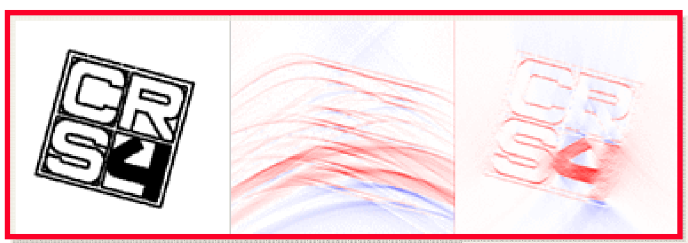
Time Seismic Imaging on a Dataflow Engine

A Preliminary Study

Ernesto Bonomi
Energy & Environment Program
CRS4, Pula, Italy



Outline

Oil&Gas Industry Challenges (1/3)

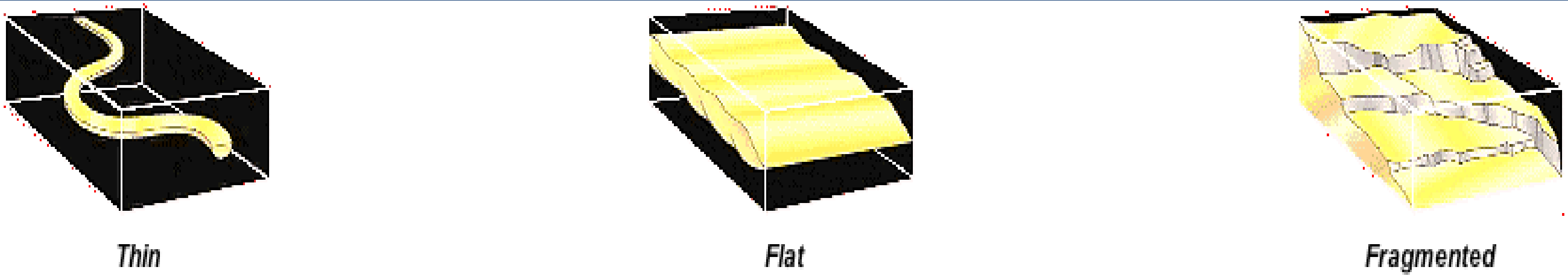
- The Oil&Gas industry depends heavily on high performance computing to support geophysical mapping
  - seismic imaging and reservoir simulation

  to identify a target and estimate the nature of reservoirs and place production wells

…to spur meaningful returns on its investments in drilling and production

- The number of wells/year drilled by a single Oil&Gas company is O(100)
- The cost of a well is O(100M$)

Oil&Gas Industry Challenges (2/3)

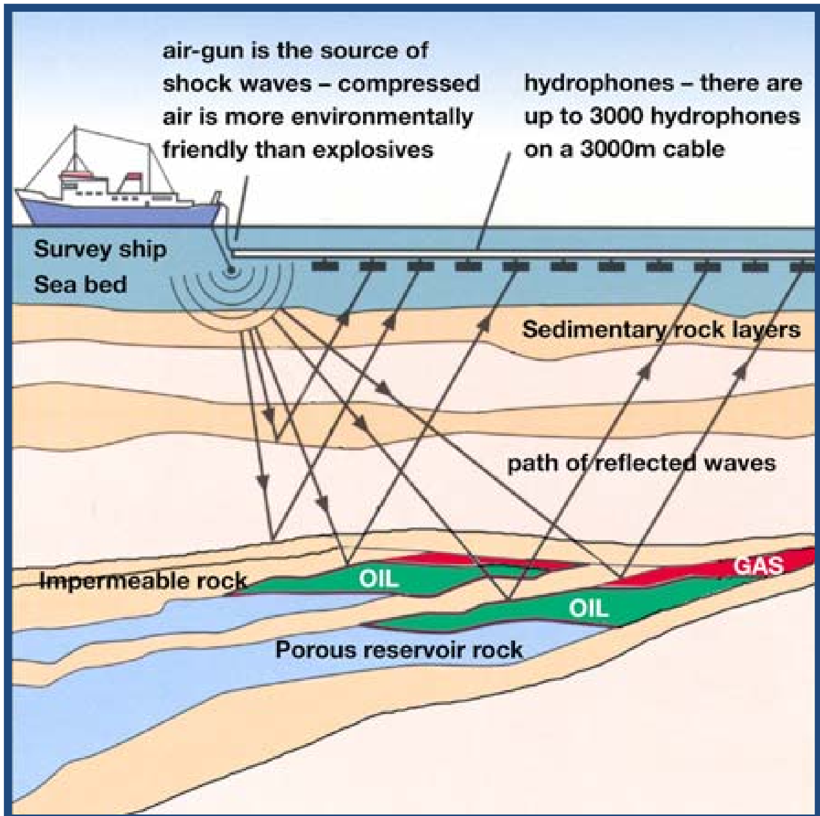


- High resolution seismic imaging tools are then necessary and extremely strategic for the reliable identification of potential reservoirs (1000m-4000m deep)

- Next generation full-waveform-inversion tools will be used to adjust the elastic medium parameters (v_p, v_s, ρ) to best-fit seismic data (FWI):
  - better petrophysical characterization of the reservoir (hydraulic properties, oil content, extension…)

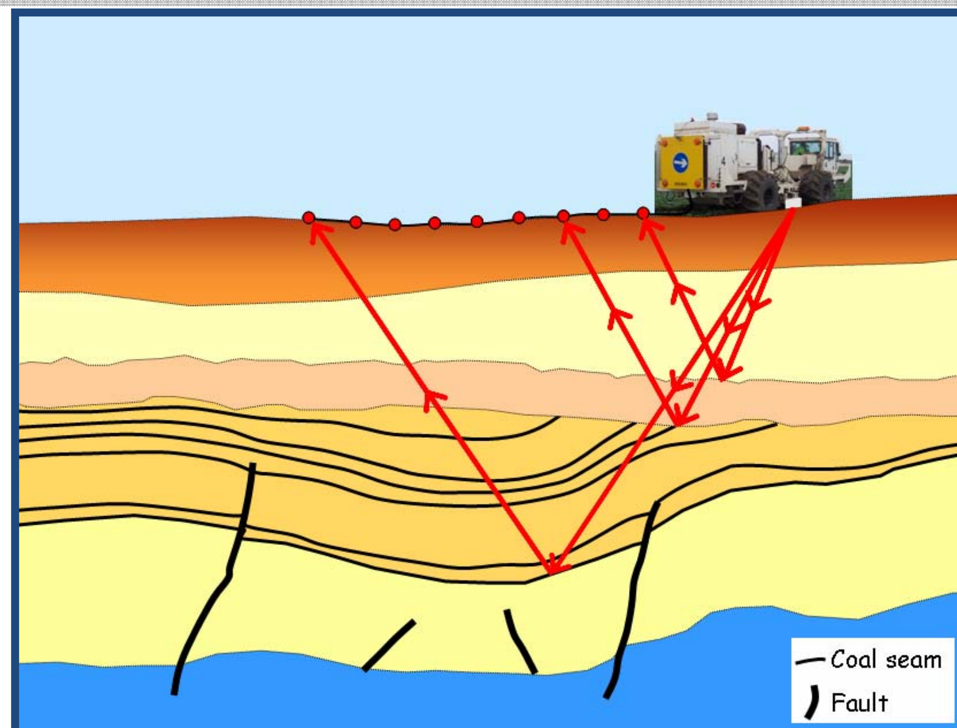
Oil&Gas Industry Challenges (3/3)

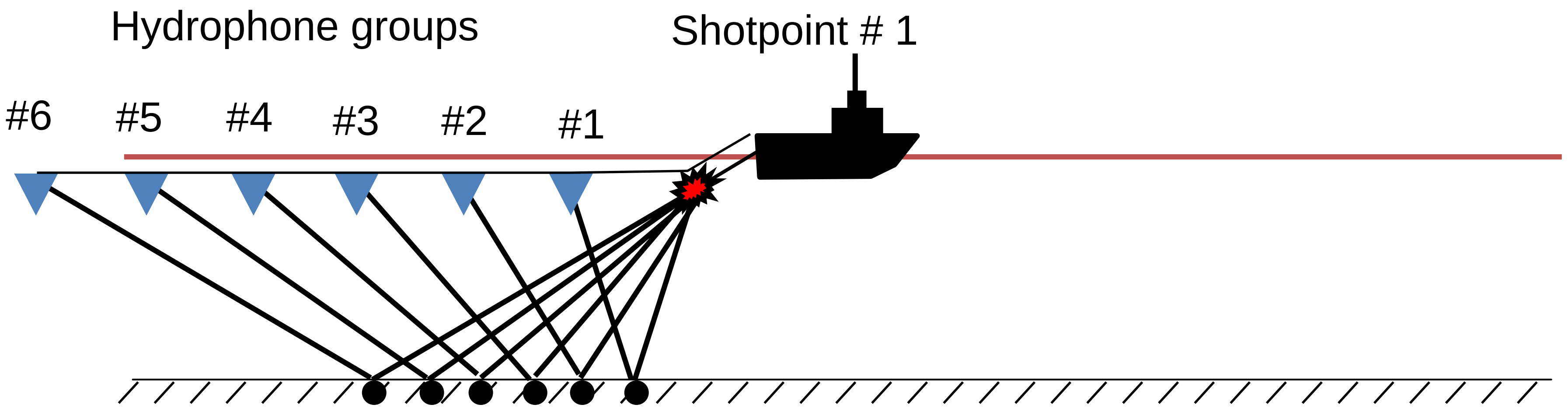
- Hardware and software technologies are also important for Oil&Gas companies to negotiate new activities with crude oil producer countries
  - innovative technologies are the companies assets that guarantee the success of the future operations



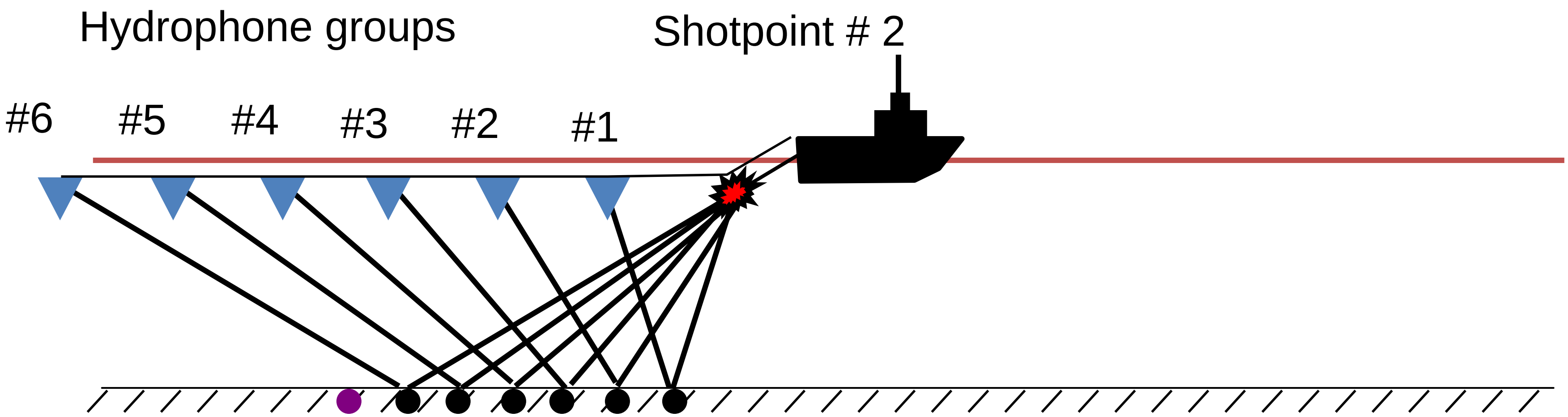
- Each new acquisition is made of multi-azimuth and long-offset data collections
- O(100K) shot gathers ≅ O(10TB) of data to process by volumes of O(1TB)

Sound waves reflected by different geological layers reach the surface at different times

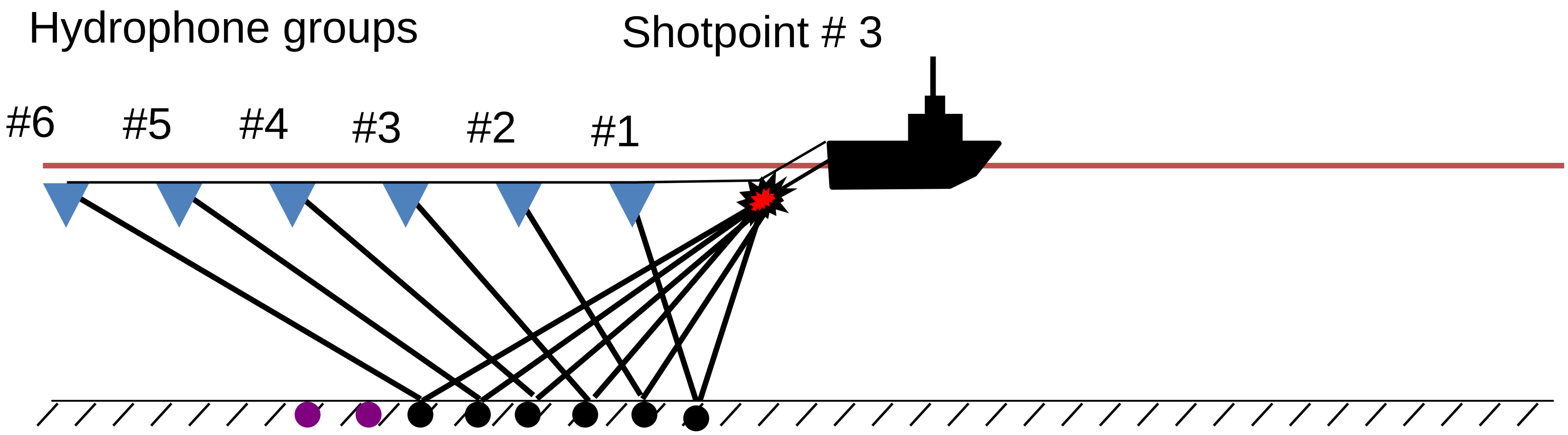



Hydrophone groups
Shotpoint # 1
#6
#5
#4
#3
#2
#1

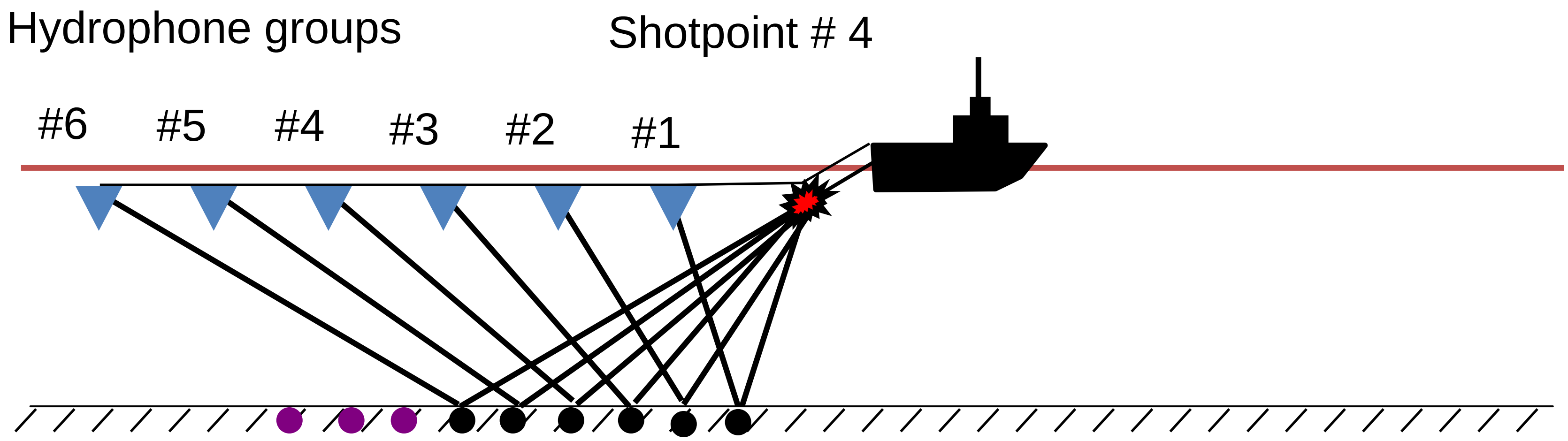
Imaging and Computational geophysics (2/4)

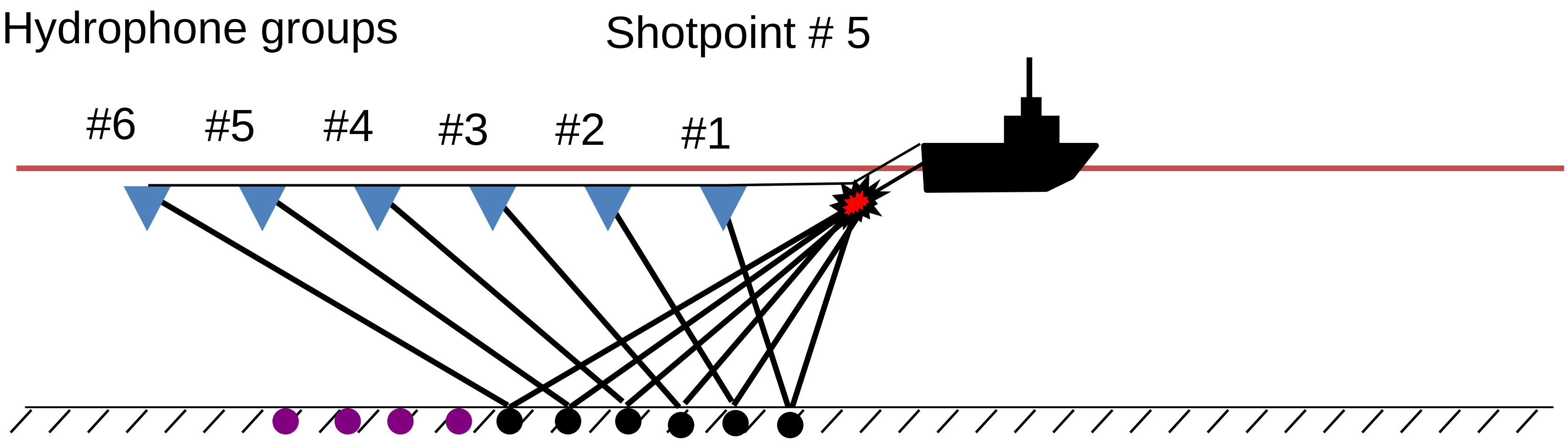


Imaging and Computational geophysics (2/4)

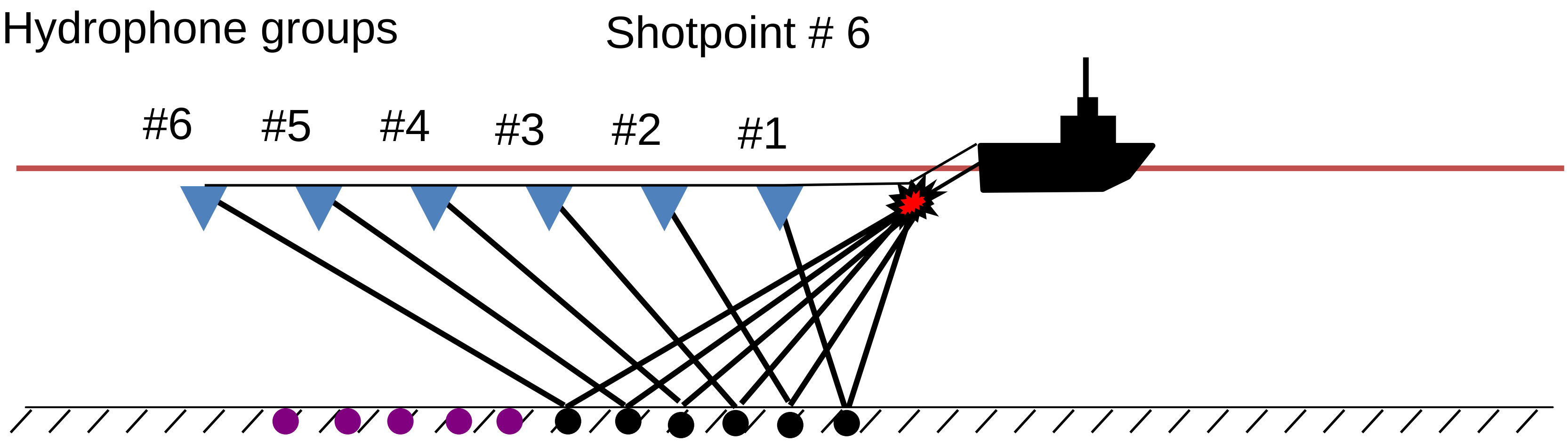


Imaging and Computational geophysics (2/4)

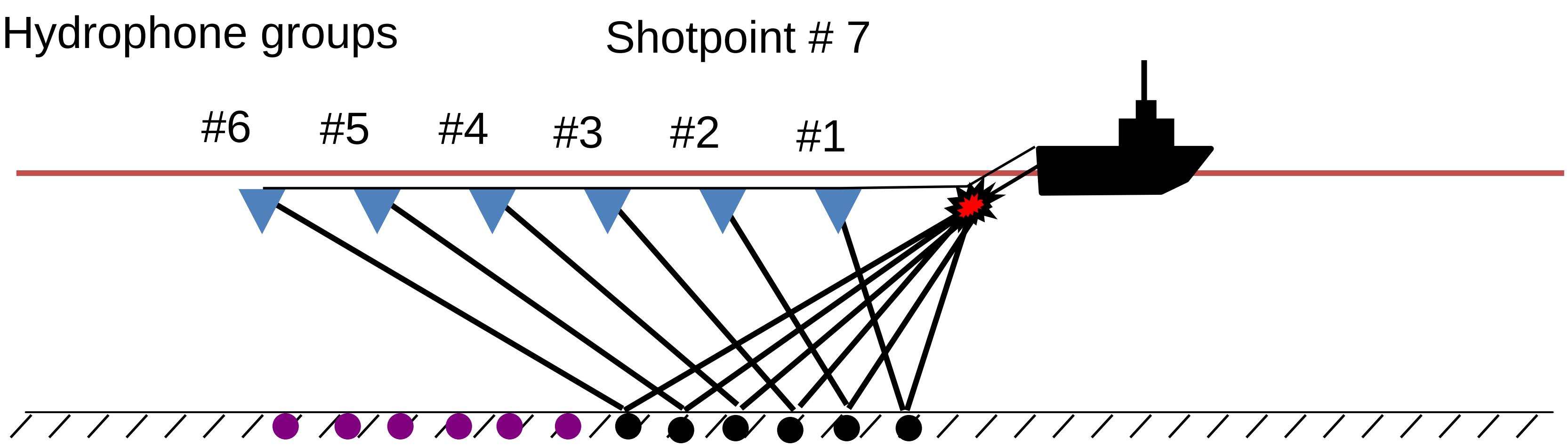


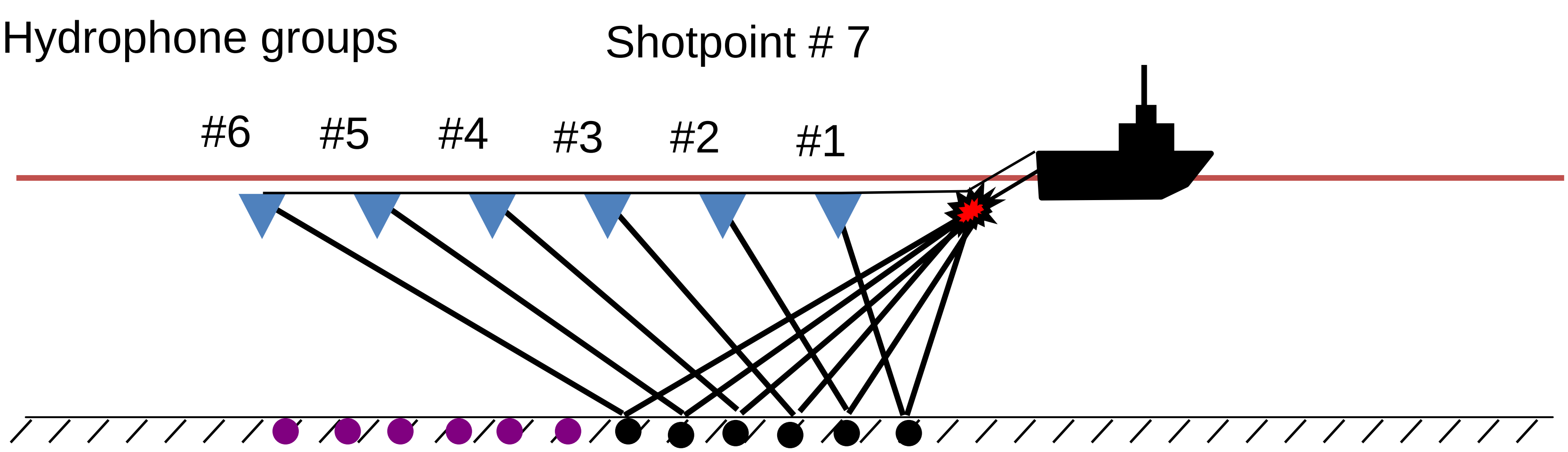

Hydrophone groups
Shotpoint # 5
#6
#5
#4
#3
#2
#1

Imaging and Computational geophysics (2/4)

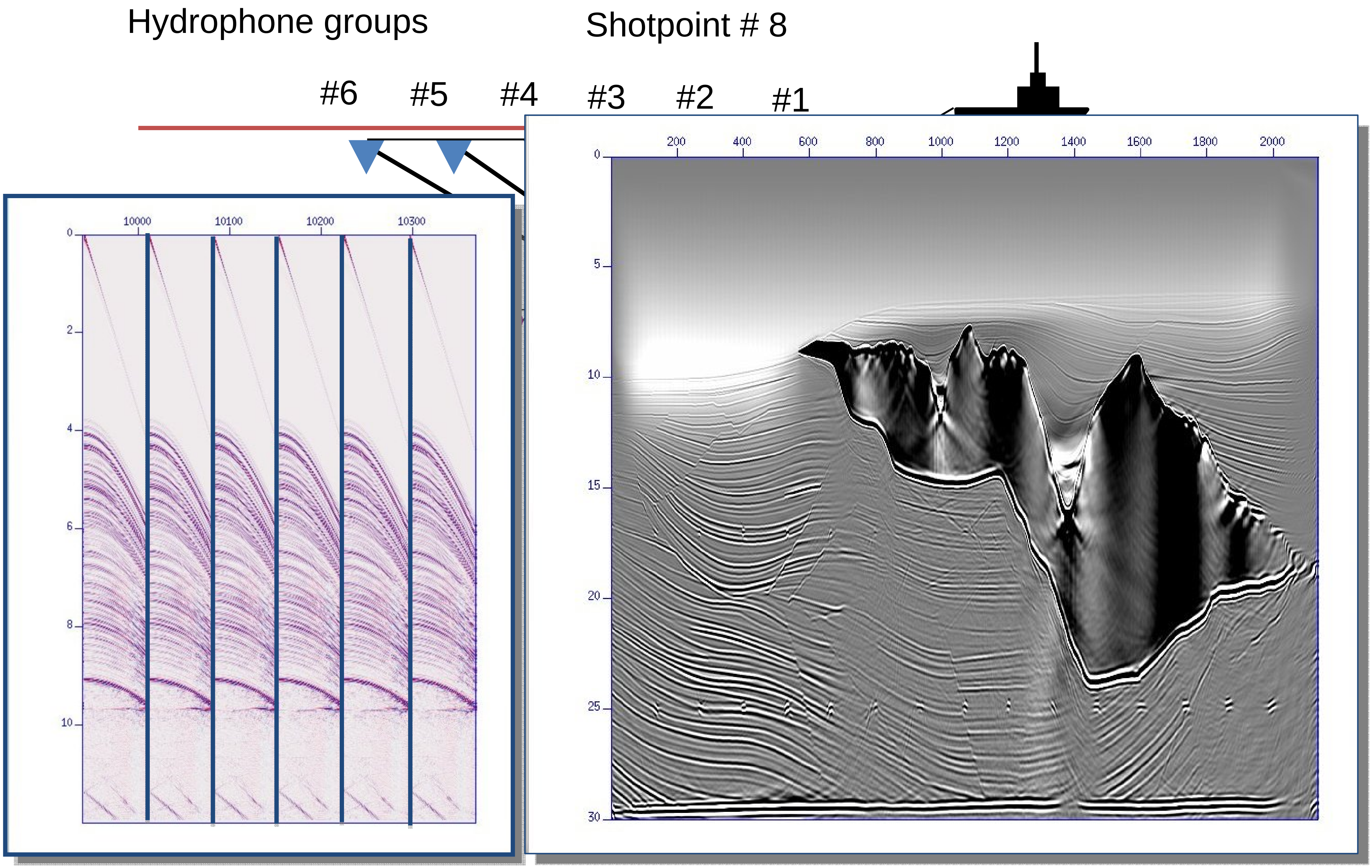


Imaging and Computational geophysics (2/4)

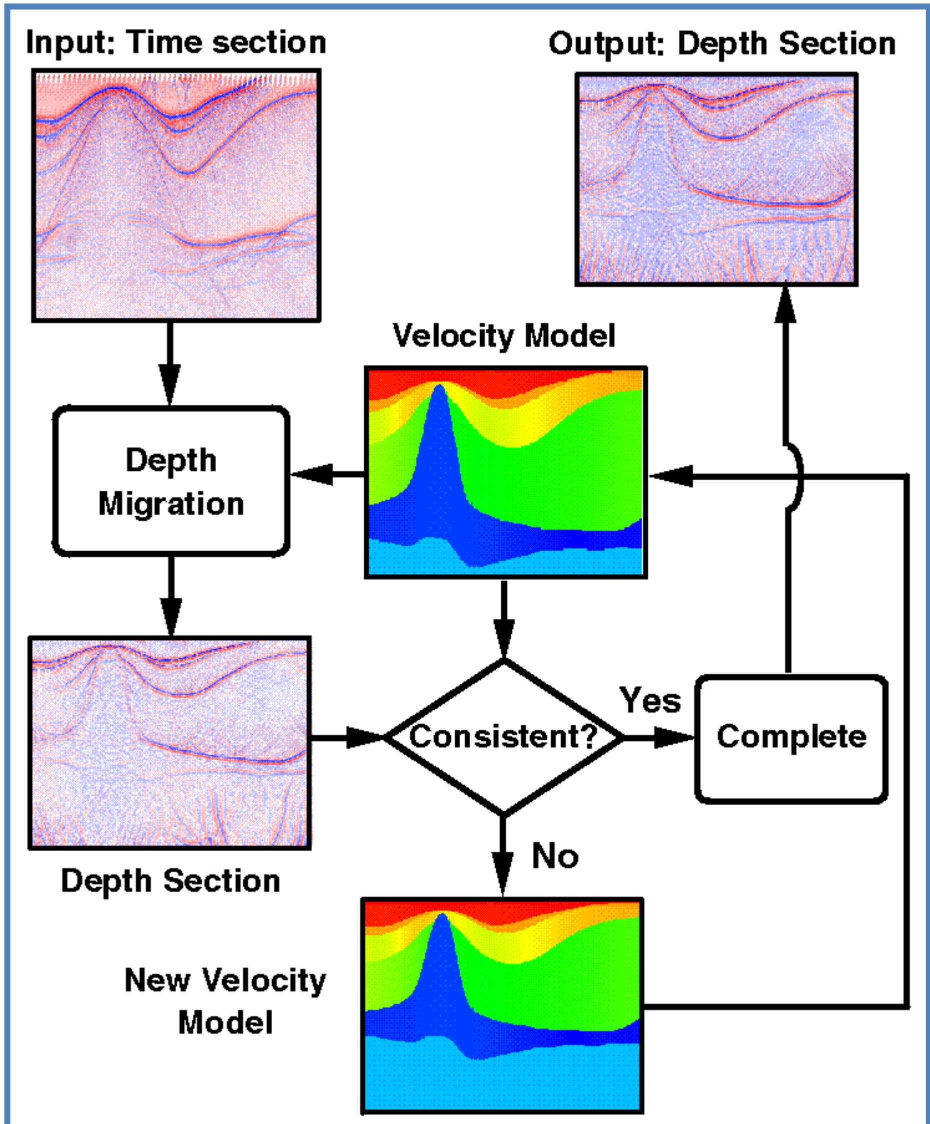



Hydrophone groups
Shotpoint # 7
#6
#5
#4
#3
#2
#1

Imaging and Computational geophysics (3/4)



Imaging and Computational geophysics (4/4)



- Progress in acquisition technology has led to the explosion of the amount of data, O(10TB), to process by volumes of O(1TB)

- Mapping seismic traces into an image requires:
  - very large computing resources, O(1K) CPUs, for a single run
  - a lot of man-machine interaction
  - The iterative improvement of the subsurface image may take several weeks of activity

- …there is a continual need for more performance in HPC…

HPC Trend in Seismics (1/3)

- Highly accurate imaging applications in acoustic media
  - data-driven stacking to zero offset
  - time imaging before/after stacking (PDEs)
  - depth imaging before stacking (PDEs)
  - reservoir characterization (FWI)

requires

1) advanced parallel programming
   - parallelism with no locality, and locality with no parallelism are extremes to avoid
2) more and more powerful computational resources…

- …whereas the performance of a single CPU is no longer increasing like in the past
  - unsustainable energy consumption and heat generation preventing frequency scaling

...as power consumption has become a concern, the dominant paradigm in computer architecture is in the form of...

Multicore systems

Hybrid multicore systems

Dataflow Engines (FPGAs)

- In spite of their spectacular performance, multicores and FPGAs obey to a restrictive programming paradigm, data parallelism or flow parallelism, that makes difficult the numerical solution of many algebraic problems
  - For example, the most efficient way to implement PDEs is to use approximations on structured grids and explicit time marching schemes

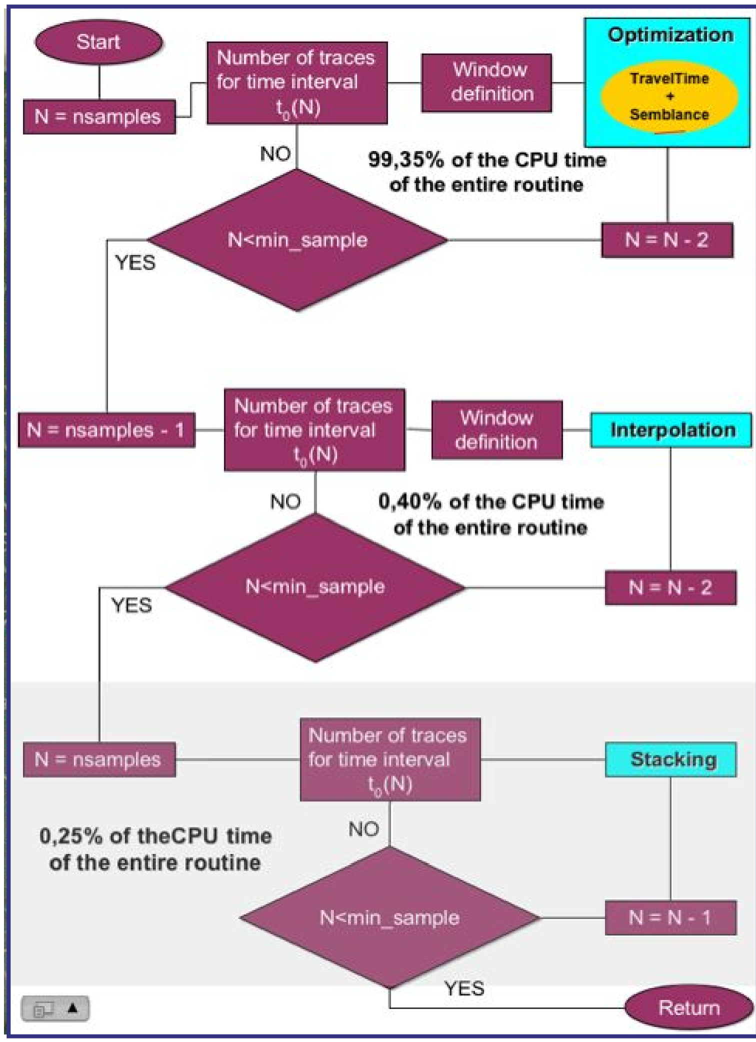
Drawbacks:

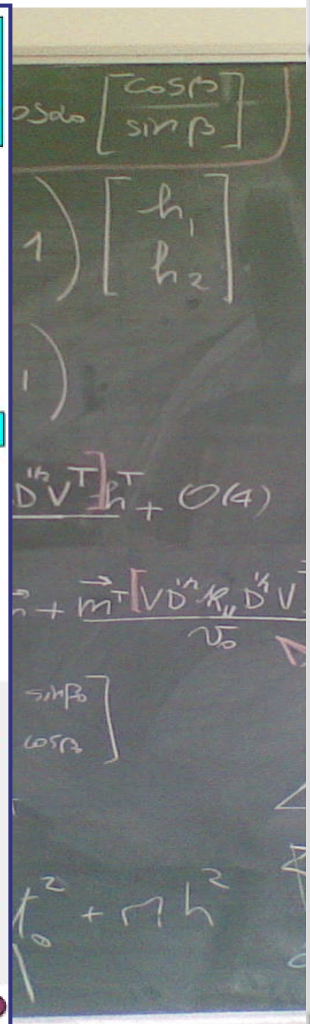
only conditionally stable → **Limit on time-step size**

numerical dispersion problems → **Decreasing time step or high-order schemes**

Time Imaging: a Data-driven Solution (1/6)

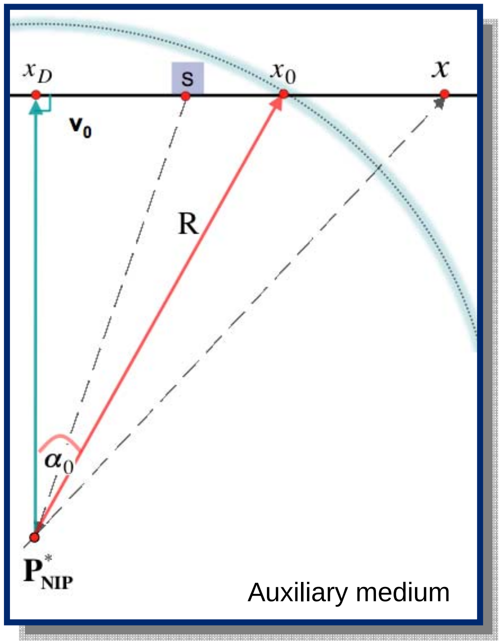
The challenge is the reduction of a mathematical model to a sequence of computational tasks that perfectly fit the paradigm supported by these extreme architectures…**time imaging** is the perfect testbed!

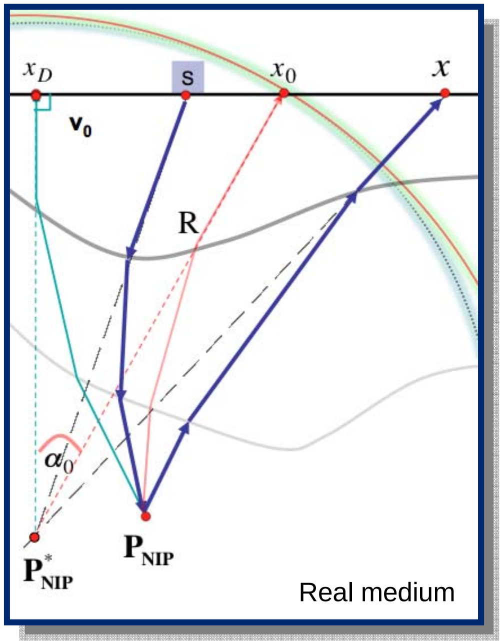


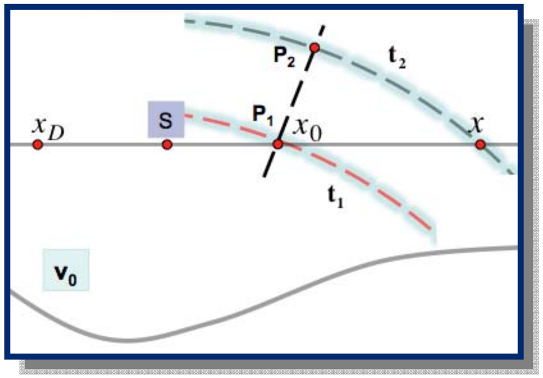


- Instead of approximating operators (finite differences) or solutions (spectral or finite elements), change the problem paradigm!!!

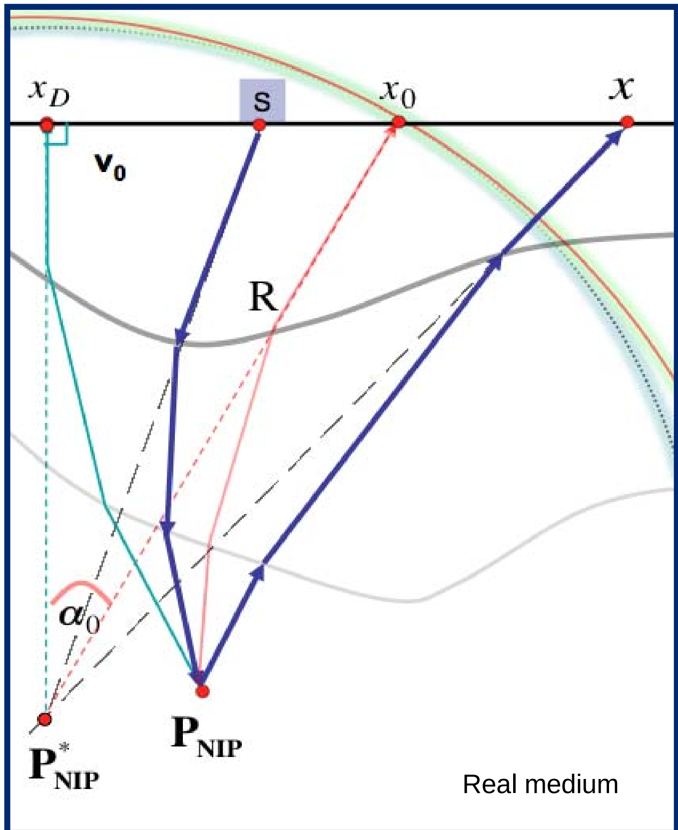
- Model the medium assuming a collection of diffracting points
- Adopt, near a reference ray, a circular approximation of each diffraction wavefront (osculating circle)







$$t_A(x-x_0)-t_A(0)\approx t(x-x_0)-t(0)$$



- Within the circular approximation of each diffracted wavefront, the resulting source-receiver time of flight $\tau_M(x_m - x_D, h)$ takes the form of a double-square root formula written in the auxiliary medium, then corrected by

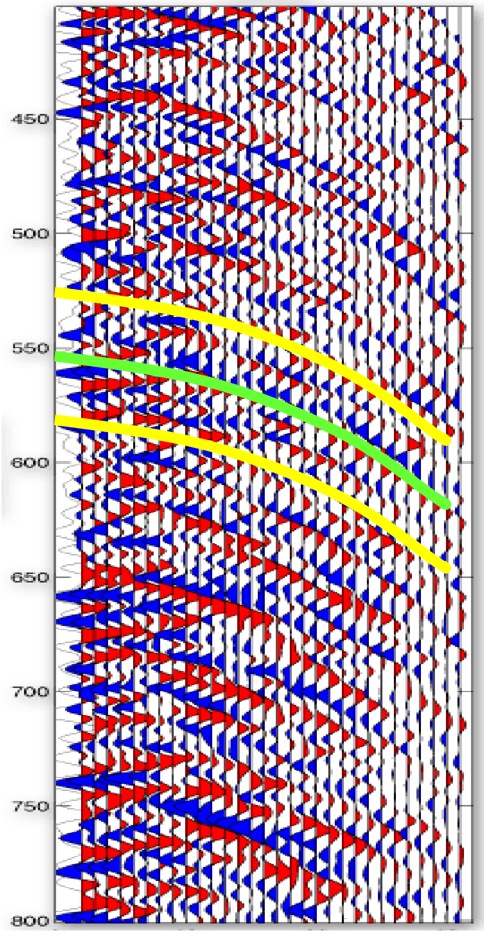
$$t_A(x - x_0) - t_A(0) \approx t(x - x_0) - t(0)$$

The traveltime is calculated with respect to the image-space coordinates (x_D, t_D)

- There are two unknown parameters, R and α_0, characterizing each reference ray tagged by (x_0, t_0)
- There is no need to know the acoustic velocity field

Time Imaging: a Data-driven Solution (4/6)

- The values of $\xi=(R, \alpha_0)$ can be determined from the data by maximizing the fitness, or coherence, between
  - the diffraction traveltimes and
  - the arrival times of seismic events

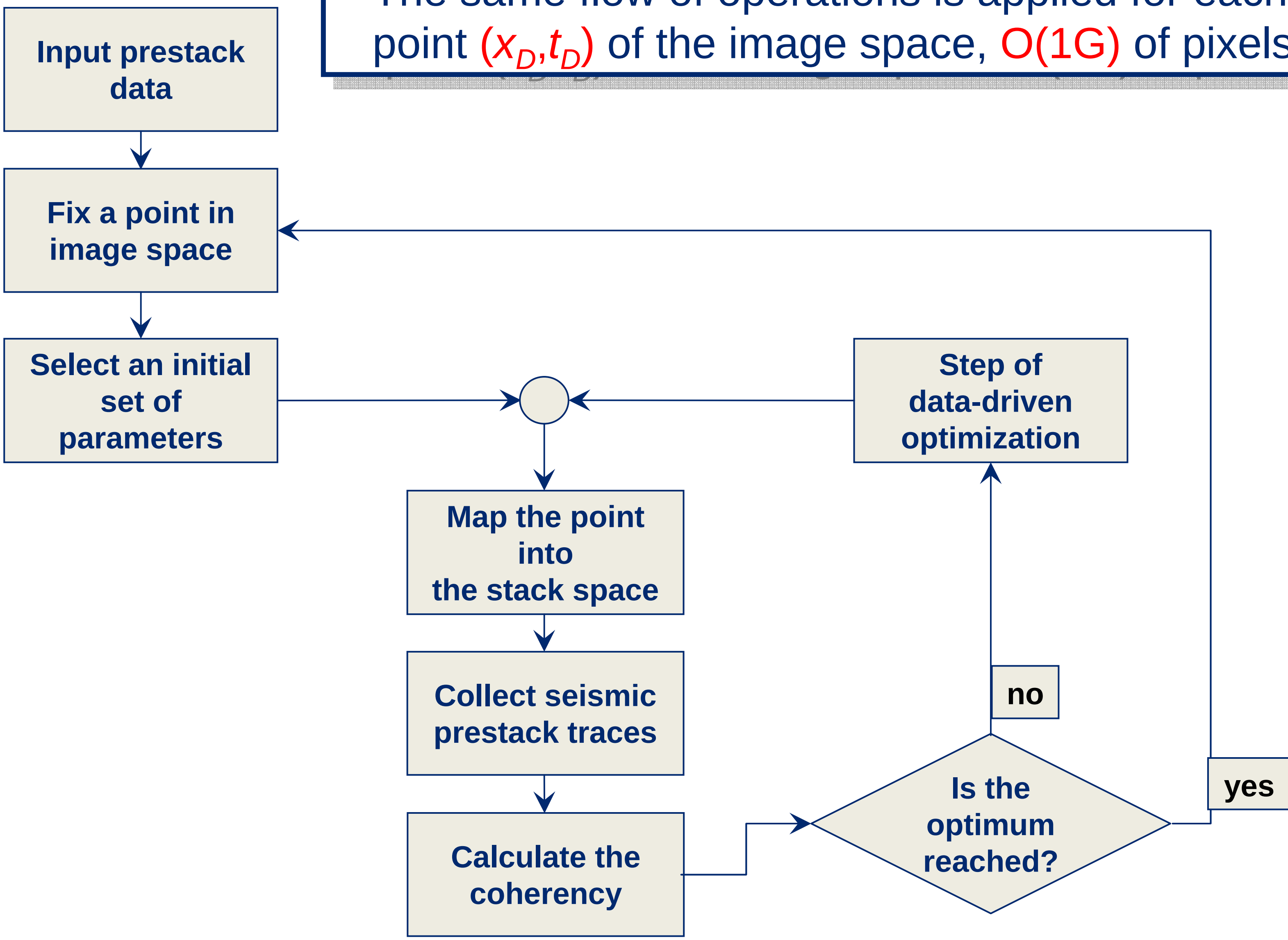


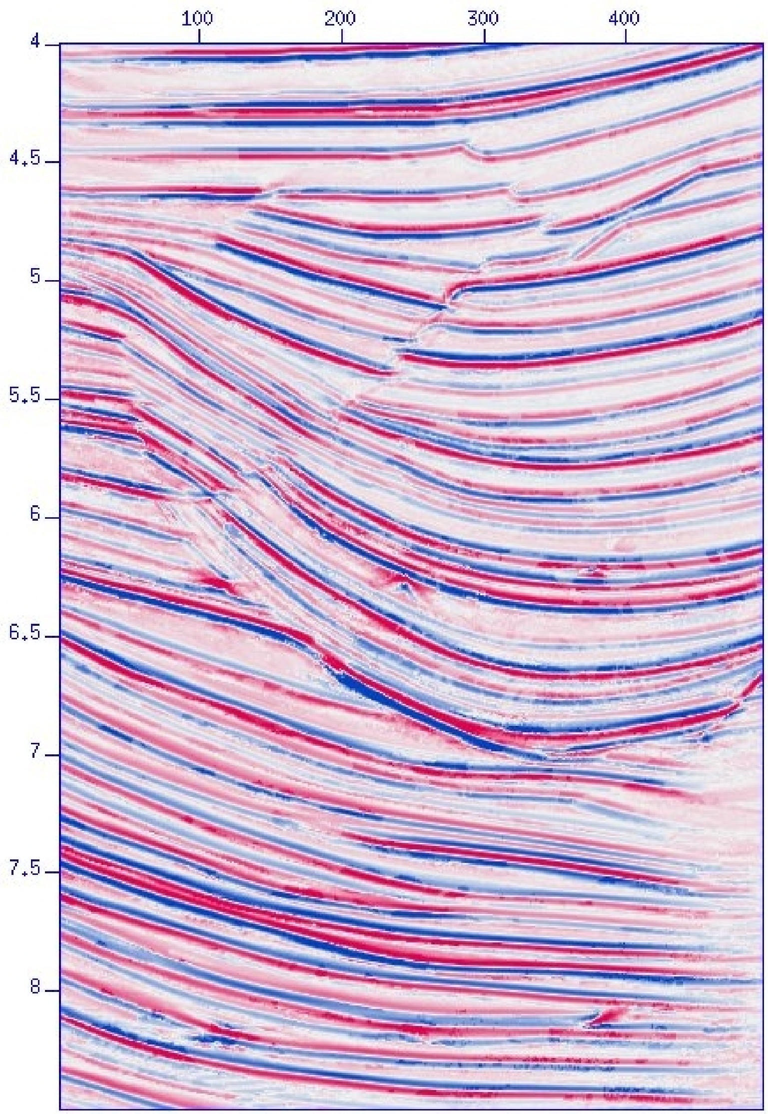
- The coherence of the traveltime prediction is measured by the semblance coefficient

$$S(\xi|x_0,t_0)=\frac{1}{M}\frac{\sum_{k=-N/2}^{N/2}\left|\sum_{i=1}^{M}a_{i,\tau^{(i)}+k}\right|^2}{\sum_{k=-N/2}^{N/2}\sum_{i=1}^{M}\left|a_{i,\tau^{(i)}+k}\right|^2}$$

- The use of a PDE to model the acoustic medium response is no longer necessary!!!

- The same flow of operations is applied for each point (x_D, t_D) of the image space, O(1G) of pixels

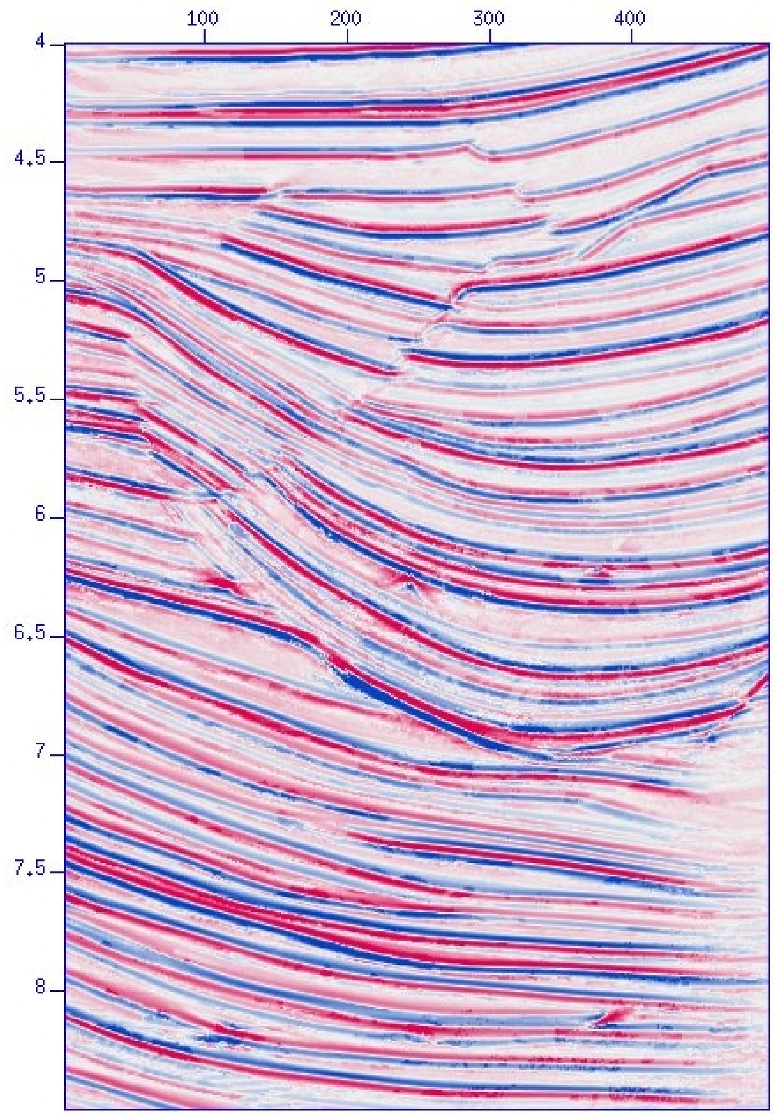




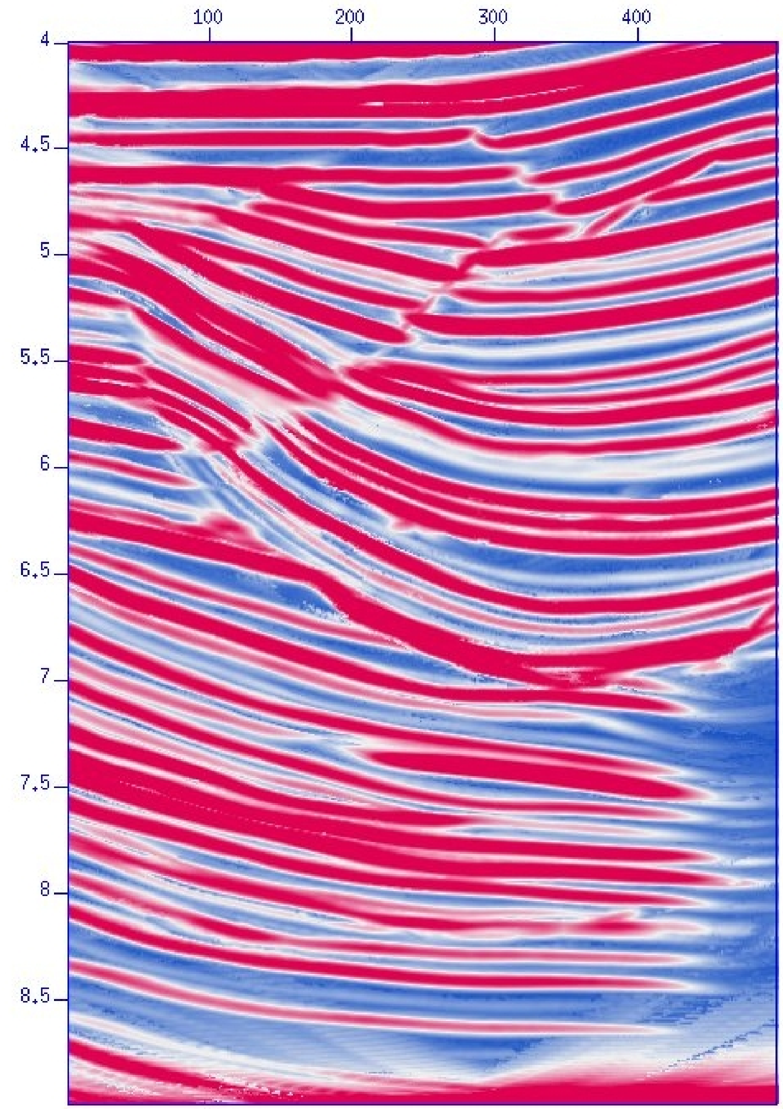
Layers, faults and diffractors

- The best set of parameters $\xi=(R, \alpha_0)$ provides reliable traveltimes
- In the image space, the content of each pixel results from the signal averaged along a traveltime trajectory (green)

Time Imaging: a Data-driven Solution (6/6)



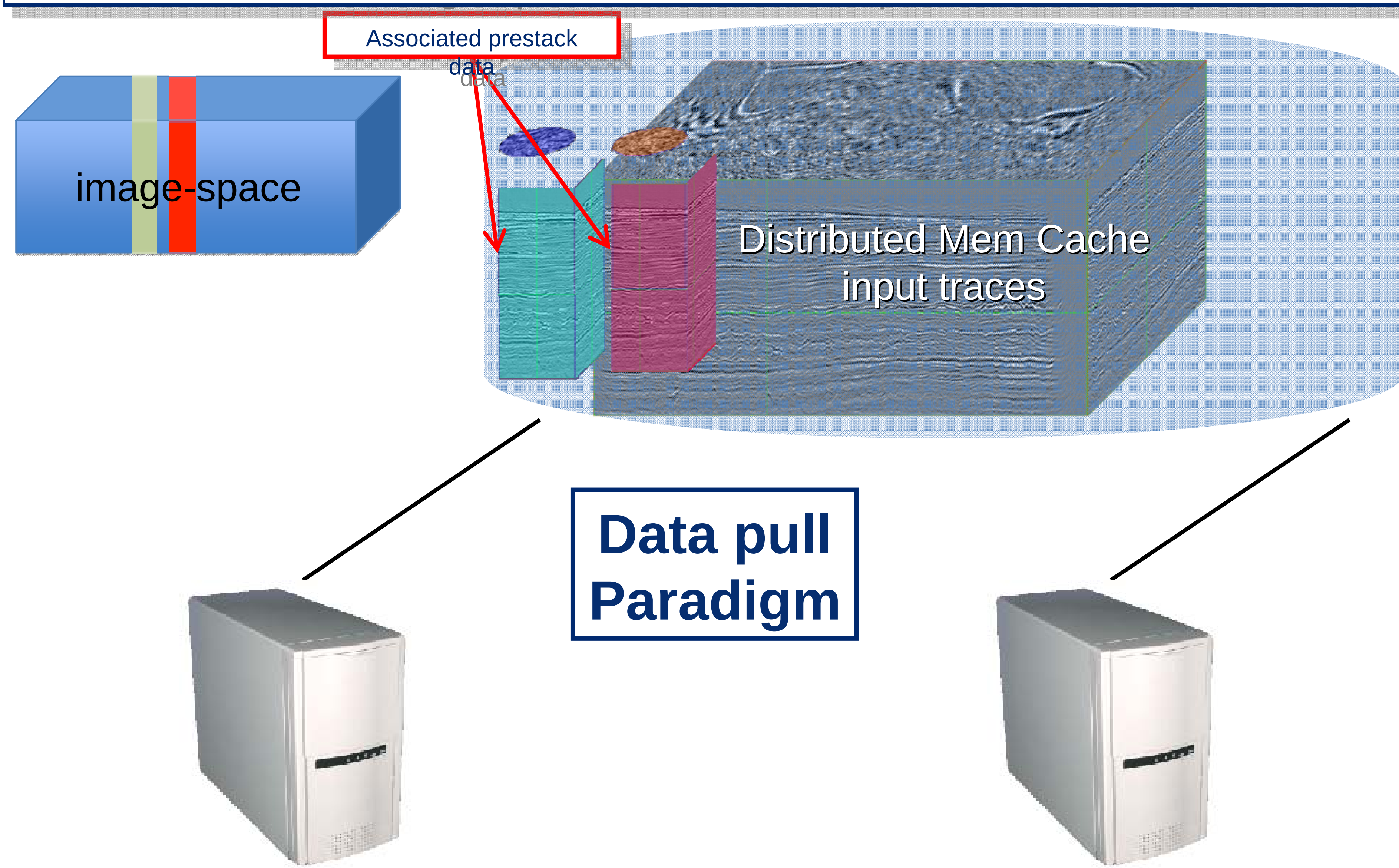
Layers, faults and diffractors



Semblance

Parallelism on Dataflow Engines (1/3)

Coarse level parallelism:
distribution of the image-space with the consequent volume of input traces

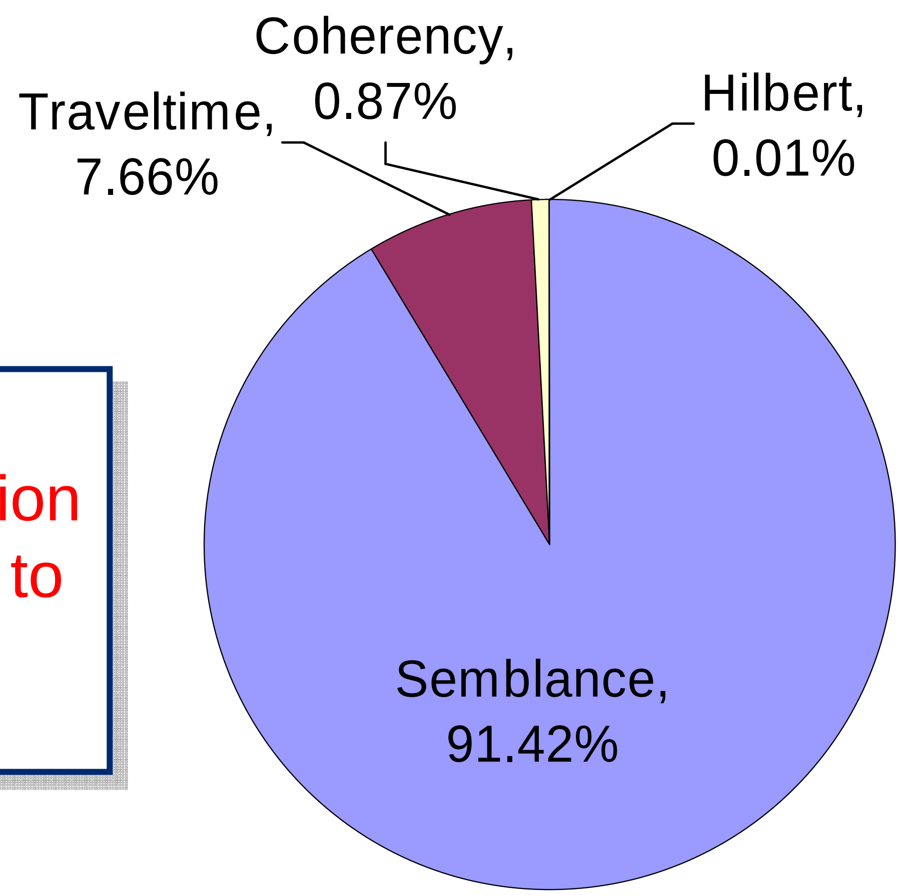


Data pull Paradigm

CPU performance on Intel Xeon 5450 3GHz for a similar application (CRS3D)

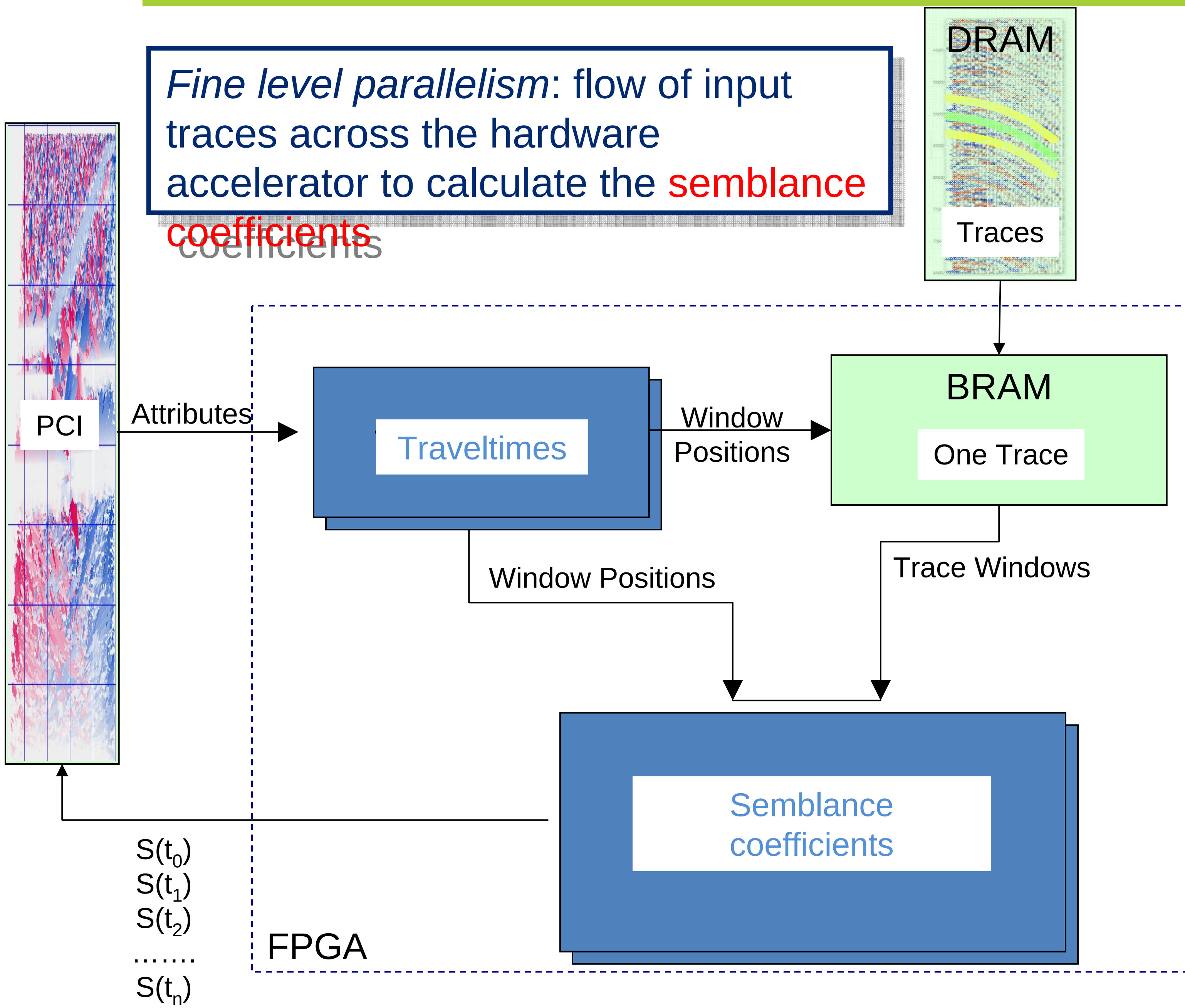
- 209 seconds per output sample
- 116 seconds in line search

- The semblance coefficient is the bottleneck of the application and then the good candidate to be processed by a dataflow engine



Parallelism on Dataflow Engines (3/3)

Fine level parallelism: flow of input traces across the hardware accelerator to calculate the semblance coefficients
coefficients



Conclusion

- The presented time-imaging tool can be determinant in the processing of datasets corresponding to areas with a severe structural geological trend

- The present application perfectly fits the restrictive parallelism paradigm imposed by dataflow engines, thus taking fully advantage of their spectacular performance

- With the explosion of the amount of seismic data, the use of extreme architectures is becoming vital for Oil&Gas companies

- This work represents a relevant scientific and technological step and constitutes an innovation which opens new perspectives toward industrial processing