High Performance Computing in Seismic Data Processing : Promises and Challenges

Presented By

Mrs. Richa Rastogi

Team Coordinator
Seismic Data Processing,
Scientific and Engineering Computing Group,
Centre for Development of Advanced Computing,
Pune University Campus,
Pune – 411007, India
Email: richar@cdac.in

Website: http://pune.cdac.in/html/secg/sdp.aspx

Contents

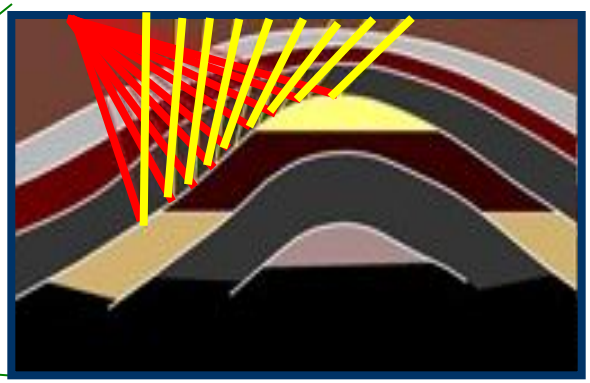
Contents

Geo-Science: Objective

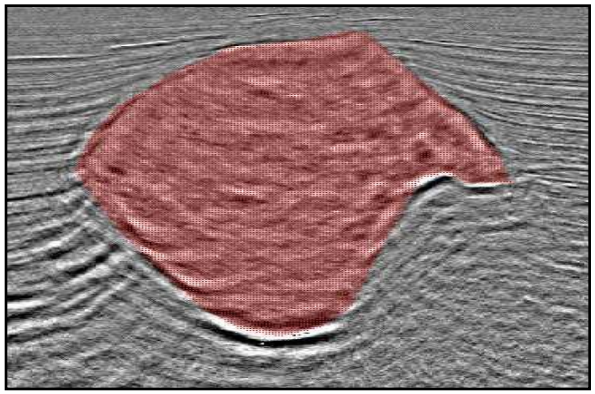
Aim of Geo-Sciences Studies

Geo-Scientists

Planet Earth



Seismic Survey



Subsurface Image

To Understand and Image Earth at higher & higher Resolution

Geo-Science: Application

Oil

Gas

Metal and Minerals

Earthquakes

Infrastructure

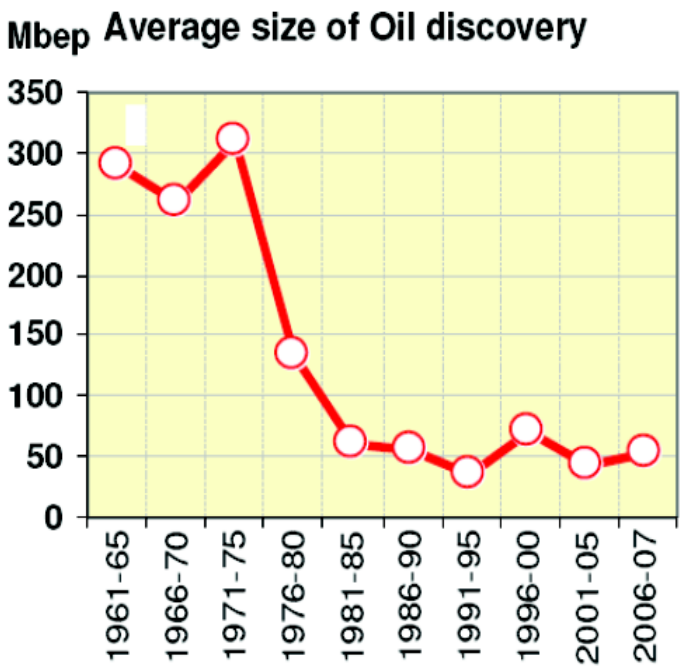
World's Economy

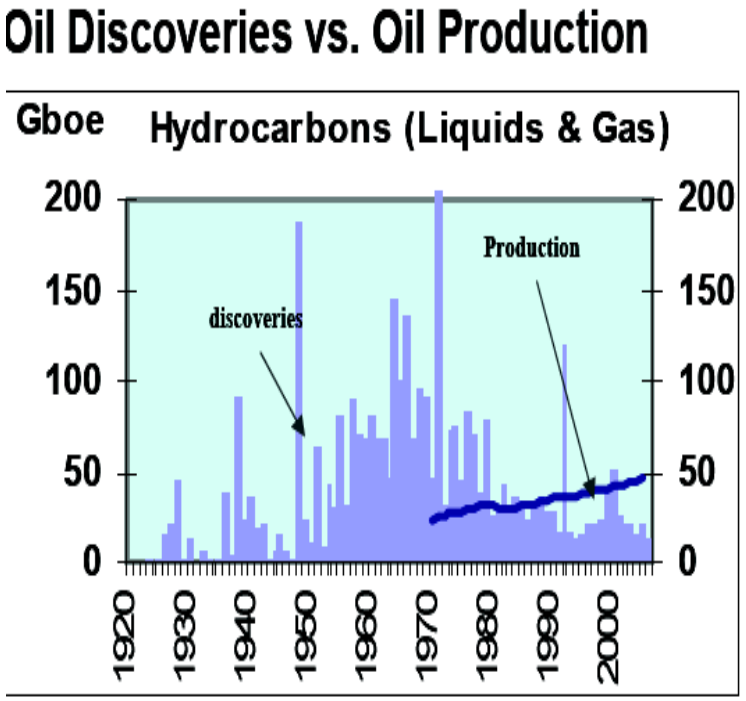
Daily Life of a Common man

Contents

Oil & Gas Sector: World Scenario

Oil Discoveries and Production



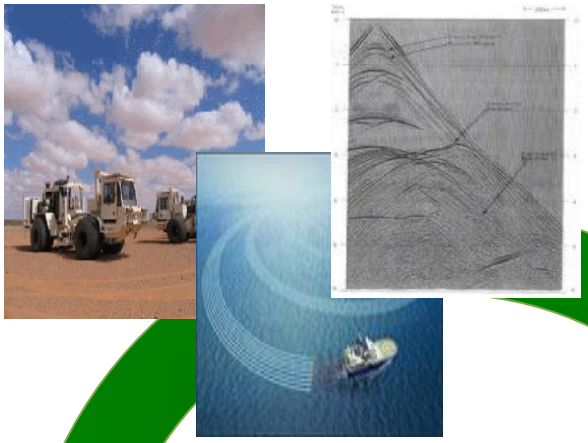


Worldwide O&G consumption/year approximately 50 Gboe

Average size of new oil field discovery ~ 50 Mboe since the last 25 years

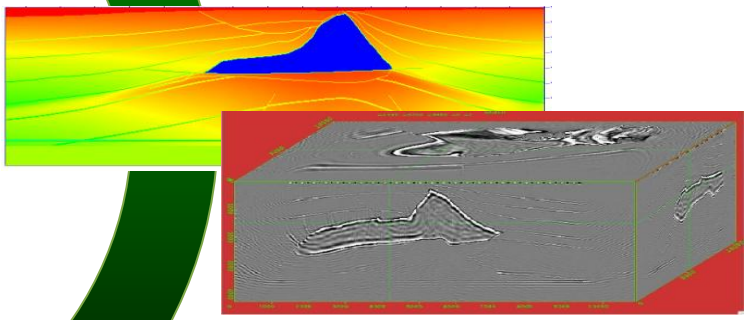
(Source: Figures projected by Total :- oil exploration and production company)

Seismic Data Processing & Hydrocarbon Exploration Cycle

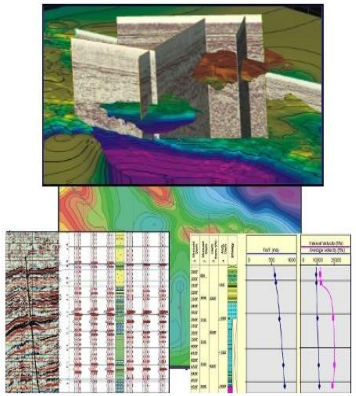


Data Acquisition,Storage & Pre-Processing

Data Post-Processing & Imaging loop

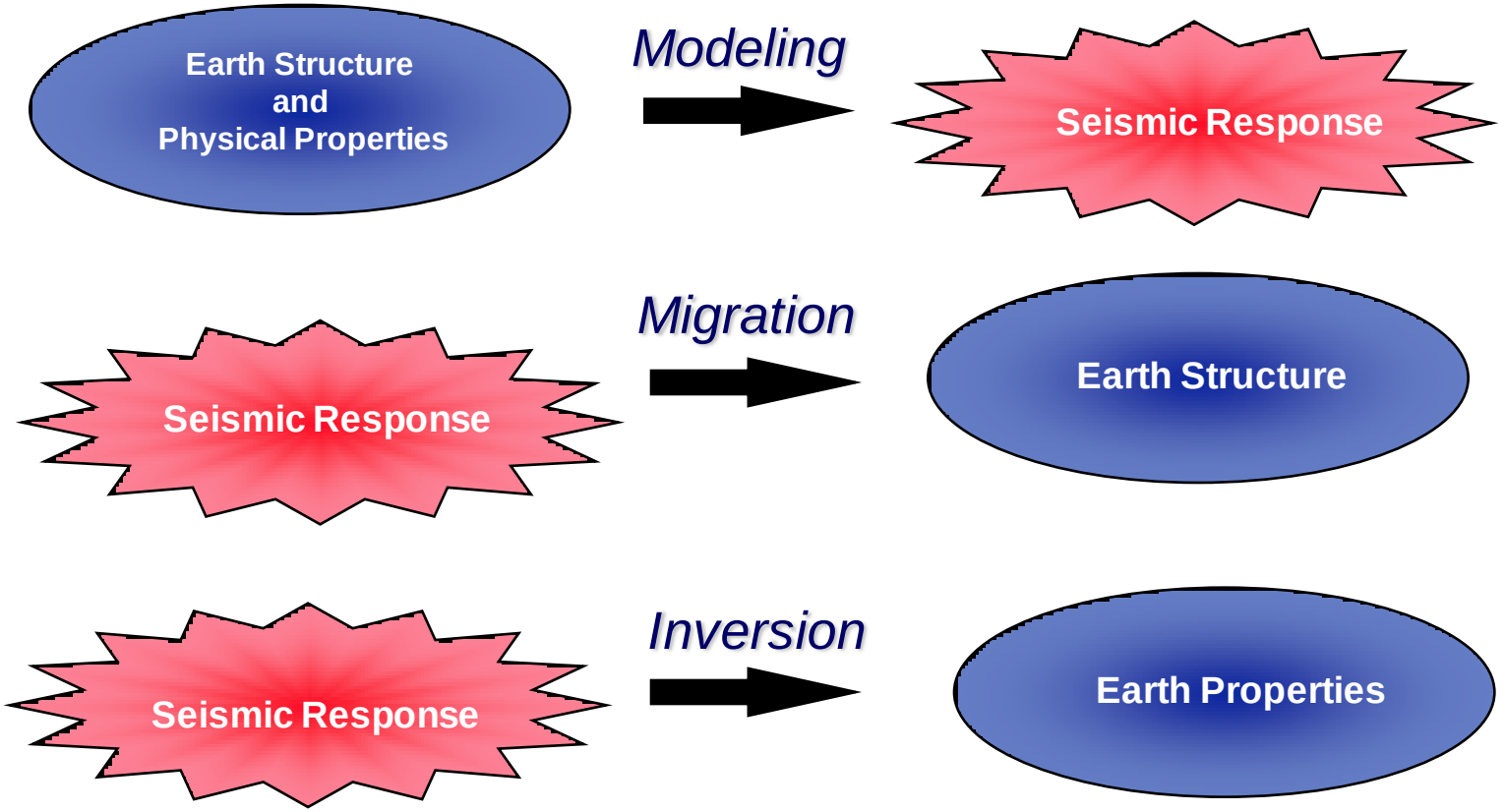


Data Interpretation & Visualization



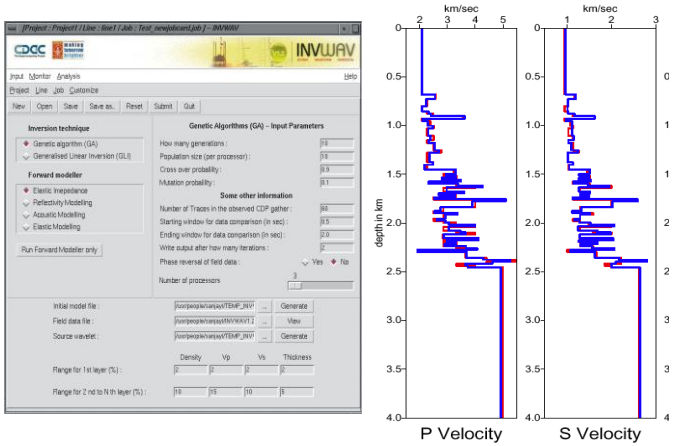
Oil Production & Reservoir Management

SDP Applications at C-DAC

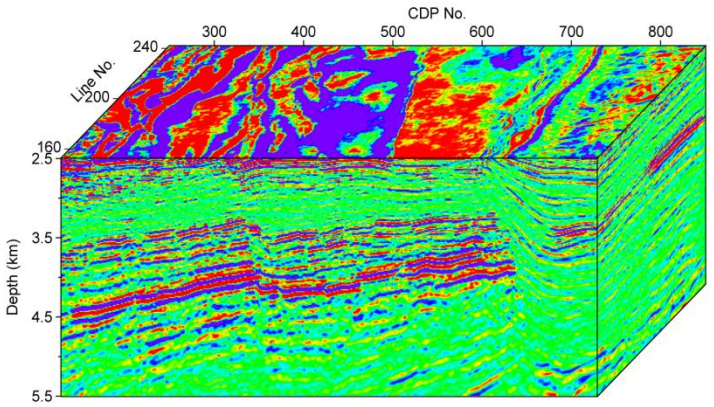


Compute and Data Intensive

SDP Projects at C-DAC

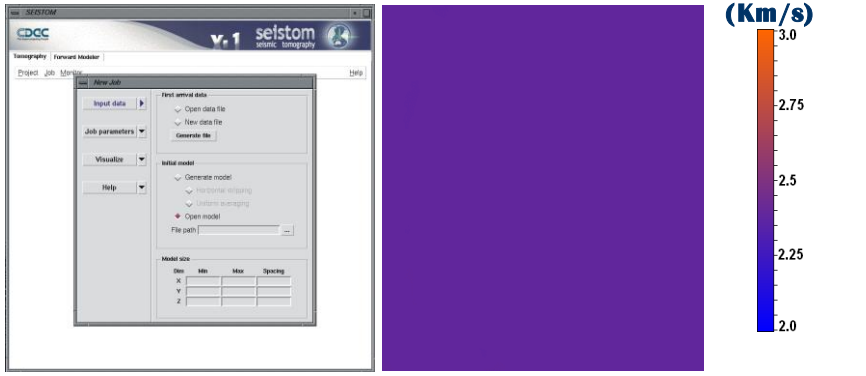


Seismic Full Waveform Inversion (INVWAV)



Seismic Migration and Velocity Analysis(WAVES)

Seismic Modeling (WAVES)



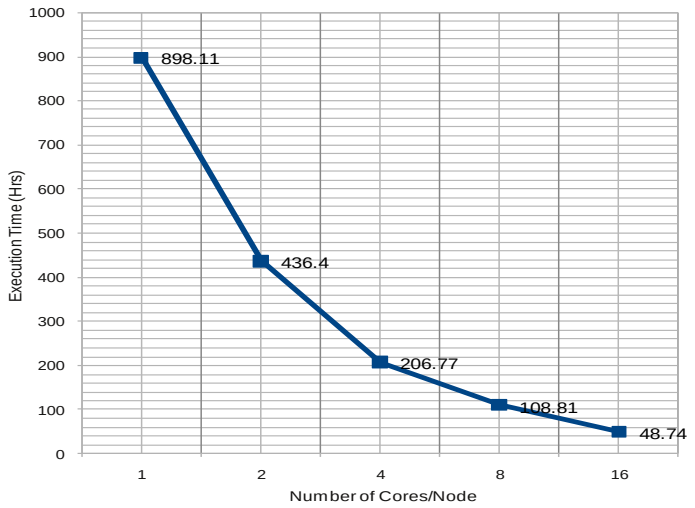
Seismic Tomographic Inversion (SEISTOM)

Full Waveform Inversion (INVWAV)

- Pre-Stack seismic data is used for sub-surface property estimation
- Reflectivity Modeling is used for calculation of forward response of earth models
- Real-Coded Genetic Algorithm is used as guided random search optimization tool

GA Parameters	Runs	5
	Iterations/Run	100
	Population/Node	334
	Unknowns	1050
Data Parameters	Number of CDPs	1 CDP (Overthrust model)
	Record length, sampling rate	2.8 secs, 0.008 secs
	Traces/CDP	80

Execution Time of INVWAV for 40 Nodes of PARAM Yuva



Rank	298 (as of Nov 2010)
Speed	54 TFlops
Cores	4608
Processor	Intel EM64T Xeon 73xx (Tigerton) 2930 MHz (11.72 Gflops)
Interconnect	Infinband DDR 4x

Methodologies & Techniques

Problems	***Algorithms***	***Parallelization***	***Solution***	***Methods***
Modeling	*Solving wave equation (diff. Eq.s)*	*Dividing Problem space*	*Finite Difference Schemes with Boundary conditions*	• *2D Elastic* • *2D & 3D Acoustic*
Migration	*Solving wave equation (diff. Eq.s, integral equation)*	*Dividing data to be processed*	• *Frequency,* • *Wave number* • *Time domain*	• *Kirchhoff* • *W-X* • *PSPI*
Inversion	*Global and local optimsation & Seismic Data simulation: Forward modelling*	*Dividing search space as well as data*	• *GA and GLI* • *Forward Modeling (EI,RM,AM,* • *EM,Acoustic Modeling, TT tomography, 2D Depth migration)*	• *Real/Binary Coded GA* • *Damp. factor optimization* • *Tomographic inversion* • *Migration Vel. inversion*

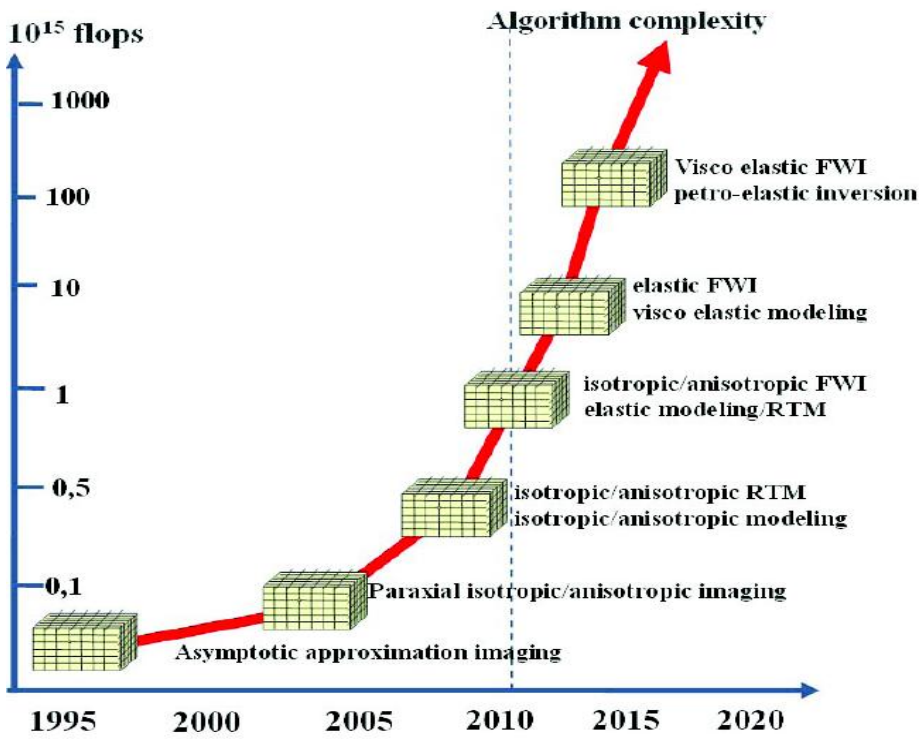
Large SDP Projects world-wide

- Hyperbeam Project by PGS – RTM in 10 minutes
- Roadrunner Project at Los Alamos National Labs
- The Repsol Kaleidoscope Project in collaboration with the Barcelona Supercomputing Center, 3DGeo and CSIC
- Sam Williams at the University of California, Berkeley

Contents

SDP: HPC Requirements

- Algorithms complexity
- Larger number of algorithm to improve S/N
- Amount of data sizes, storage and retrieval
- Demand of higher and higher resolution
- Larger and Deeper area of survey
- Fast Processing & Interpretation
- Interactive and 3D/4D Visualization
- Exploration and Exploitation Management



Algorithmic complexity and corresponding computing power

Reduction in time cycle from survey to drilling and increase in production

(Source: Figures projected by Total :- oil exploration and production company)

SDP: Major HPC Challenges - A Geo-Scientist's Perspective

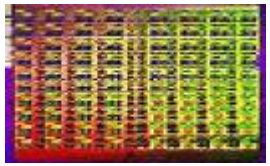
- Porting and scaling of legacy algorithms across different hardware platforms
- The data transfer along with memory and process management across powerful computing systems
- In-memory handling of Peta bytes of data for processing, interpretation & visualization
- Quick delivery of data through the caches to the processing units from remote location
- Need of high resolution visualization and real time rendering of computed results in a interactive fashion
- Failures due to the large number of components in a single HPC system
- Re-engineering large number of algorithms to cater to new computing paradigms
- Computing sensitiveness to environment and community

SDP: Major HPC Challenges - A Geo-Scientist's Perspective

The real HPC challenges today are to link all the compute units together (efficient networking), to transfer large data volumes to and from the processing units (efficient data access), and to process data in small chunks distributed over many processing units (efficient algorithms)

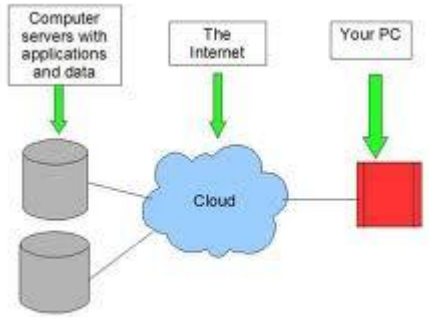
HPC Promises: Emerging Technologies

SMP's and MPP Cluster

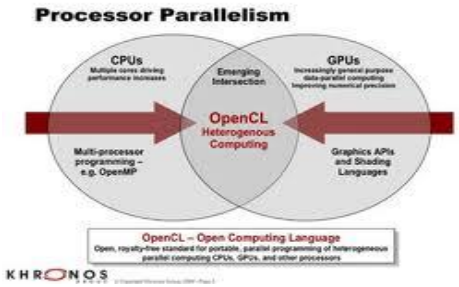


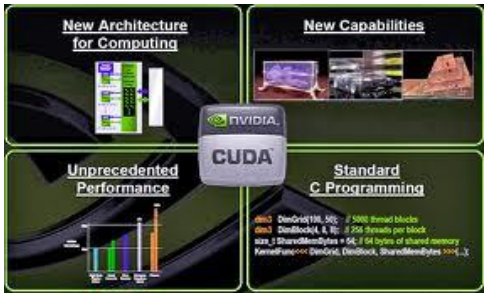
Multi-core Processors

GPUs and FPGA's



www.irunabiz.com





HPC Promises: Accelerate R&D and Production

Hybrid (GPU's and Multi-core together)

- GPU and multi-core clusters can be a good option
- OpenCl can be used for such type of clusters, effectively
- A hybrid cluster on one interconnect can offer new programming paradigm for application acceleration in a environmental friendly way
- Certain issues need attention:
  - Porting efforts
  - Standards of programming
  - Memory managements
  - Bandwidth and Latencies for large clusters
  - Auto handing of hardware failures
  - Power and heat aware job schedulers

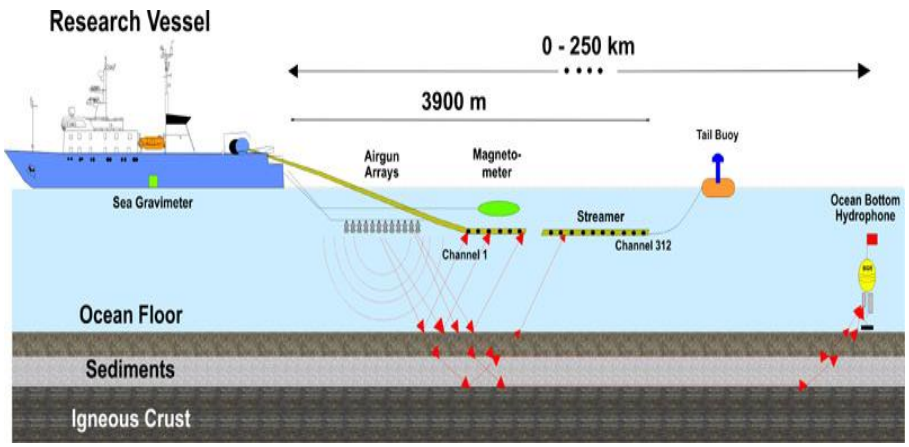
Such kind of system and technology should be made available for use in a shared way through Grid or Cloud. As replicating of many super computer of such caliber is not going to be planet and human friendly.

Conclusions

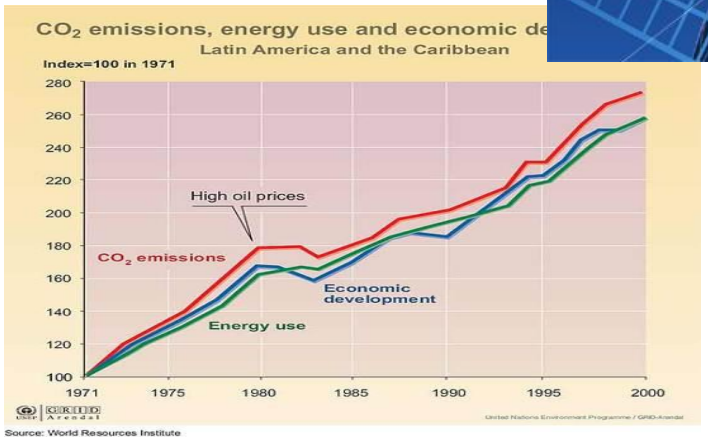
If there are challenges, there are promises also from the technology side

- Hybrid Computing is emerging as solution for the current scenario
- Sharing of resources through Cloud/Grid/Bandwidth is essential
- Power aware scheduling of jobs, use of accelerator like FPGA's with better memory management algorithm, resilient programming are the need of the hour
- And last but not the least, the technology growth has to be society sensitive (Green Computing)

THANK YOU


Research Vessel
0 - 250 km
3900 m
Tail Buoy
Airgun Arrays
Magneto-meter
Sea Gravimeter
Streamer
Channel 1
Channel 312
Ocean Bottom Hydrophone
Ocean Floor
Sediments
Igneous Crust


iStockphoto


CO_2 emissions, energy use and economic de
Latin America and the Caribbean
Index=100 in 1971
280
260
240
220
200
180
160
140
120
100
1971
1975
1980
1985
1990
1995
2000
High oil prices
CO_2 emissions
Economic development
Energy use
Source: World Resources Institute