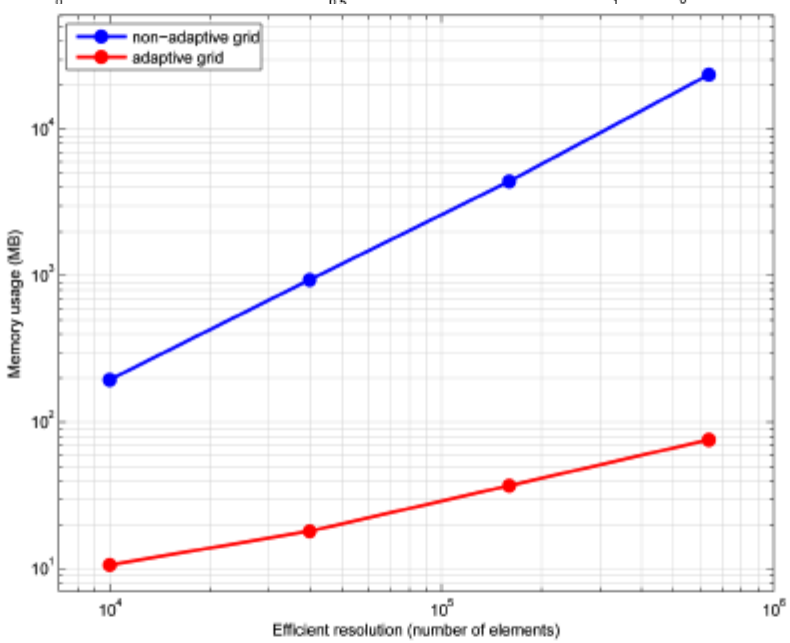
FEM-AMR for Stokes problems with large viscosity contrast, spontaneous localization and free surface (Q1Q1 and Q2P1 elements)

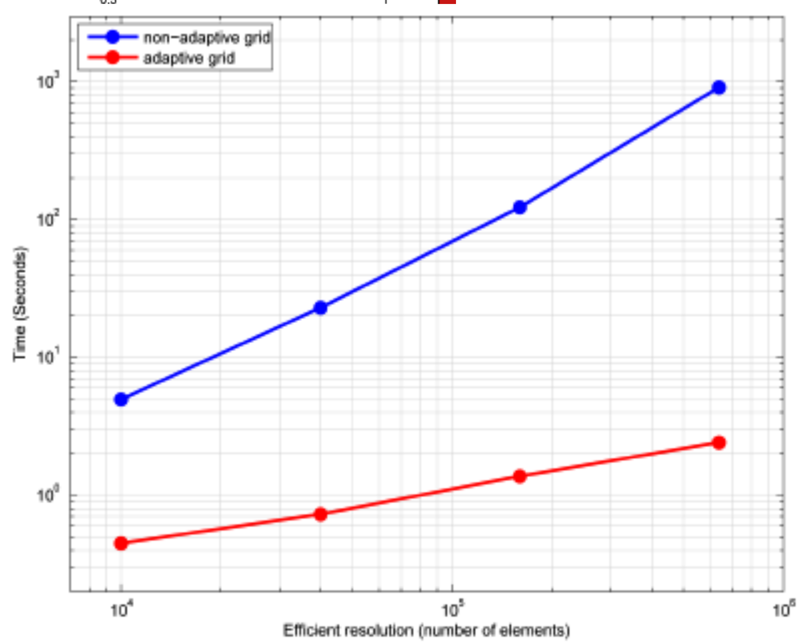
Gerya, G. ,F.M. May , T. Duretz, 2012



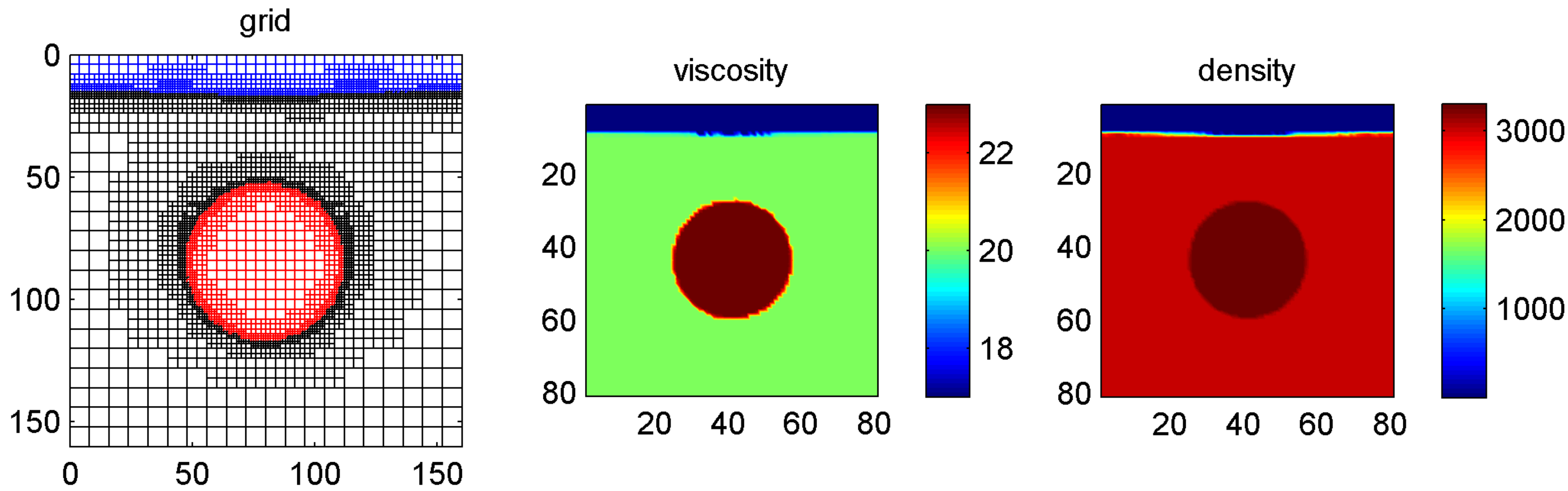


FEM-AMR for Stokes problems with large viscosity contrast, spontaneous localization and free surface (Q1Q1 and Q2P1 elements)

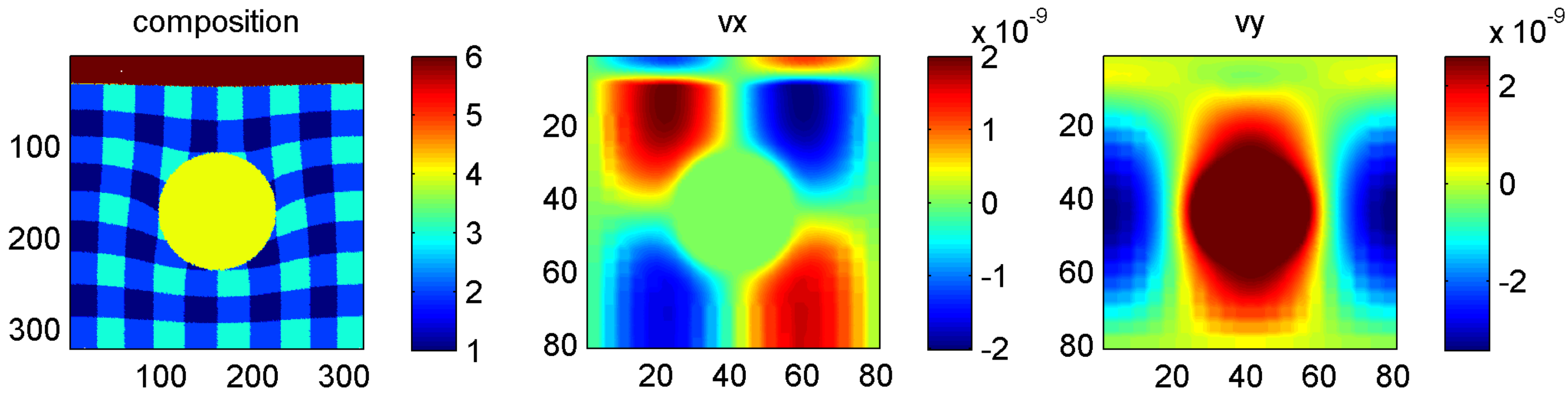




Gerya, T.,F.M. May, G**3, 2012



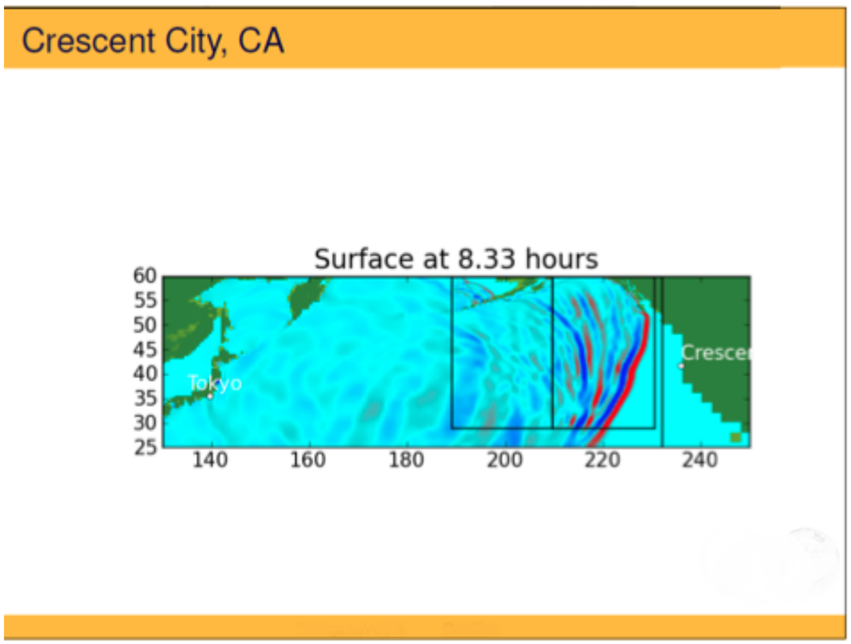
FDM-AMR on staggered grid
for Stokes problems
with large viscosity contrast and free surface
(conservative finite differences and marker-in-cell)

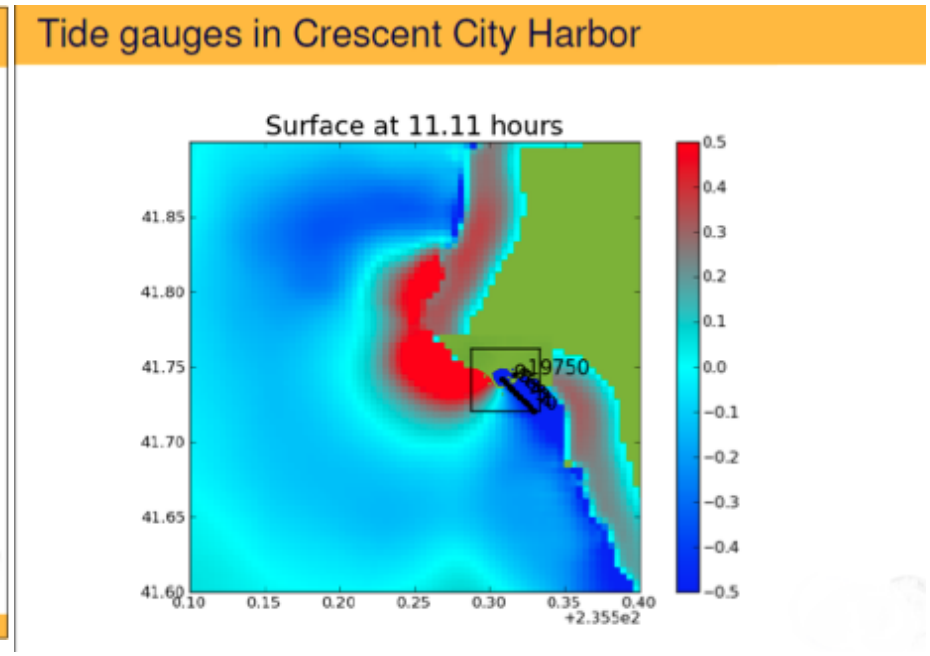


Gerya , T., F.Me. May. (G**3, 2012)

AMR on Tsunami wave propagation

FROM TOHOKU Earthquake

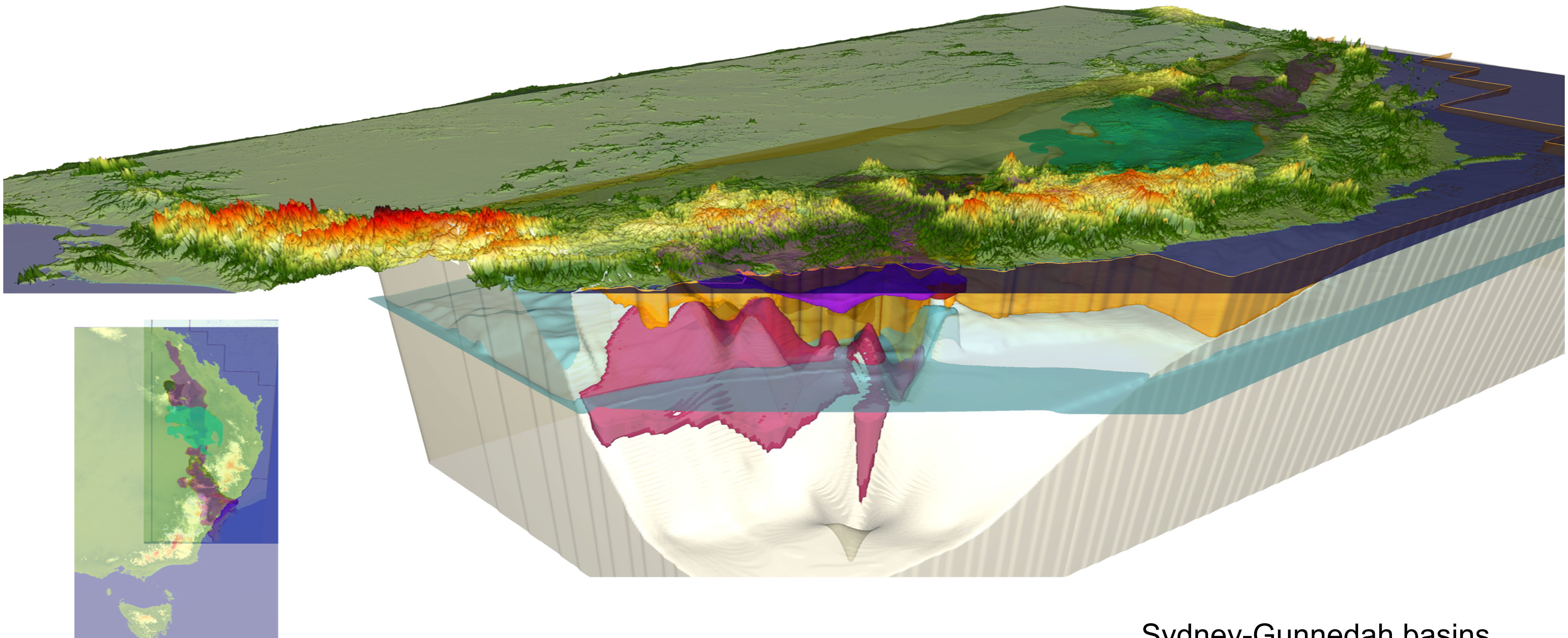


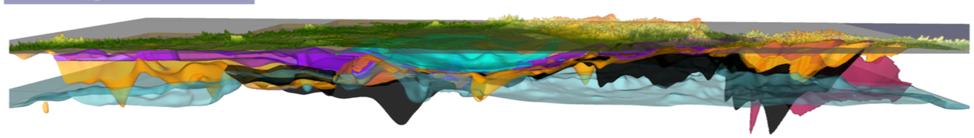


Up to 5 different levels were used in this trans Pacific effort by David L. George, USGS

Sydney-Gunnedah-Bowen

A thermal model of the east-coast basin system of Australia
… methods enabled by peta-scale computing (and beyond)





Sydney-Gunnedah basins
600km x 300km x 12km
based on initial work in Danis et al. 2010

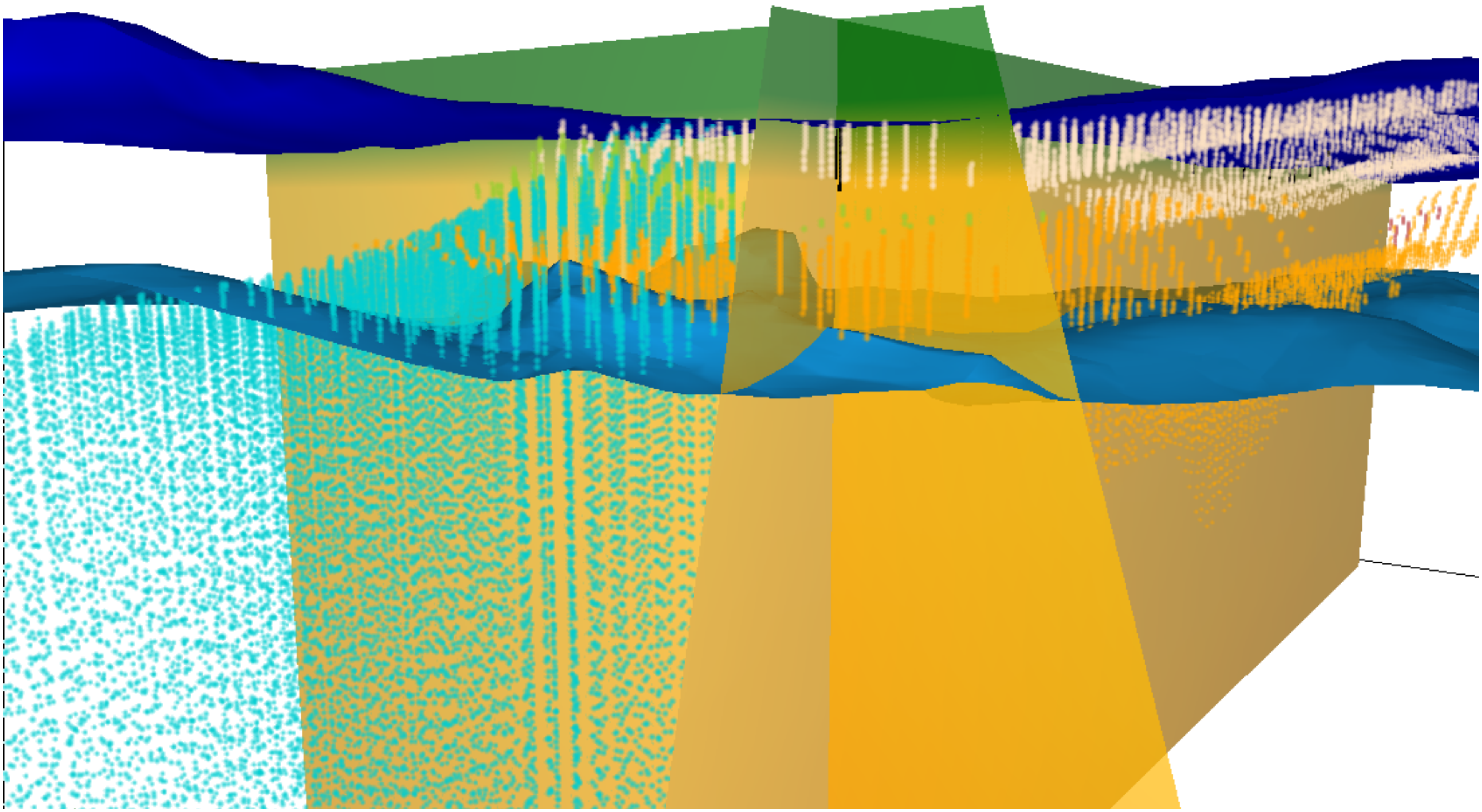
A collaboration between - Steve Quenette, Cara Danis, Owen Kaluza, John Mansour, Craig O'Neill, Louis Moresi

Computational Geophysics

- Use *geodynamics* (e.g. 3D, anisotropic, heterogeneous heat flow) in the *inversion process*
  - Why? assimilate more data with models (physics)
  - Assessment is in terms of confidence
  - Obtain the heat flow / temperature at depth for free

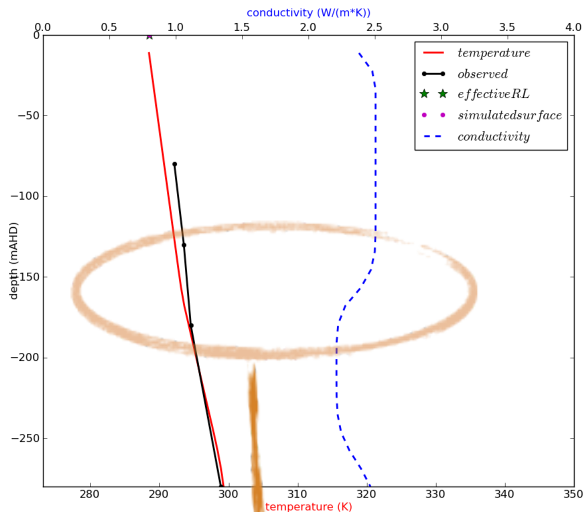
- Some needs:
  - Very high spatial resolution
    - (capture basal boundary (MOHO?) but also deal with coal in the order of 1meter)
  - Very short wall-clock time on the supercomputer
    - (doesn't matter how many cores we use – each realization must be in the order of 1 minute)
  - Import GoCAD, GeoModeller, …

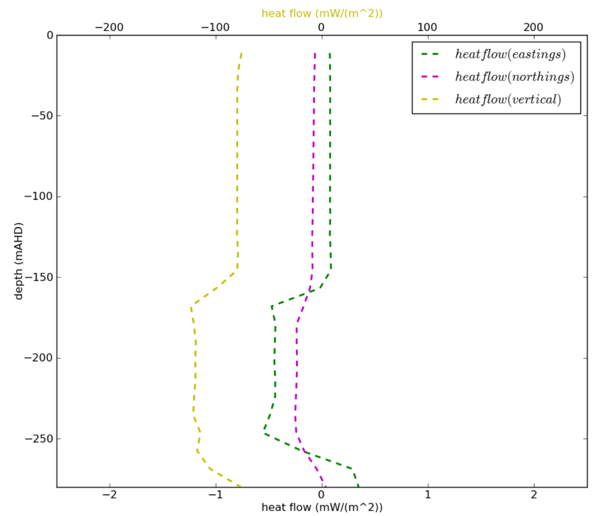
Ambi-dextrous: particle or mesh frame



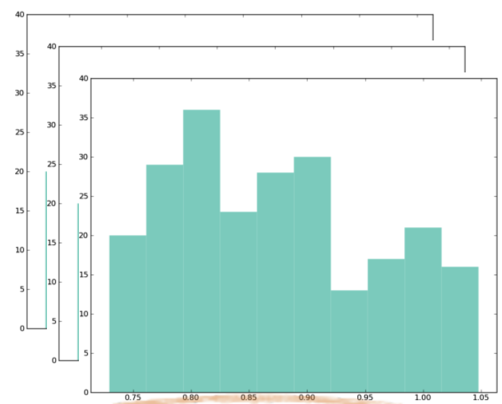
A collaboration between - Steve Quenette, Cara Danis, Owen Kaluza, John Mansour, Craig O'Neill, Louis Moresi

Numerical Experiments…





- Using 10s of drill cores of 5+ observables each. Most are shallow. Some are deep.
- Resolution: at least same order as ave vertical observation
  – 128proc @ 3GB/core: takes 2-4mins (e.g. above)
  – 1000proc desired



... Over 1000s of realisation

Petascale enables this!!!

HPC is now moving beyond the capability of a single programmer, even **Joe Scientist** (T.V. Gerya) must now seek other resources for help

on HPC platforms

We must use software **Libraries（图书馆）** to go further

PETSc from Argonne National Lab has been around for 21 years and is mature enough to support non-linear partial differential equations on massively parallel computers, even GPU,

See next talk by Matt Knepley

Others include Portland Group Inc. for GPU（图形处理器）and MIC（迈克）.

Other problems in geophysics also require considerable computational resources

(1.)Mantle convection, Crustal Dynamics and geothermal resources(Interdisciplinary areas)

(2.) Hybrid Inverse（混合反演）problems involving different complimentary physics, acoustic waves generating electric signals. Data assimilation and seismic imaging

(3.)3-D Seismic wave propagation in heterogeneous media for energy resources and hazard mitigation

(4.) Mineral physics, Density functional theory and molecular dynamics of earth materials, liquid and solid states. Surface mineral reactions.

(5.)Tsunami waves and Landslides, Floods, tidal bores in rivers.

INFRASTRUCTURE（基础建设） FOR PROMOTING PROGRESS IN HPC IN GEOSCIENCES

TRAINING FOR YOUNG SCIENTISTS

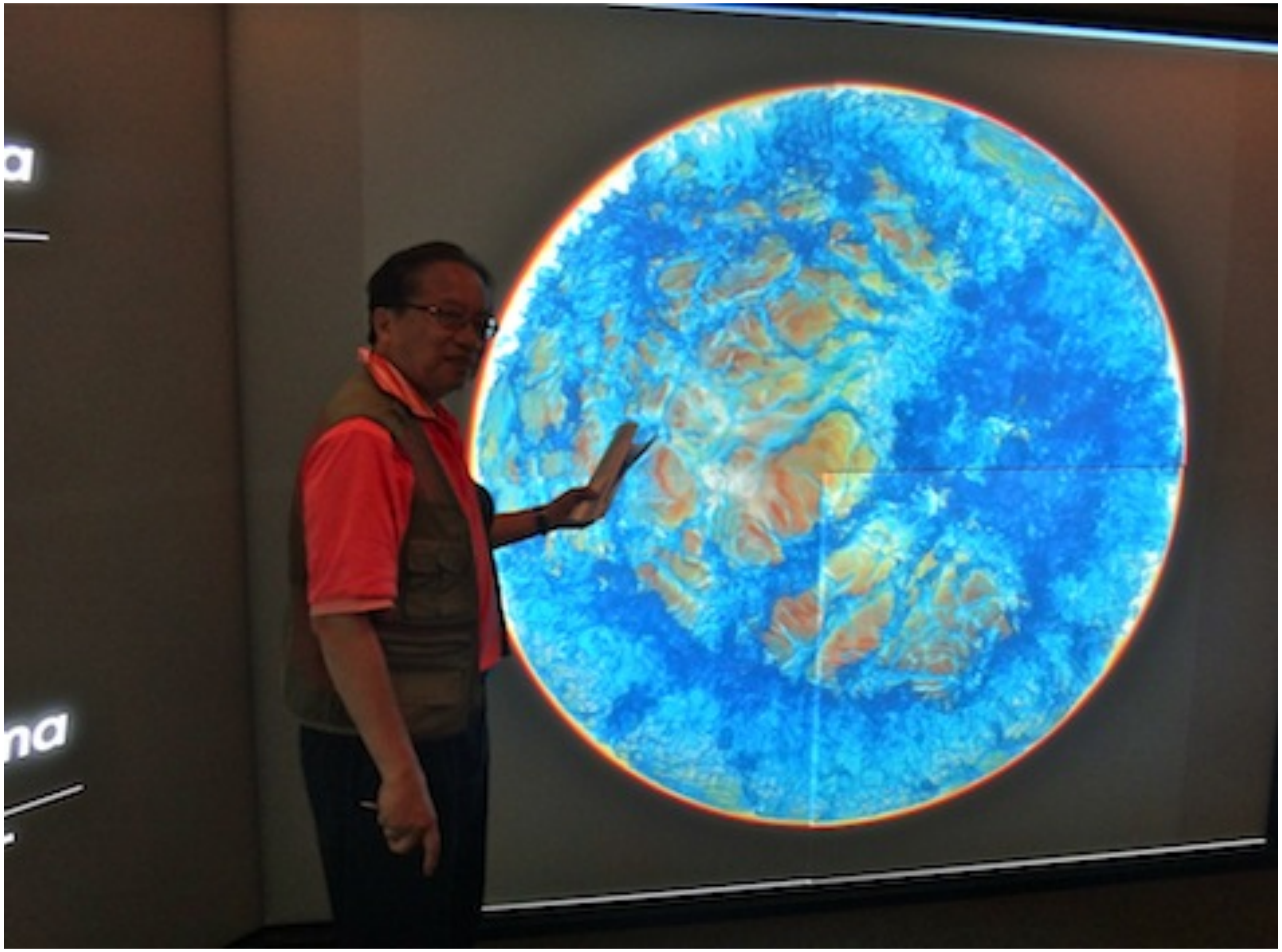
SUMMER SCHOOL, SHORT COURSES

A PLACE FOR PEOPLE TO MEET AND SPEAK

Riken, Kobe

Japan

Need to **Visualize（可视化）** Numerical Simulations of Time-Dependent Dynamics , goes hand in hand in HPC



D. Yuen in Laboratory of Computing. Sciences. Engineering, University of Minnesota

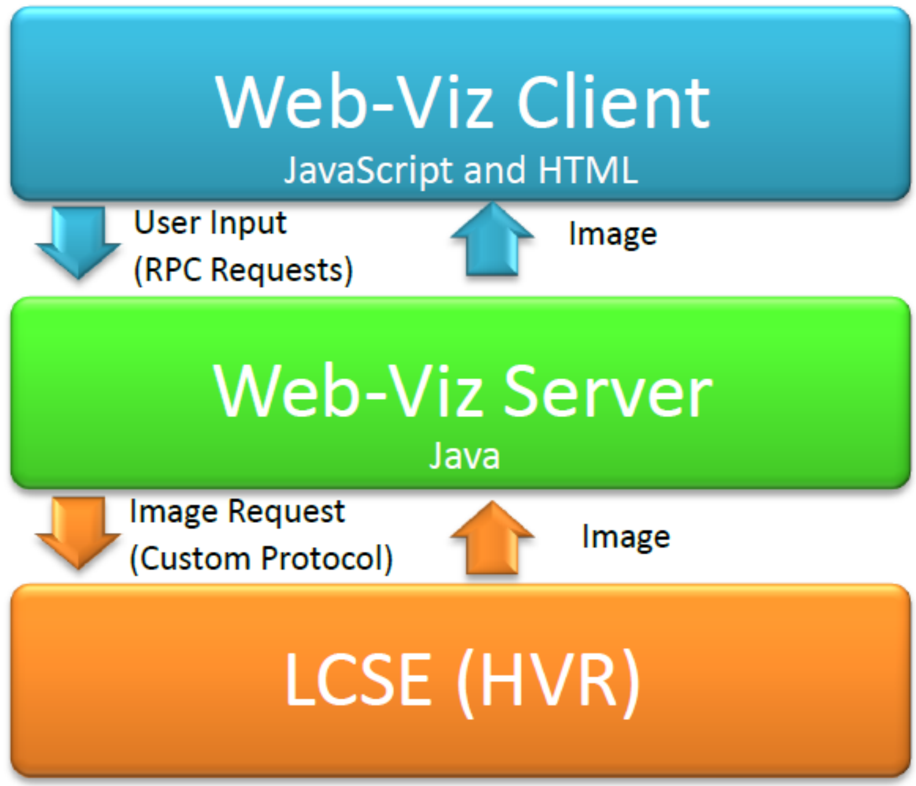
Hardware Hierarchy

Web-Viz includes 3 layers of applications:

1 Client layer gathers user input and returns server response ,using standard internet protocol.

2 Server layer provides interface services and act as a media between visualization server and Client layer.

3 HVR layer processes and renders data and return visualization result to above layers.

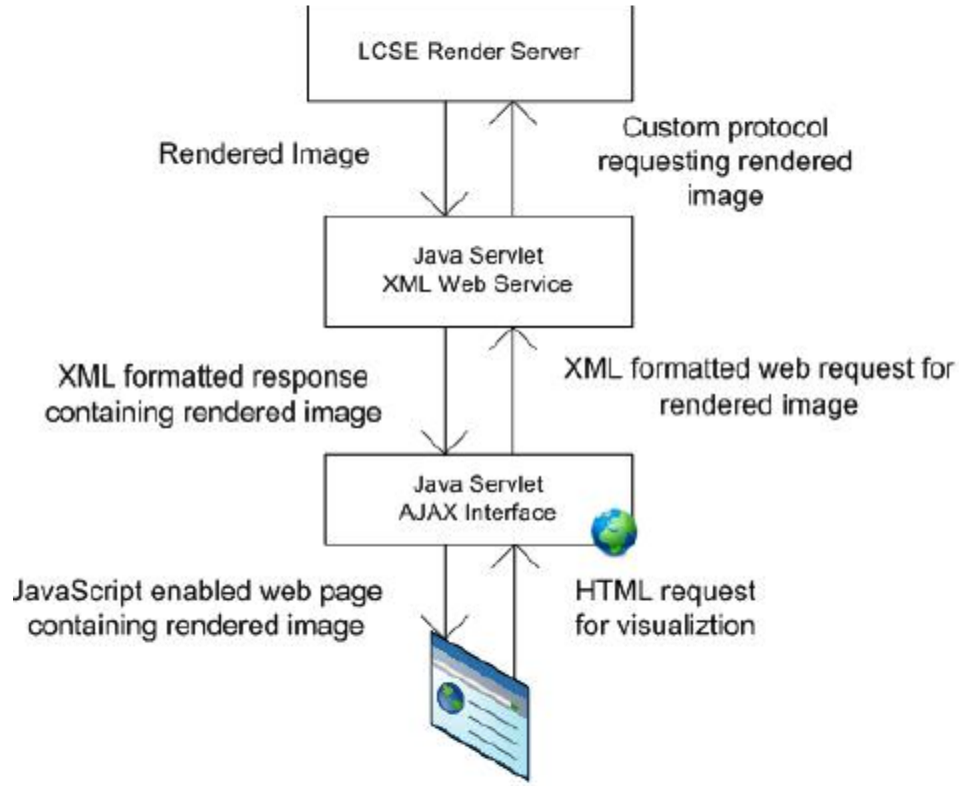


Software Hierarchy

On the client layer, JavaScript is used to provide in-browser interface.

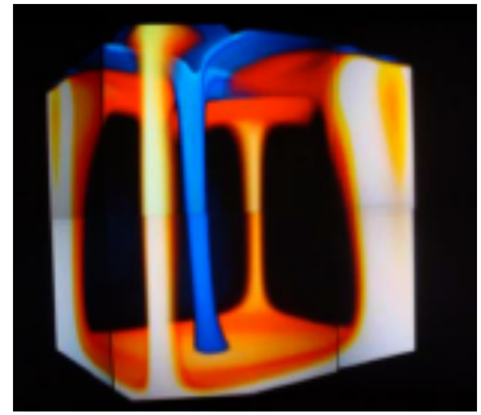
On the server layer, we use XML web service to formatted request from user and pass it on to the lower layer.

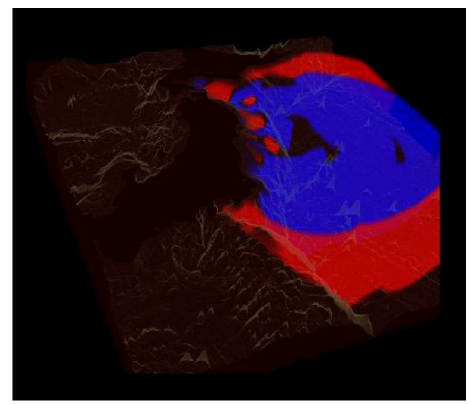
On the rendering layer. We apply Hierarchical Volume Rendering (HVR) to process raw data.



Examples from Mantle Convection, Tsunami

- Numerical simulations generate a lot of data for a 500 x 500 x 500 grid, and 1000 time steps. This can lead to tens of Terabytes of data. So we need to visualize this on the fly .Examples, we have used come from 3-D mantle convection and tsunami wave modelling of the Tohoku-oki earthquake from 2011.

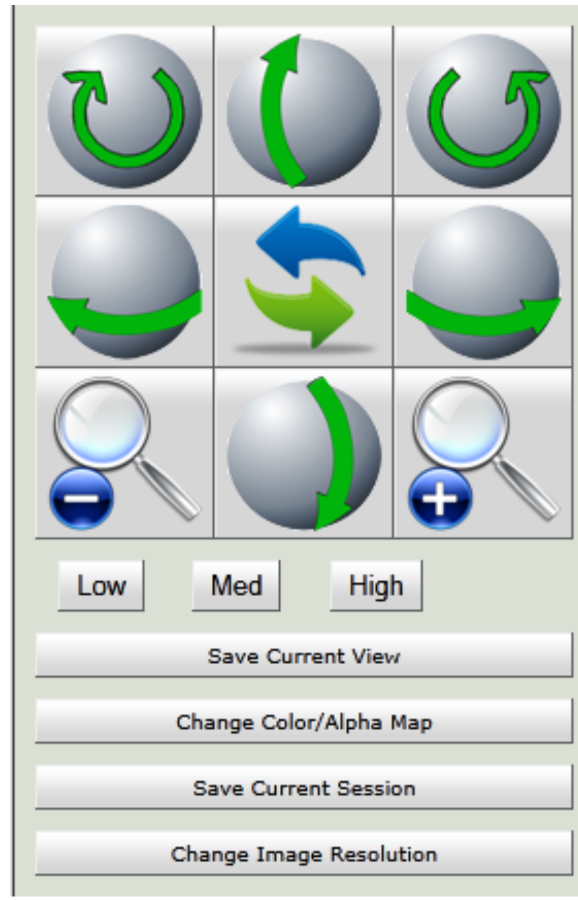




Interactive and Collaborative Features

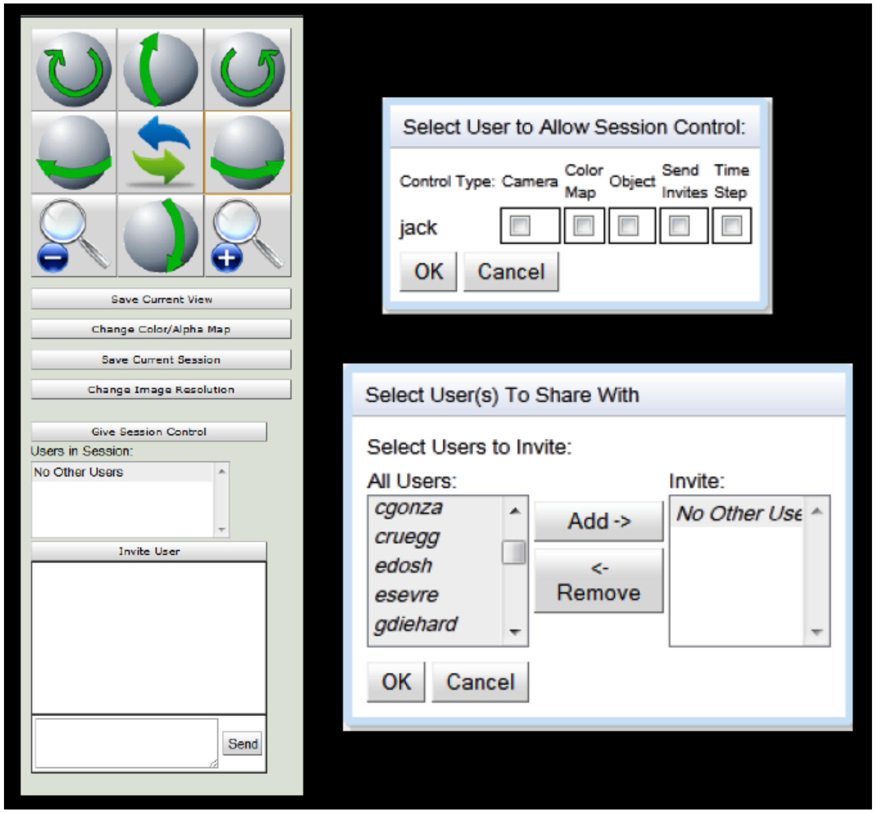
- User Interface provides camera position and orientation relative controls, color and opacity set-up ,and user authentication control.

- User authentication control provides different permission for different types of users within the visualization session.



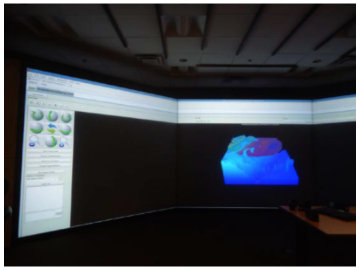
Interactive and Collaborative Features

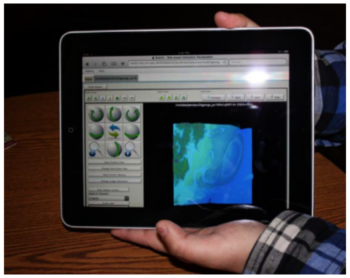
- Session can be shared among multiple users. Users within the same session can collaborate using chat Panel.
- In default setting, control panel is locked to users other than session owner. Session owner can release control to designated users.
- Control panel and other session control function can be released partially to maintain stability of the session.
- Control and Collaborative Panel can be used seamlessly on different platforms with different input equipments



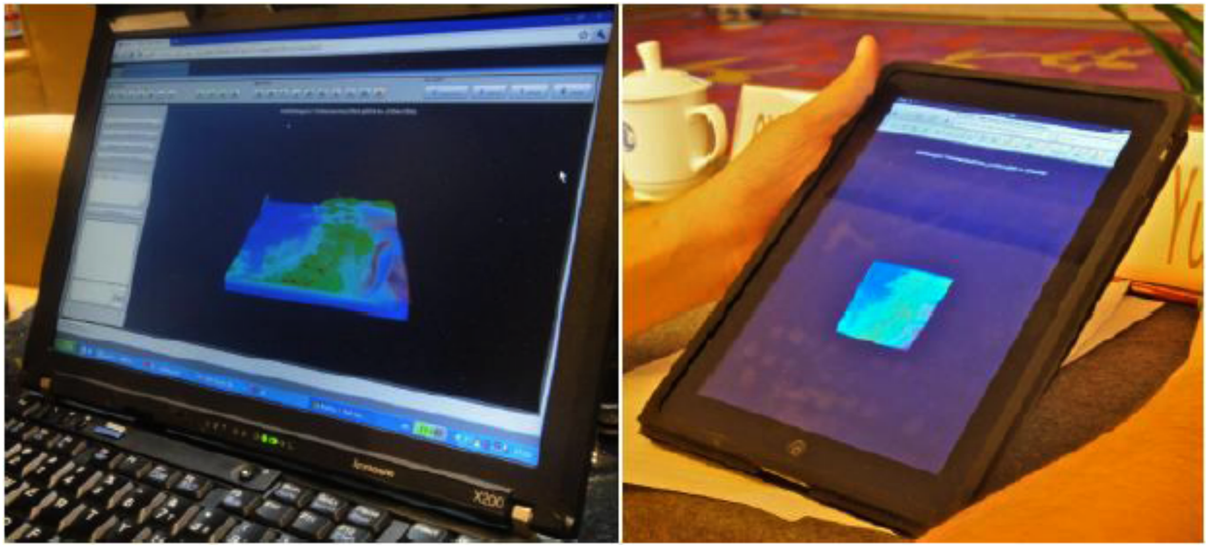
Multi-device Capability

- WebViz can be used on variety of platforms, ranging fro large PowerWall, Tablet, PC and smart 4G phones.
- Visualization session can shared among different platforms. Users can set-up resolutions to fit their platform.
- Multiple sessions can be run simultaneously with parallel rendering server.





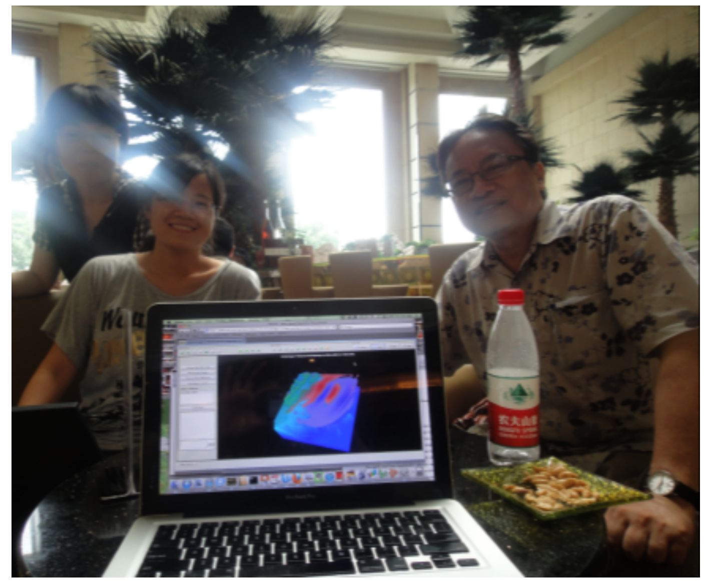
Experiments carried out globally



We have tested Web-Viz’s collaborative features on several remote locations around the globe that far away from University of Minnesota. These locations include **Harbin** (9,300km from U of M), **Beijing** (10,200km), **Shanghai** (10,900km), **Lanzhou** (10,800km), **Kiev** (8,100km) and **Perth** (17,100km).

Global Experiments

The visualization under examination is the result of GeoCLAW modeling of the Sendai tsunami waves, showcasing negative (blue) and positive (red) displacement of the water surface, as well as seafloor topography. The rendering speed , when all things are considered, from Beijing was only a factor of 2 to 2.5 slower than at University of Minnesota.



Collaborative visualization test at Beijing, China

High-resolution runs involving 2500x2500x2500 points involving 4 fields and 2000 time instants

at 4 bits would require close 100 Tbytes !

Storage of data is a problem.

Making streaming videos would be one solution

Doing things in multi-resolution (different scales)

would be another solution to reduce the amount of data.

We need to get computer scientists , applied mathematicians visualization experts involved.

Data compression. Compressive Sensing(sparse

Data representation, been around since 2004)

IN ORDER TO DISCOVER NEW PHENOMENA WITHIN THE MULTI-SCALE SCHEME OF THINGS

(1.) USE HPC to the utmost and not be satisfied with running jobs on your desktop, unless it is a GPU or a cluster of Digital Signal Processor (DSP)

(2.) Visualization goes hand in hand with large-scale numerical simulations of time-dependent flows. Matlab is not **up to speed（加速）**, we need fast volume rendering, together with fast edge-detection multi-resolution algorithms

like wavelets or curvelets, again **GPU and/or MIC would help in these directions**

WHAT LIES AHEAD IN THE NEXT DECADE?

(1.) We will still have Petascale for another 6 to 8 year .Today we are at 16.7 Pflops (Sequoia machine from Livermore National Lab. In California)

(2.) What are needed will be new developments in software and algorithms.**Software（软件）**holds the key in petascale and exascale efforts , if new science is to be discovered. We need to make compilers easier to use for common users. Wider library usage, such as **PETSc**, for many-core computing. They will help even Joe Scientist to use more cores.

(3.) Visualization will always remain a challenge .It needs to be both fast and **collaborative（协同）**in order to make some new science. Networking among scientists (**Open-Source** policy) holds the key for new discovery.

4. EFFORTS IN EXASCALE COMPUTING WILL ALSO BRING GREAT BENEFITS TO LOWER RANKS IN COMPUTING Hierarchy AS WELL, SUCH AS LAPTOPS WITH **MIC（迈克）/GPU**,DESK-TOPS and CLUSTERS OF **CPU-GPU-MIC**. NETWORKING WILL ALSO BE FASTER.

5. A VERY BRIGHT FUTURE IN COMPUTATIONAL GEOSCIENCES IN THE NEXT DECADE.

THANK YOU

谢 谢！