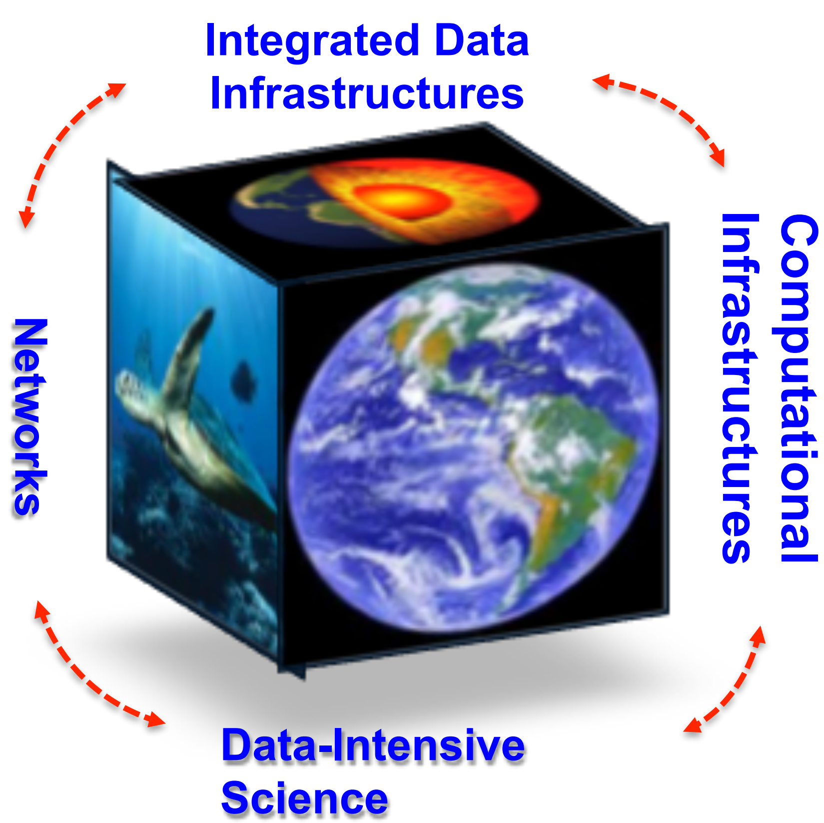
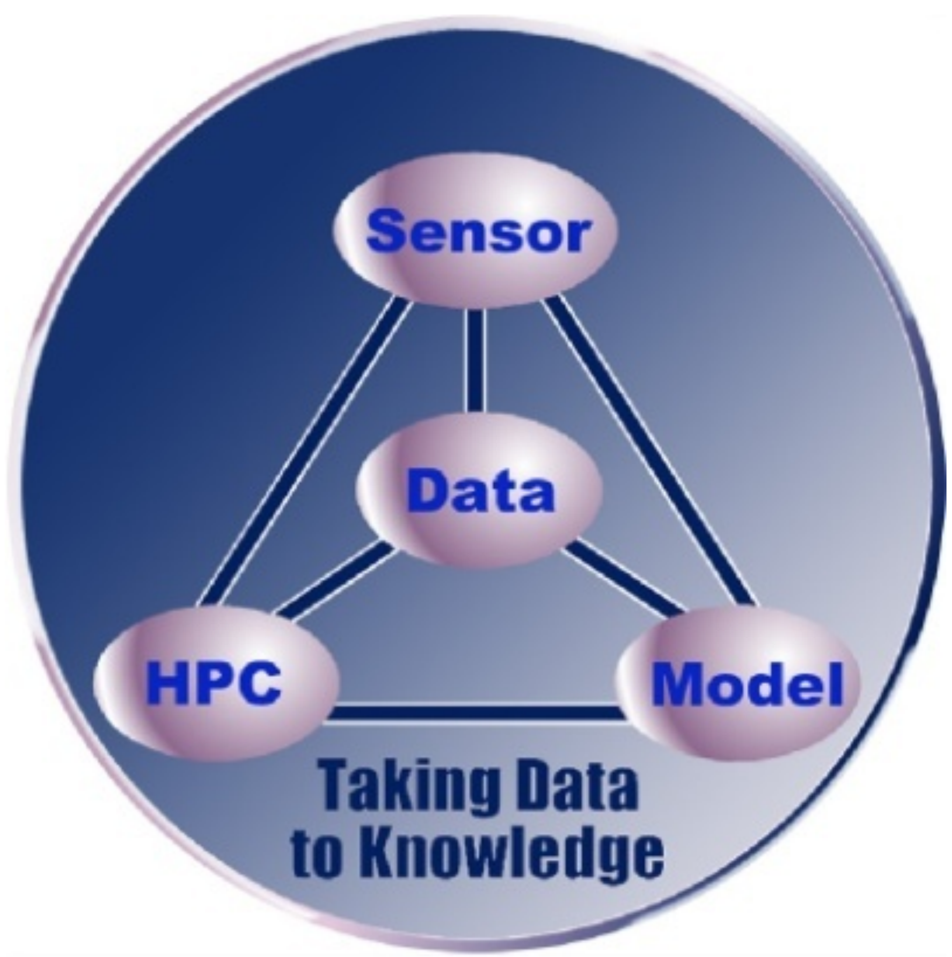
Big Data and Data-intensive challenges in the Earth and Universe science: a change of paradigm

Jean-Pierre Vilotte

Institut des Sciences de l'Univers (CNRS-INSU)
Institut de Physique du Globe de Paris (IPGP)

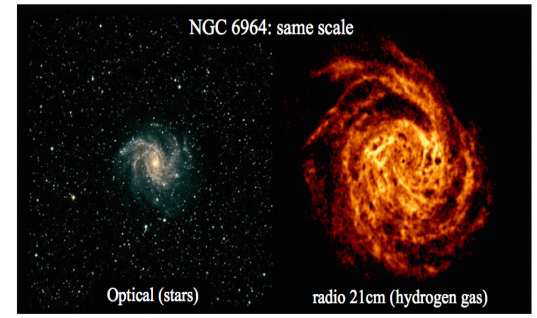


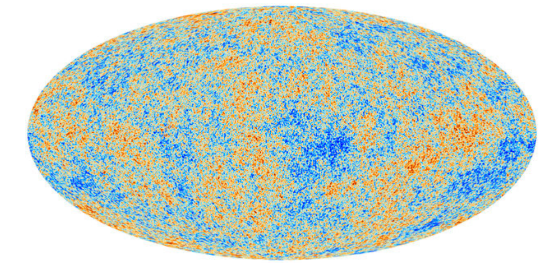


Big Data and HPC: a strategic challenge

Drive Scientific discoveries

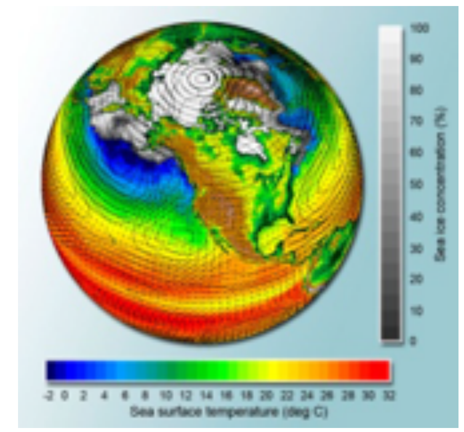
- structure of the universe; climate evolution, seismology; geophysical fluids; high-pressure and high-temperature materials; earthquakes and tsunamis

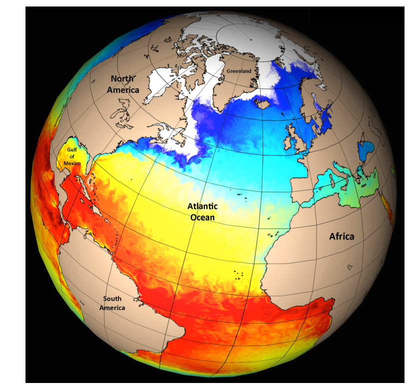




Across multiple disciplines

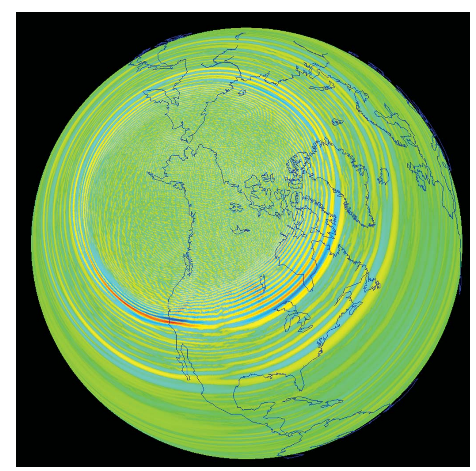
- Astronomy & Astrophysics
- Climate, Atmosphere, Ocean
- Solid Earth Sciences
- Continental surfaces and interfaces

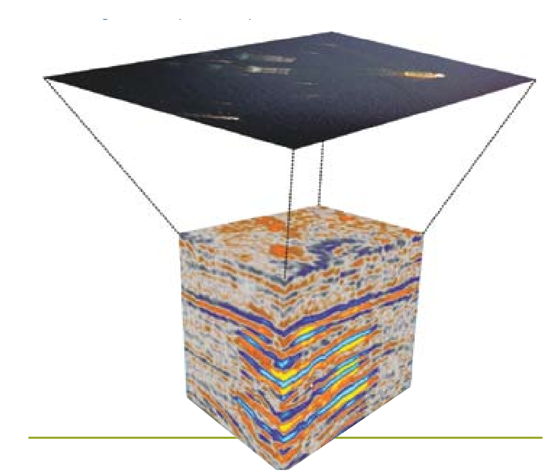


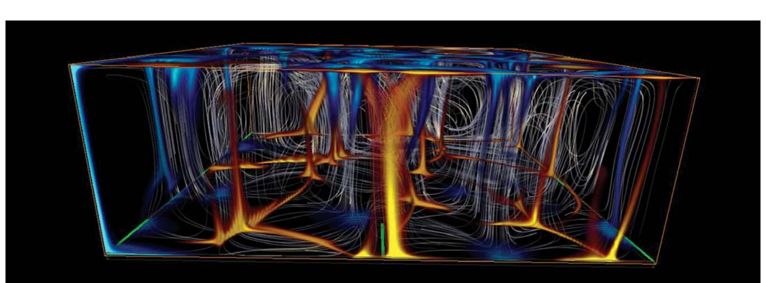


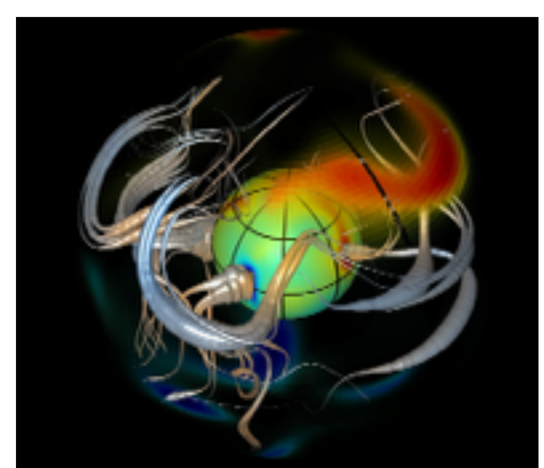
Socio-economical applications

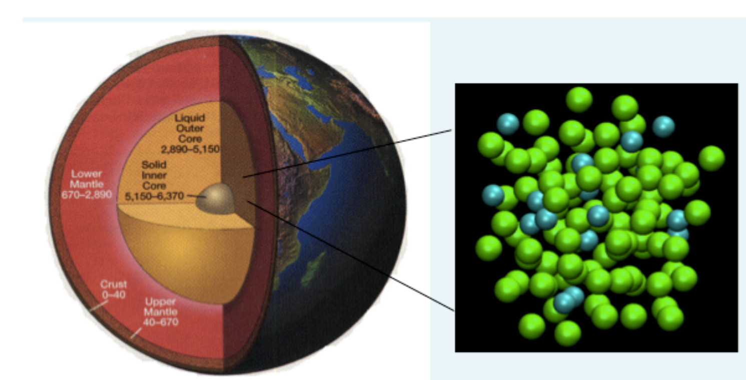
- Climate evolution and prediction
- Natural hazards (earthquakes, volcanoes, tsunamis, landslides …)
- Exploration of new energetic resources
- Environment changes







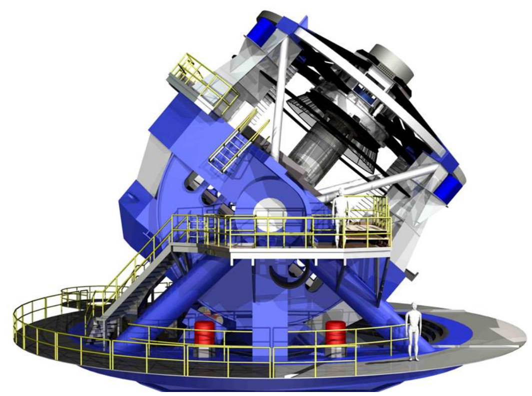




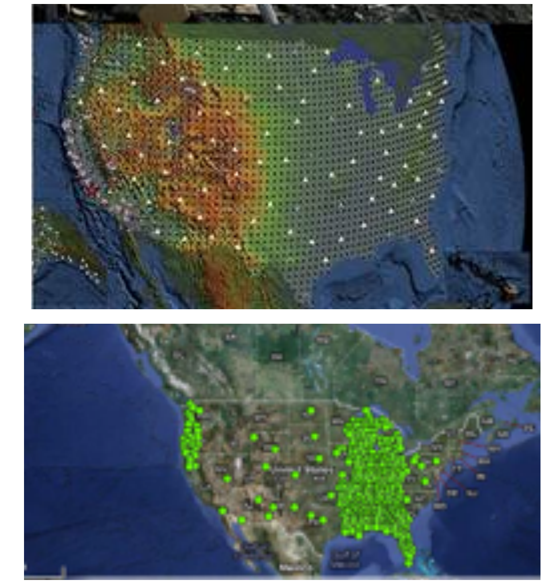
Extreme scale science

Ubiquitous data explosion: 100 PBs era

~10 exabyte/jour



~20 PBs/night



~4 PBs

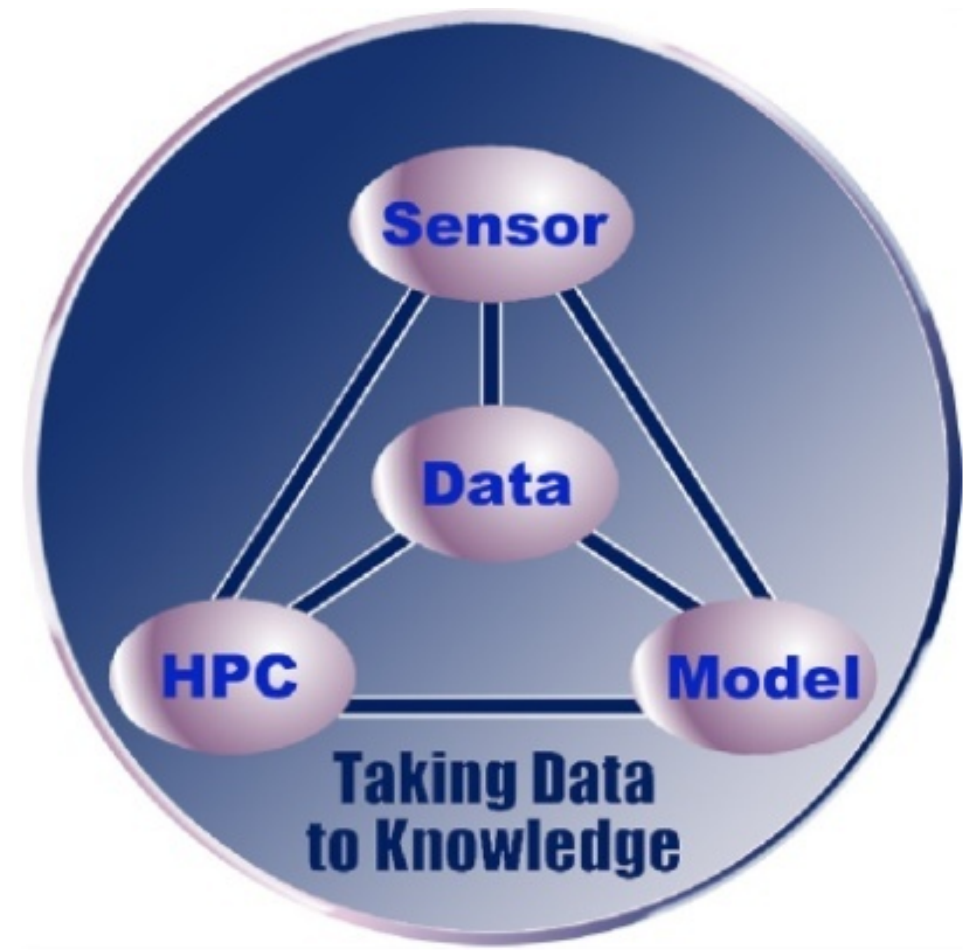
Data explosion generated by:

- Large throughput instruments; dense observation and monitoring global and regional arrays, space observation
- Large HPC simulation

Next generation discoveries require:

- Dealing with big data : new systems of large instruments and of observation and monitoring (spatial, land, ocean and ocean bottom), data archiving, curation and preservation
- Big data analytics: innovative processing and analysis methods of massive data generated by large throughput instruments, observation systems and simulation
- Data-intensive simulation: innovative methods for large-scale simulation of complex multi-physics and multi-scales systems
- Data inversion and assimilation: breakthrough for Bayesian methods in high-dimensional parameter spaces
- Statistics and Stochastic methods: quantification of direct and inverse uncertainties, together with large extreme events

Science is changing

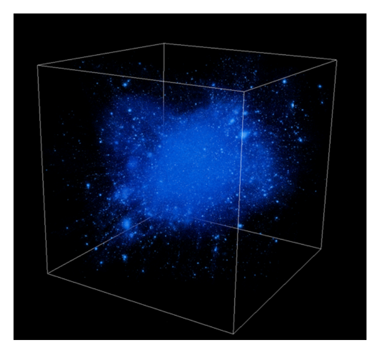


THOUSAND YEARS AGO
science was empirical
describing natural phenomena

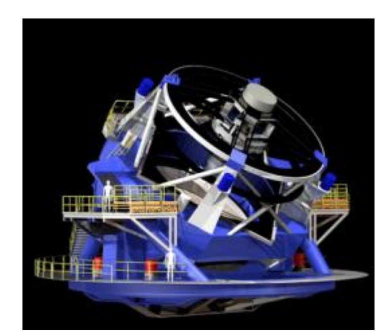
LAST FEW HUNDRED YEARS
Theoretical branch using models
generalisations

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$

LAST FEW DECADES
A computational branch
simulating complex phenomena

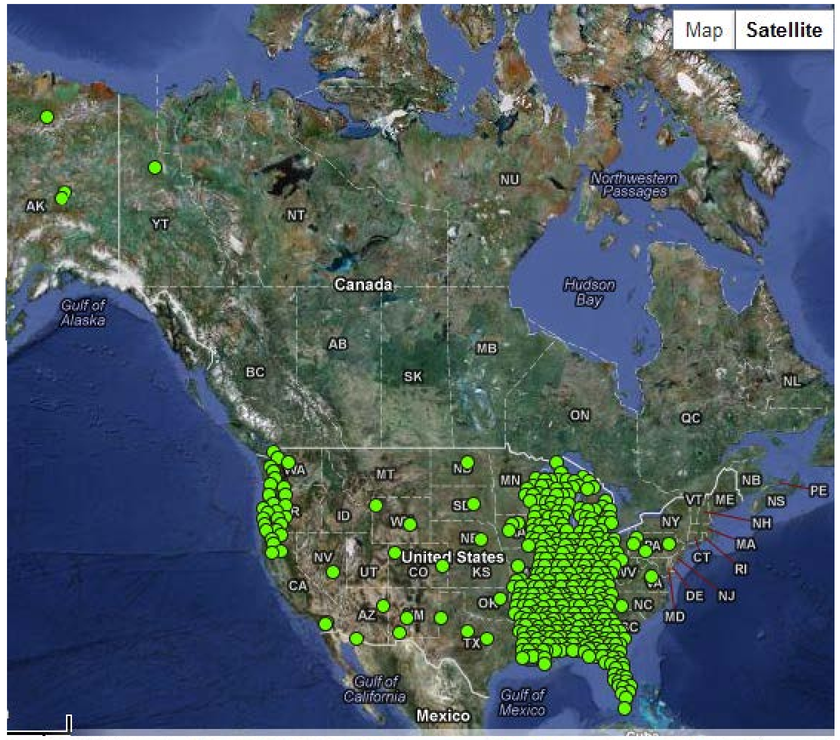


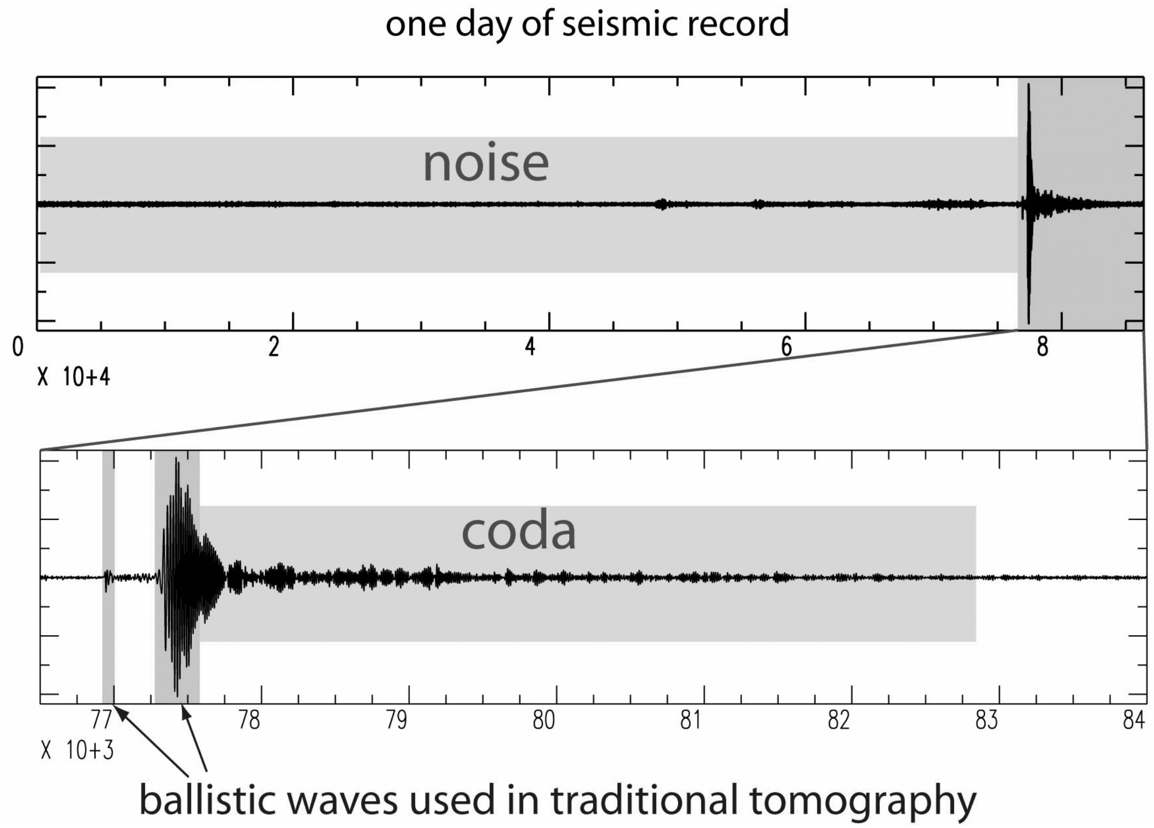
TODAY
Data-intensive science, synthesising theory, experiment,
observations and computation with statistics
new way of thinking required



Big data analysis

Seismology: data-intensive seismic noise analysis





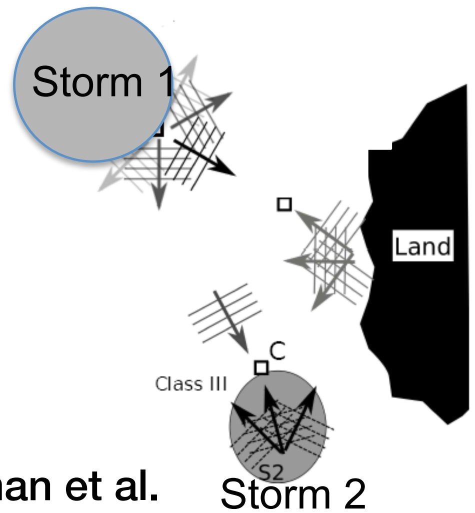
For a random wave field with homogeneous sources distribution everywhere in the medium

$$\frac{d}{d\tau} C_{A,B}(\tau) = \frac{-\sigma^2}{4\,a}\left(G_a(\tau, \vec{r}_A, \vec{r}_B) - G_a(-\tau, \vec{r}_A, \vec{r}_B)\right)$$

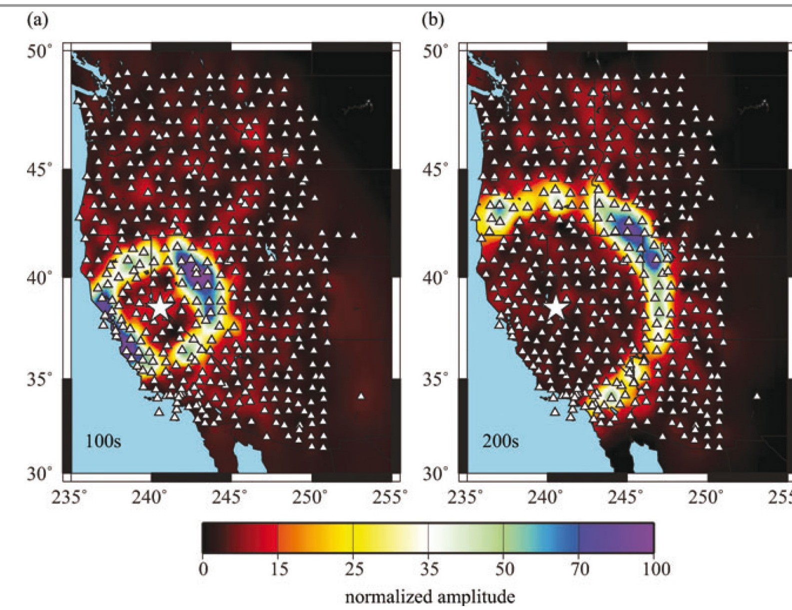
noise cross-correlation — Green function

- Cross-correlation of seismic noise between two stations from long enough records is equivalent to an experiment when a source is acting at location of one station and recorded at another
- Repetitive computation of noise cross-correlations are equivalent to using repetitive sources and can be used to detect time changes in the medium

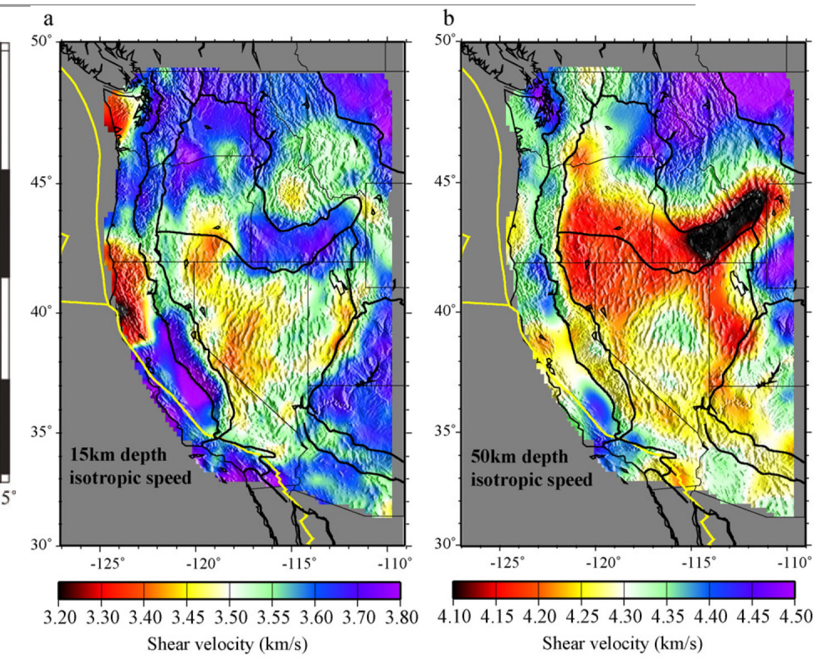
Seismic noise applications



Stutzman et al.

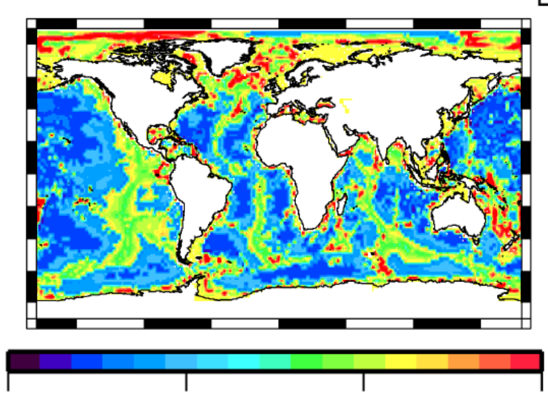


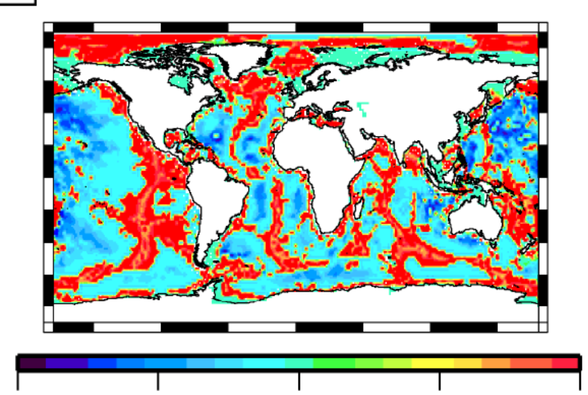
Shapiro et al.



Seismic noise tomography

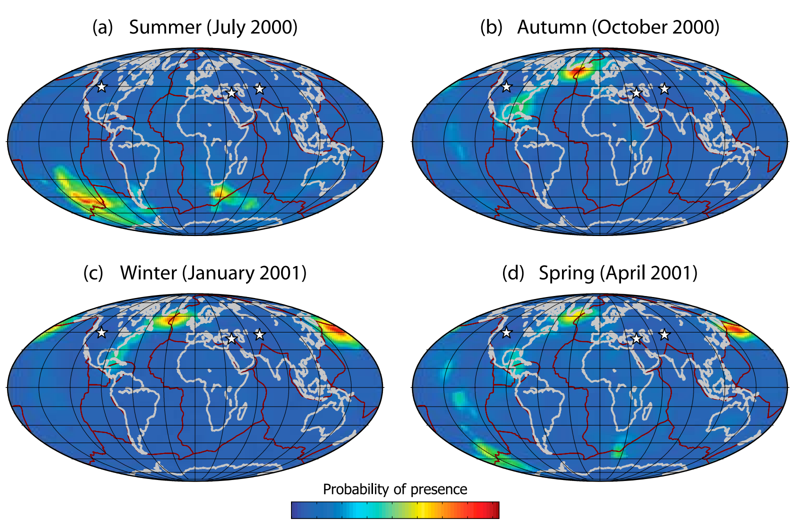
Rayleigh waves | 6 sec | Body waves



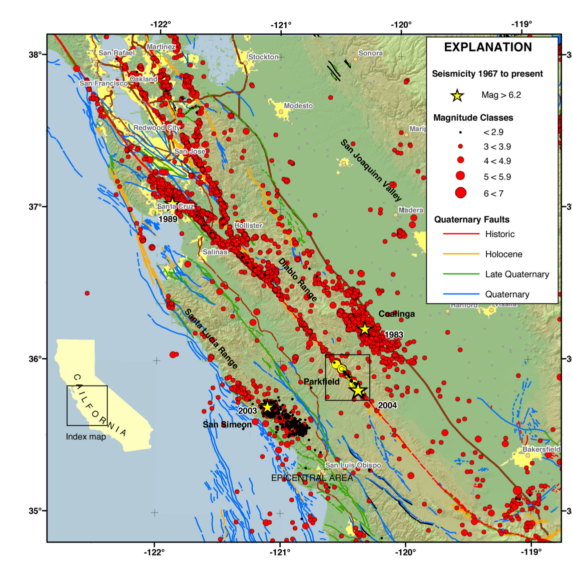


Ocean seismic sources

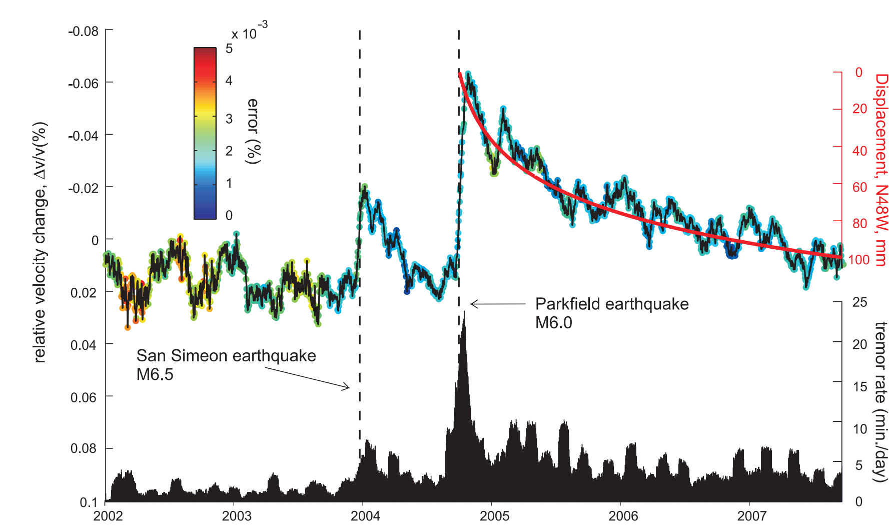
Earthquake-induced property changes



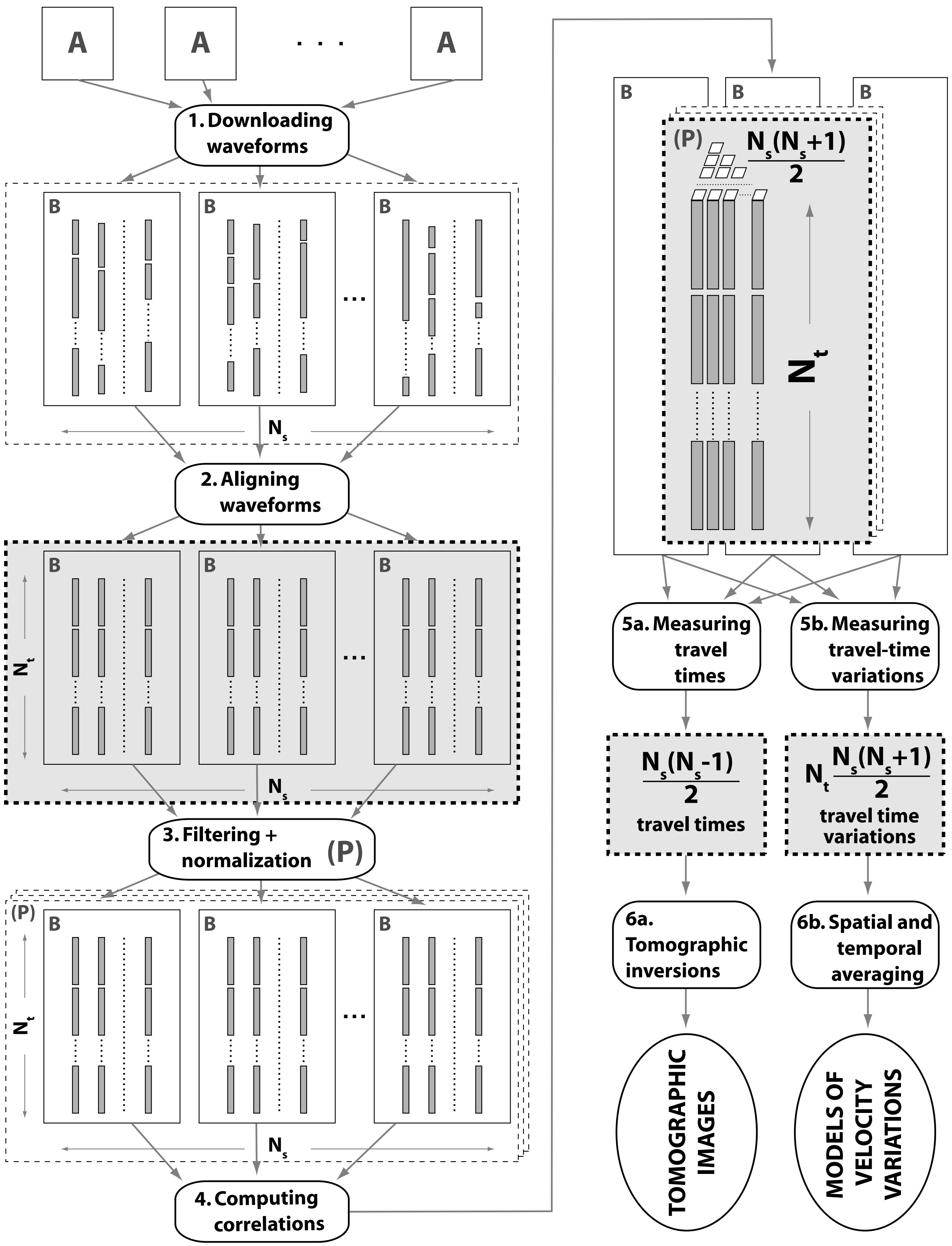
Landes et al



Brenguier et al.



Seismic noise data analysis



Data ingestion / quality control

- N-dimensional ***time series***
- ***binary large objects* (blob): > 100 TBs**
- ***fine granularity: variable chunk sizes (GBs)***
- Partitioning, indexing, replication

Data processing

- **Low level data access pattern**
- **Linear complexity**
- Streaming data workflow
- Provenance and metadata management

Data analysis

- **Cross-correlation** and higher order statistics
- **Quadratic complexity** and CPU intensive
- Thread-blocks CUDA and CSP
- **Secondary data : ~ 6 * N^2 * N_t**
- Provenance and metadata management

Data-intensive analysis chalenges

Intrinsic infrastructure mismatch

- **Data volumes increase 100x in 10 years**
- **I/O bandwidth improves ~3x in 10 years**
- **Data analysis resources close to the data**

Need for efficient data crawling strategy

- **data locality**
  horizontal and vertical re-use
- **memory/IO bandwidth and latency**
  hierarchy of data storage (SDD,HDD)/memory, optimized aggregate sequential IO bandwidth

Data Architecture:

- **Data Center infrastructure:** archiving and distribution -> archiving synthetic models
- **Data analysis infrastructure**: new data-intensive paradigms enabled by HPC, Hybrid architecture (GPU), PFS, HDFS, Hadoop-MapReduce; NoSQL DBs, CUDA-SQL

A Data-scope environment and framework:

- **Analyze and model 100 TB+** of data in academic setting;
- **~ PB+ of storage with safe redundancy**;
- **High sequential IO throughput ~ aggregated disk speed**;
- **Streaming data analyses on par with data throughput;**

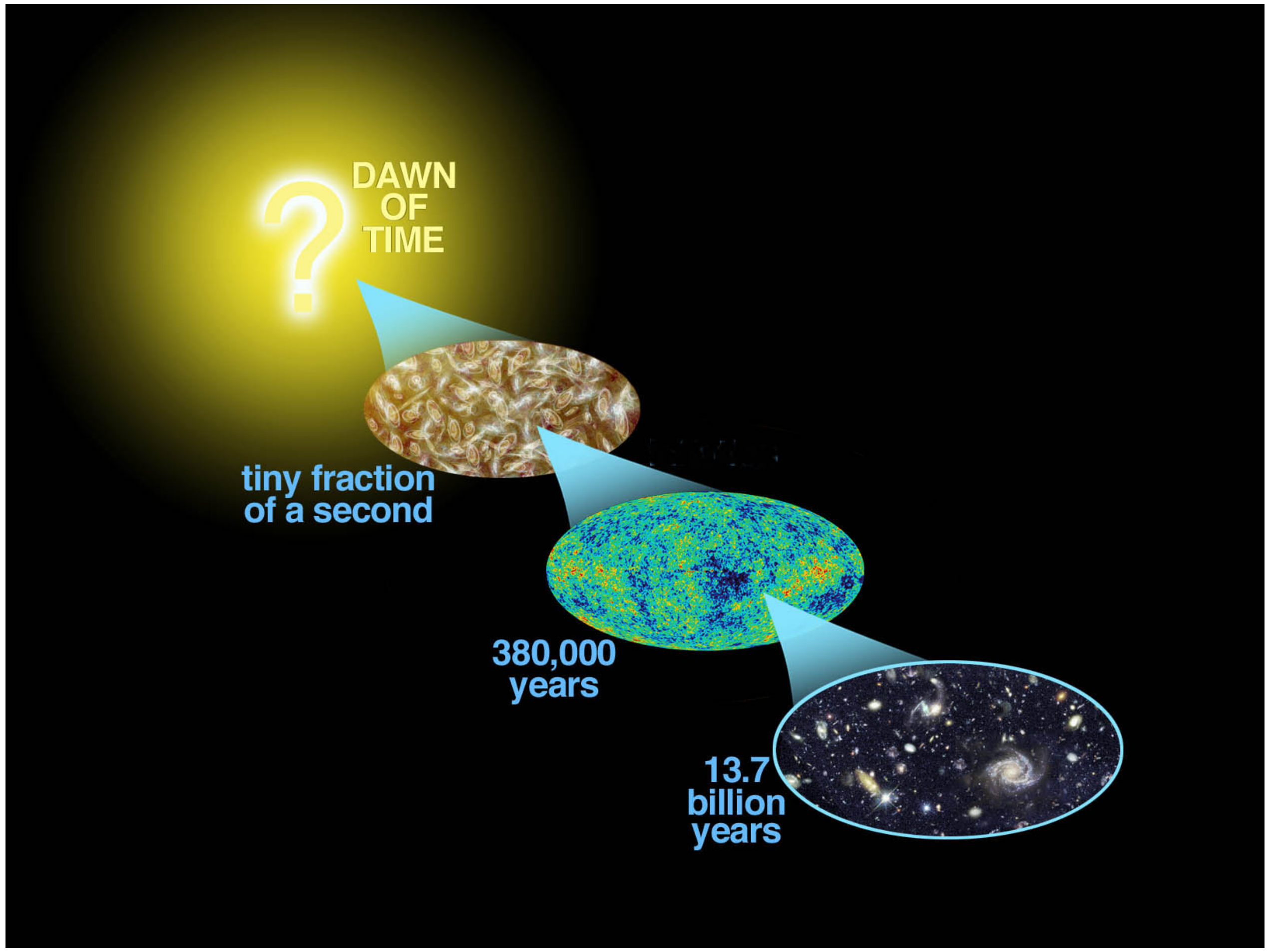
Infrastructure architecture:

- **A storage layer**: maximize capacity with enough disk bandwidth per server
- **A data-intensive processing layer:** maximize low level data access bandwidth and fabrics; fast sequential IO, large local disk storage, parallel file systems
- **A performance layer:** memory fabrics and bandwidth, CPGPUs, memory/disk hierarchy, interconnect bandwidth/latency
- **A development environment:** data and work flow engines with optimized data streaming, virtualization

Data-life cycle: from 10 years (community services, archiving) to months (personal reuse)

Turning large simulations into numerical laboratories

Numerical cosmology: a Big data HPC challenge



Universe is in rapid expansion, and mostly invisible (95%)

Dark energy (73%) is responsible of the cosmic acceleration

Dark matter (23%) is responsible of the visible structure in the universe

What is the nature of the dark energy?

Is it a supplementary exotic energy component (infinitely small!) or a modification of gravity laws at the scale of the Universe (infinitely large !)

The DEUS project

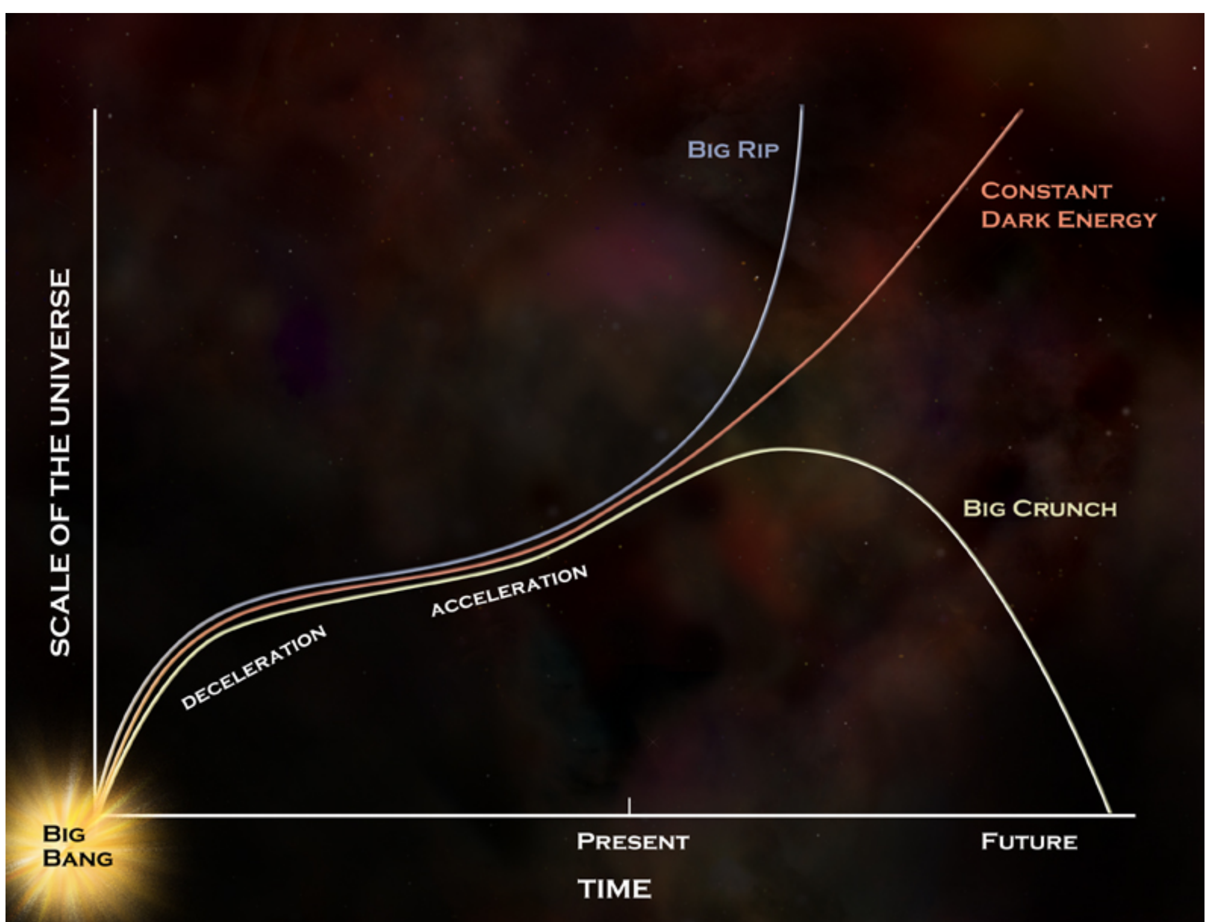
models of Dark Energy?

ΛCDM	WCDM	RPCDM	Etc.
Cosmological Constant	Phantom Model	Quintessence Model	

PROBLEM : The corresponding cosmological models cannot be distinguished with present data

BUT : Dark Energy has an influence on the formation of large scale structures

Compute the print of the Dark Energy on the structuration of our universe ?



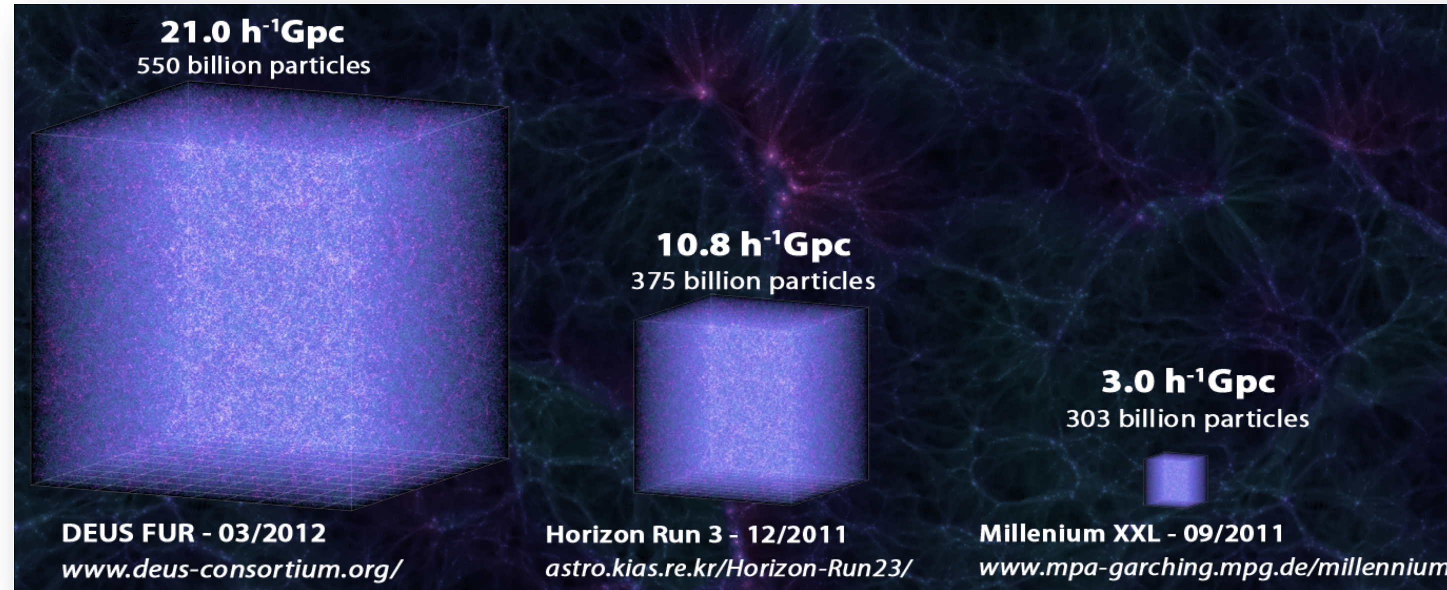
Require to follow the gravitational collapse of the matter all along the history of the universe and over ALL the observable universe

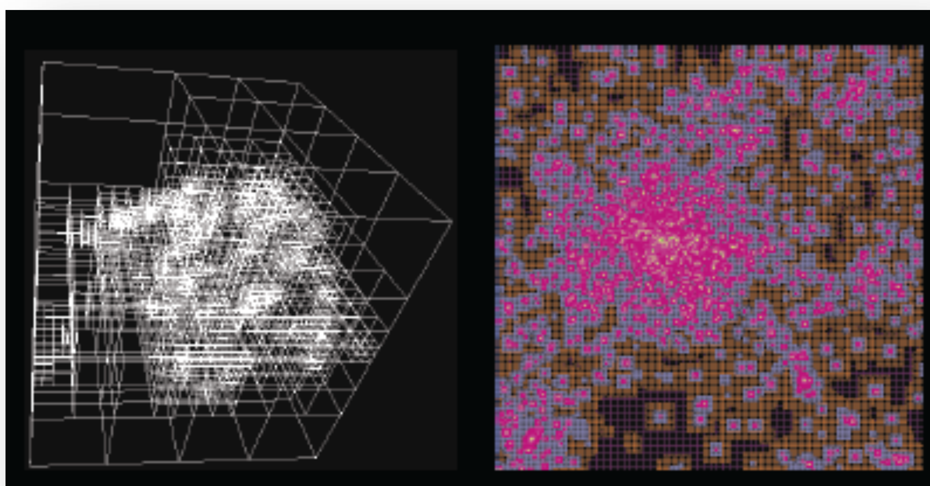
The DEUS project: J.-M. Alimi et al.

The AMA-DEUS application: N-Body simulation

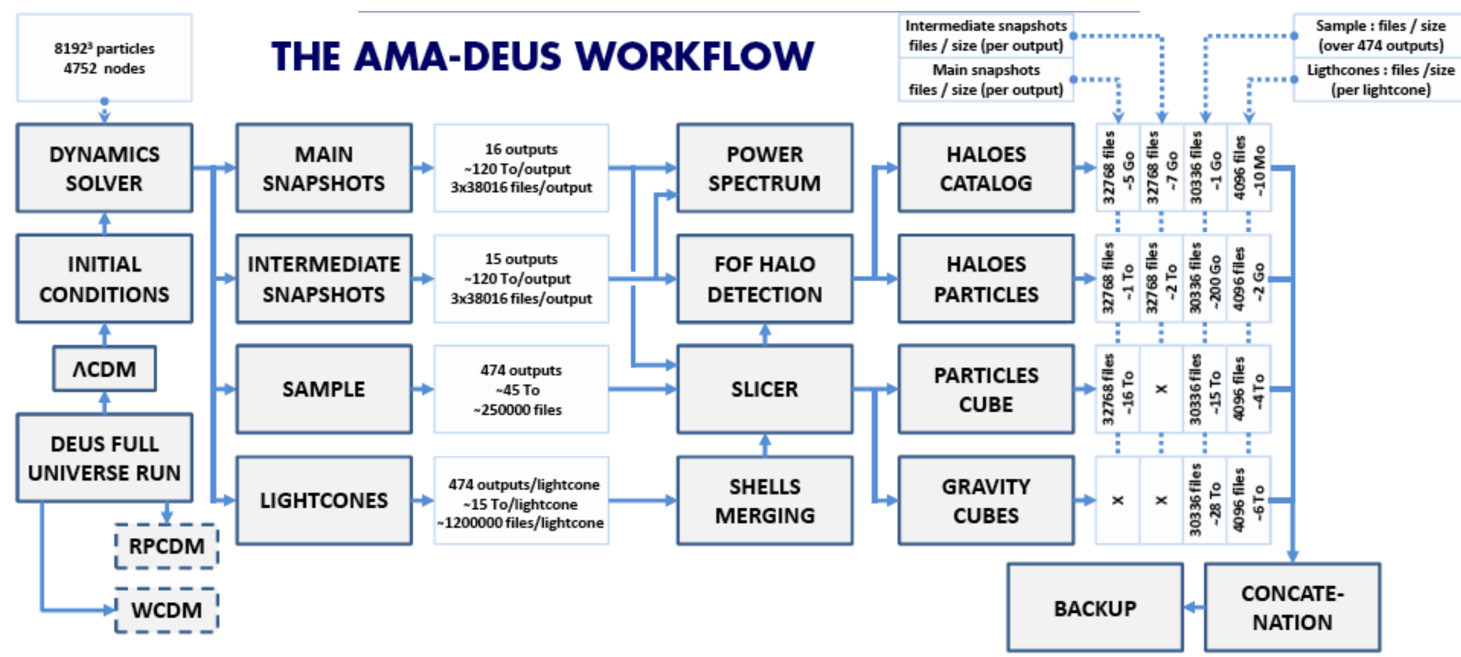
A TGCC-CURIE grand challenge

- 550 billion particles
- 2.5 trillion computing points
- 50 million CPU hours (> 5700 years)
- 76 032 cores & 300 Tb memory
- > 50 Gb/s data throughput (PFS)
- 1 500 Pbs reduced on fly to 1 500 Tbs



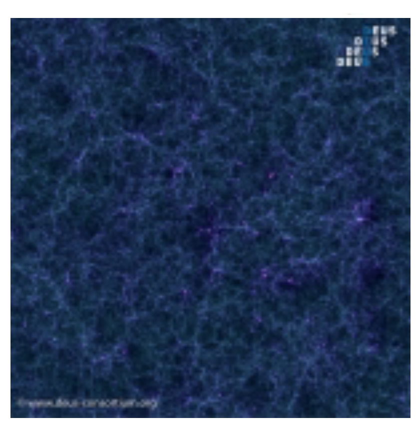


An end-to-end workflow !

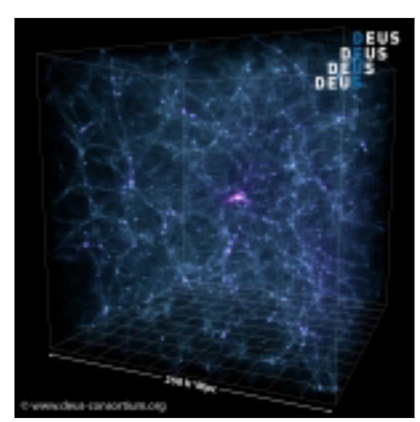


Challenges

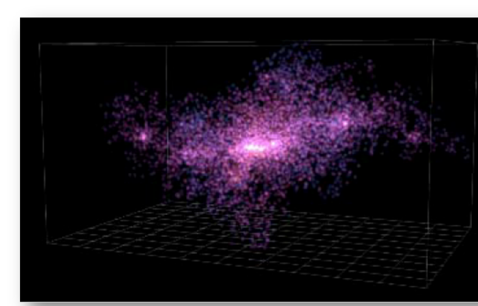
- dynamic load balancing
- smart parallel I/O optimisation through
- reduction of raw data (time) -> direct post-processing
- physical objects -> on-the-fly processing workflow



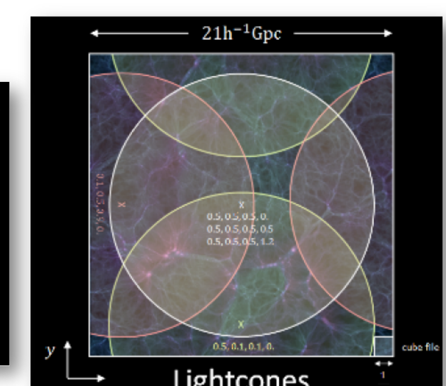
Snapshots
~16 x 16 TB



Samples
~40 TB

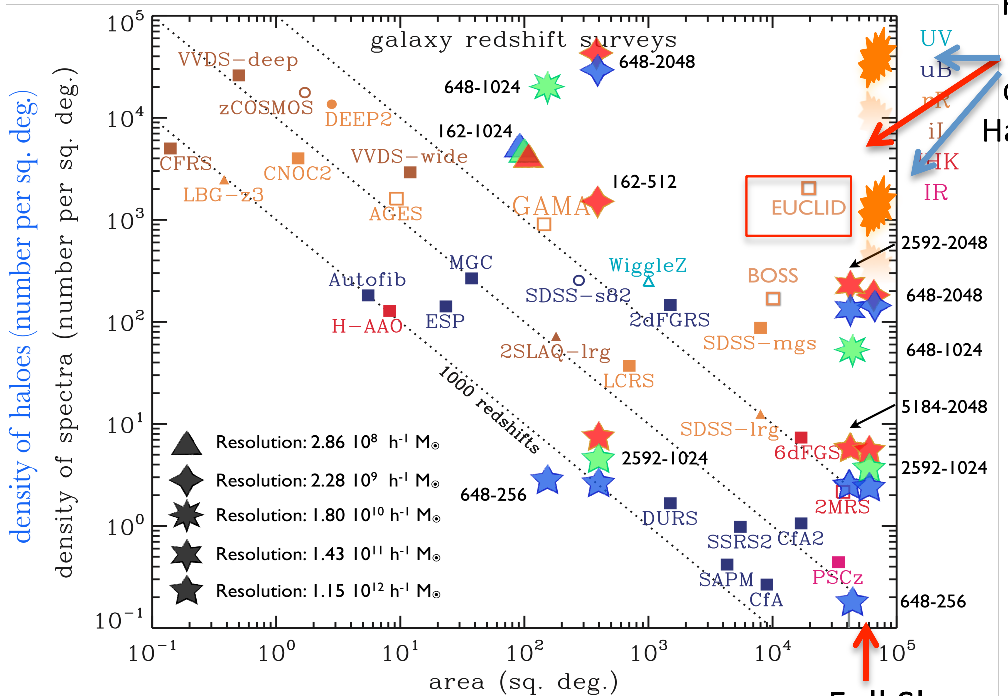


Halos/catalogs
~50 TB



Lightcones
~ 5x10 TB

HPC and large-scale instruments

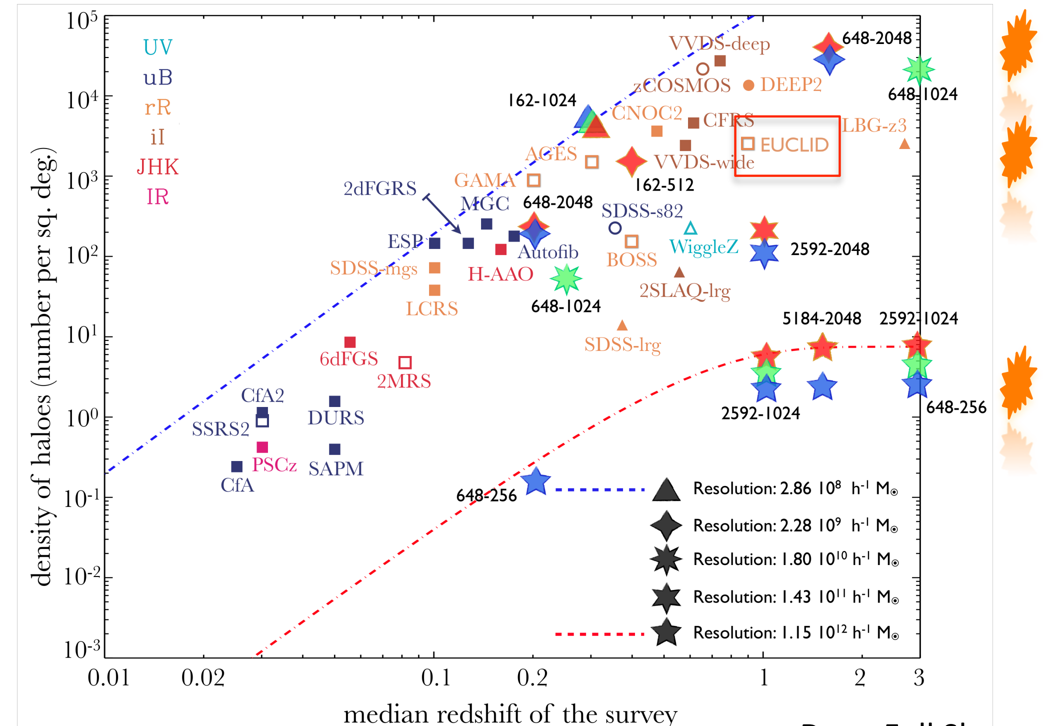


Full Sky
Halo
density
Halo mass

Full Sky

Density of Halos on the sky

Observational motivations of extreme scale computing in cosmology

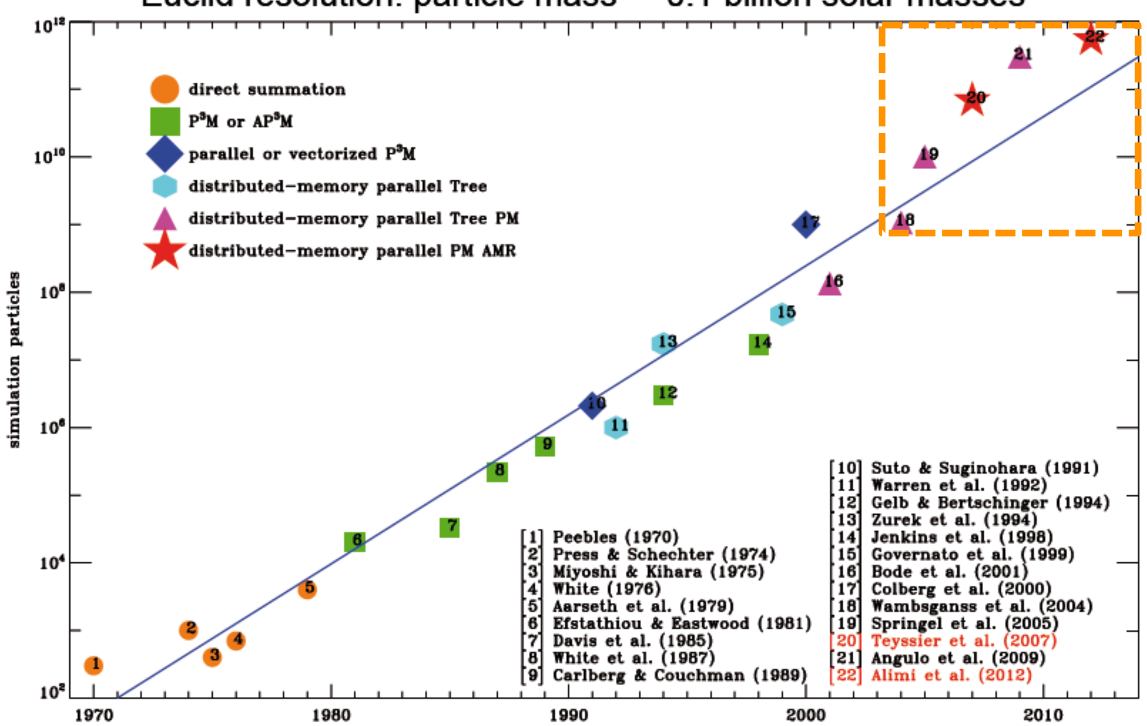


Density of Halos in depth

Deep Full Sky
(10^{10} h^{-1} $M_{O,}$ z≈7-8)
halo density halo mass

Cosmological N body simulations

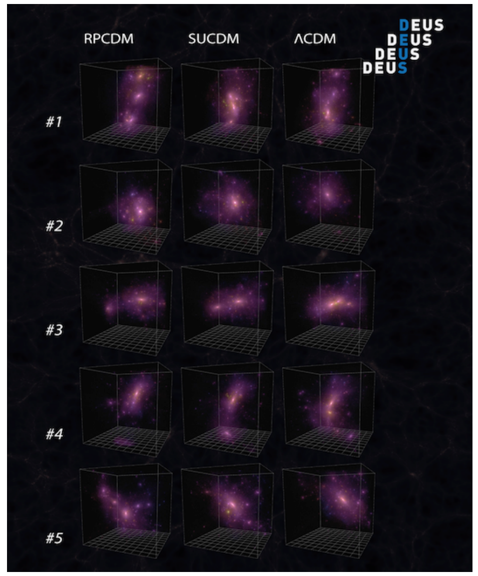
Euclid resolution: particle mass = 0.1 billion solar masses

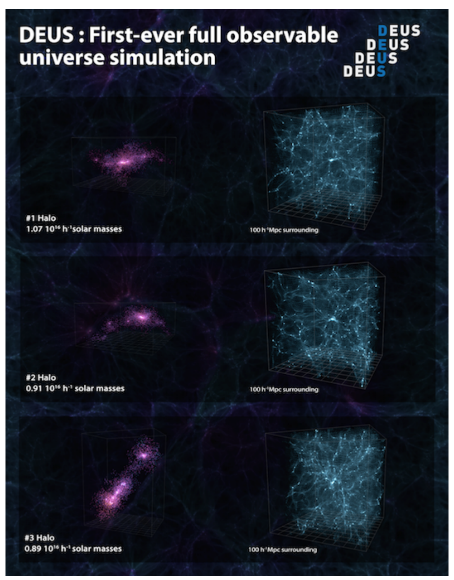


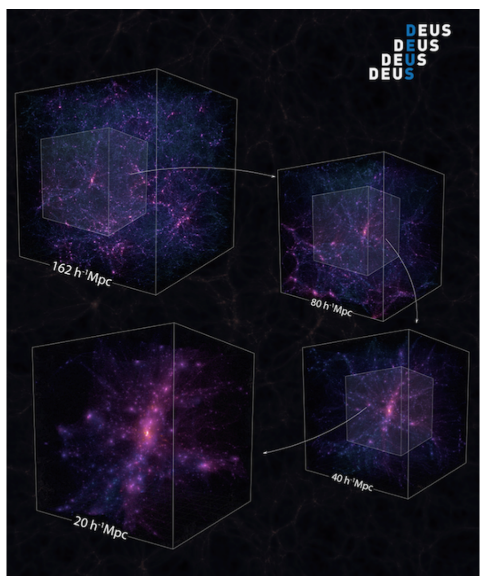
550 billion particles
(1 particule = Milky Way)
2.5 trillion computing points

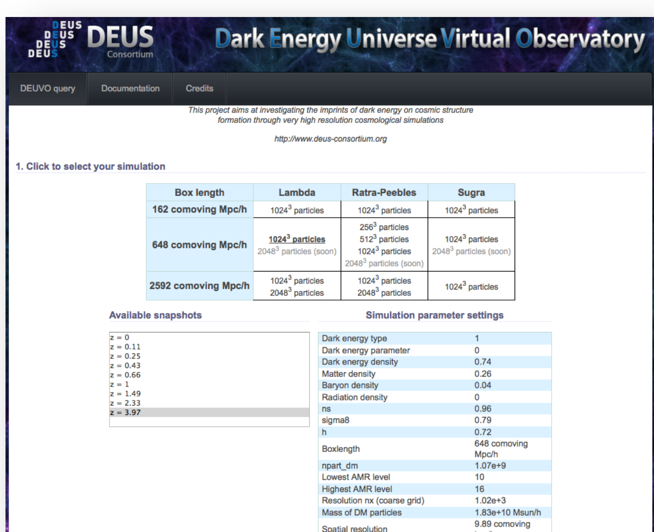
Link between big data HPC and large-scale instruments becomes closer and closer

Numerical laboratory: Data Analysis









Consortium DEUS

- scientific teams coordination
- DEUVO DB: physical objects and some raw data

Data Analysis

On-the-Fly

- MPI-based power spectrum
- MPI-based parallel Halos finder
- Halos properties

Post-Processing

- Higher-order statistics on matter field and Halos
- Topological analysis
- Dynamical analysis
- Visualisation

…

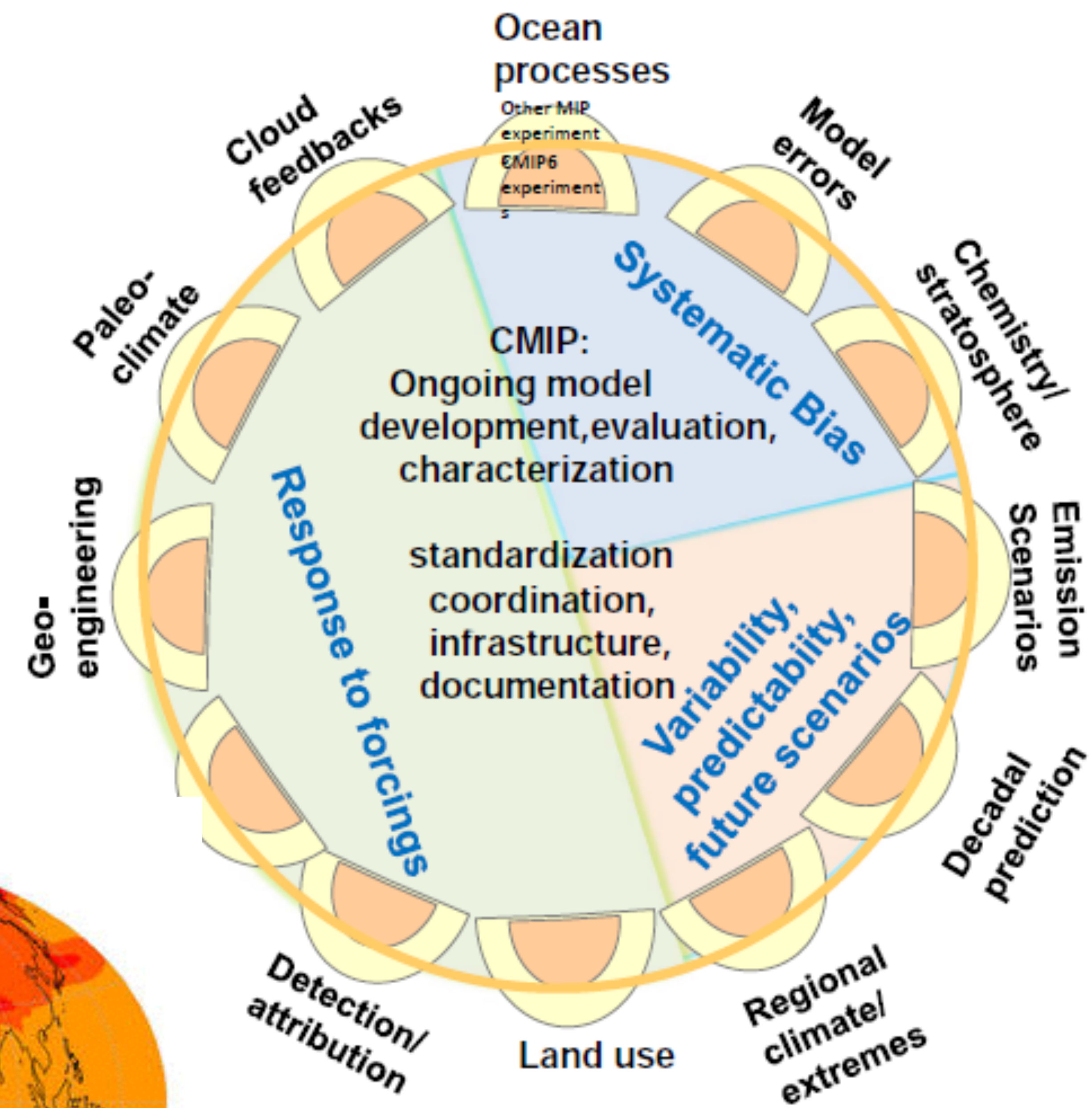
Data life-cycle: long-term preservation, annotation, publication

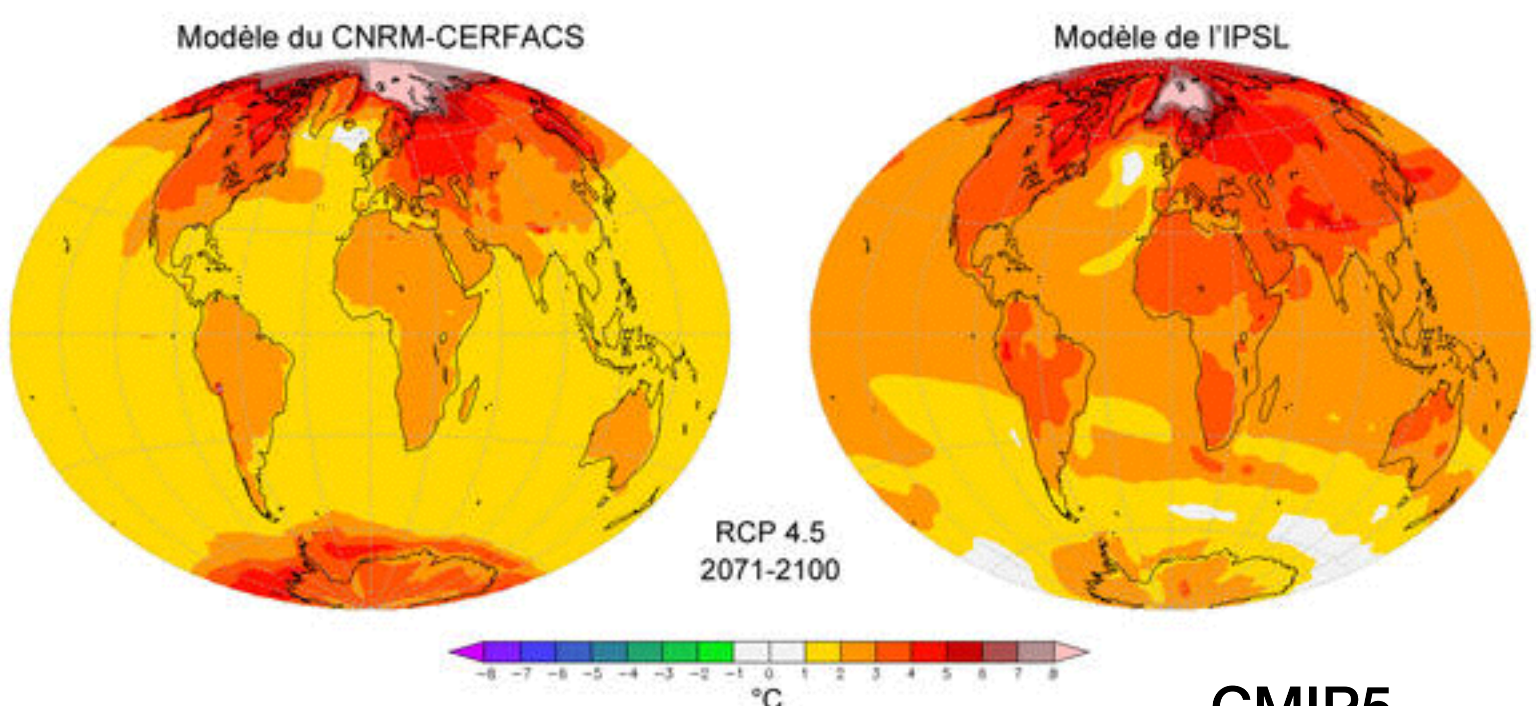
Climate simulation: next CMIP6 exercise

On-going discussion at the international level
Agenda: starts in 2016 and ends mid 2018

Distribution and file structure (ESGF, net-cdf)
Similar number of variables that previous CMIP5

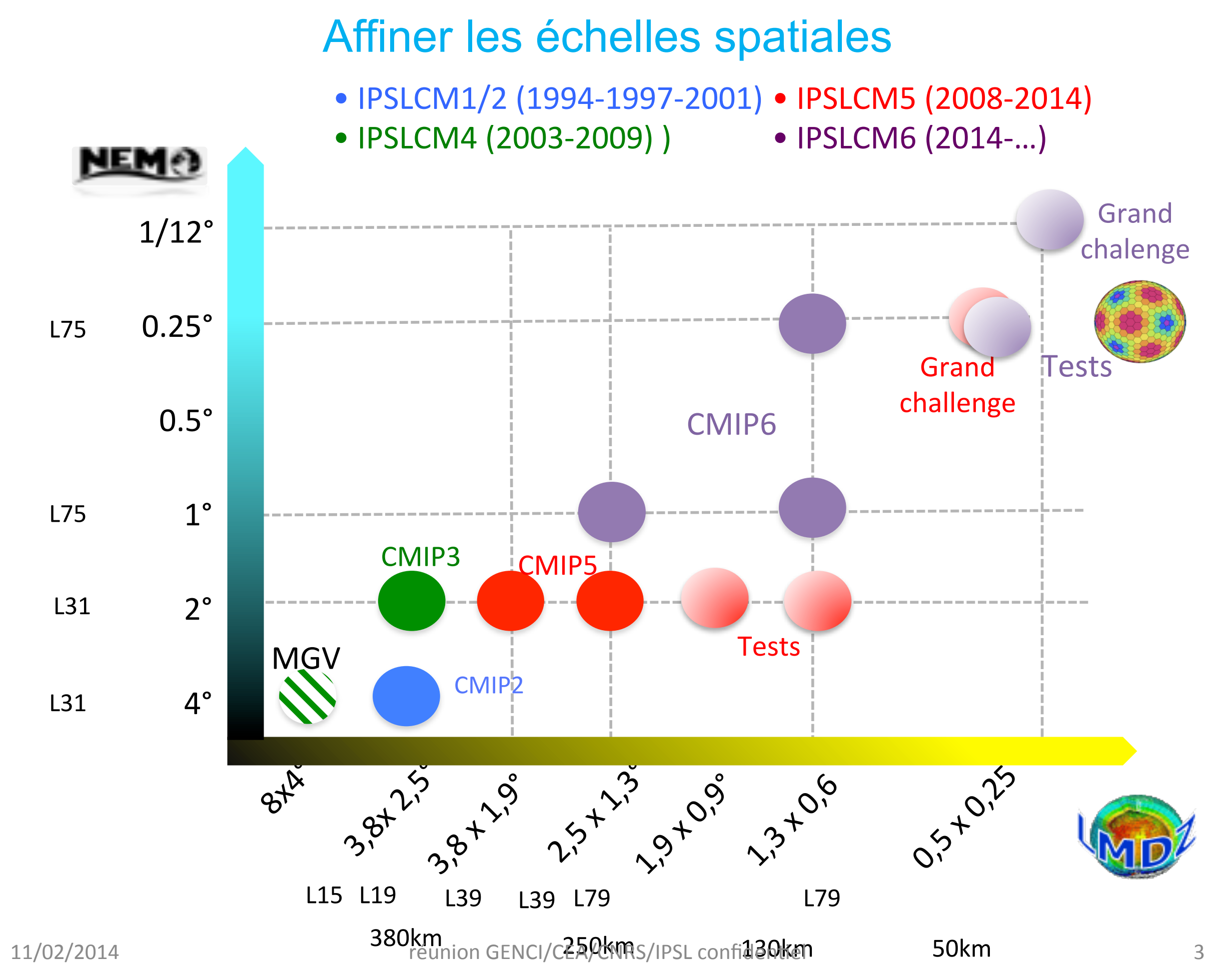
Research activity and scientific exploitation of the results

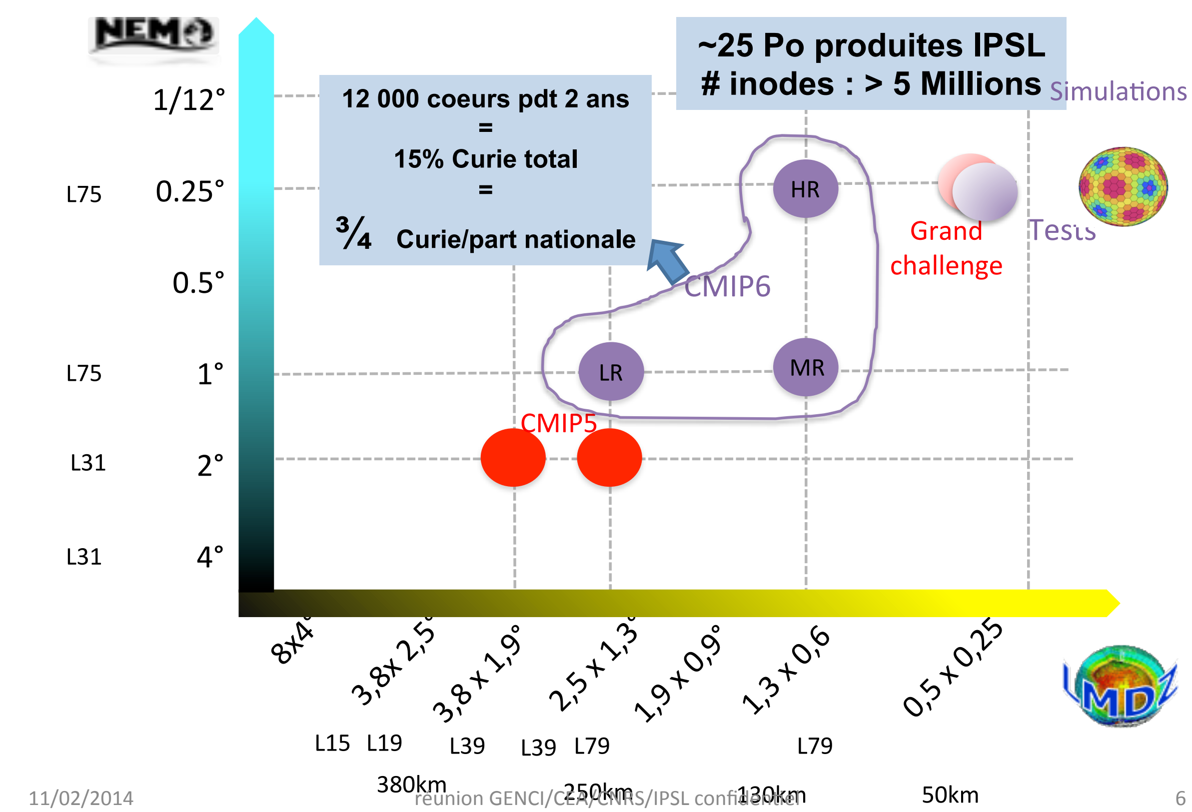




CMIP5

CMIP6: a big data HPC challenge



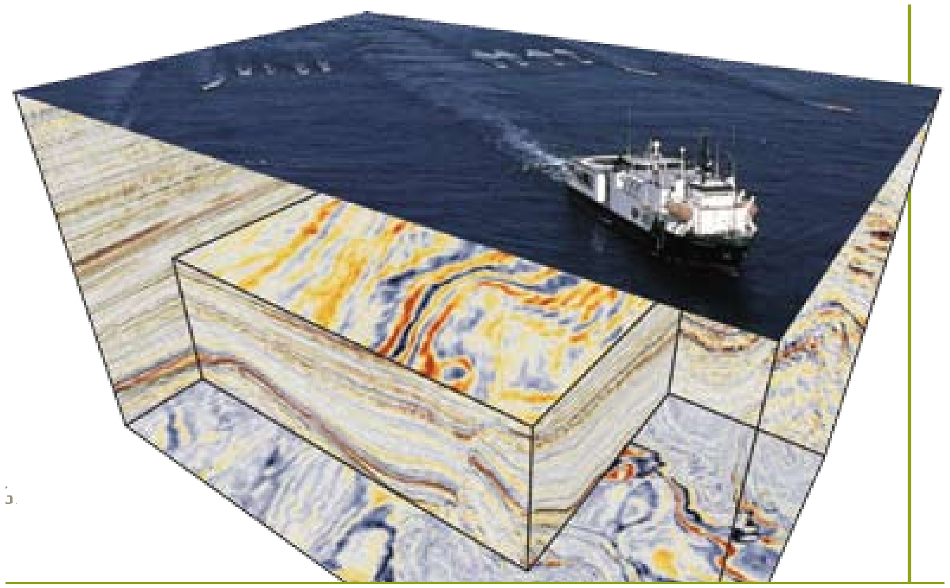


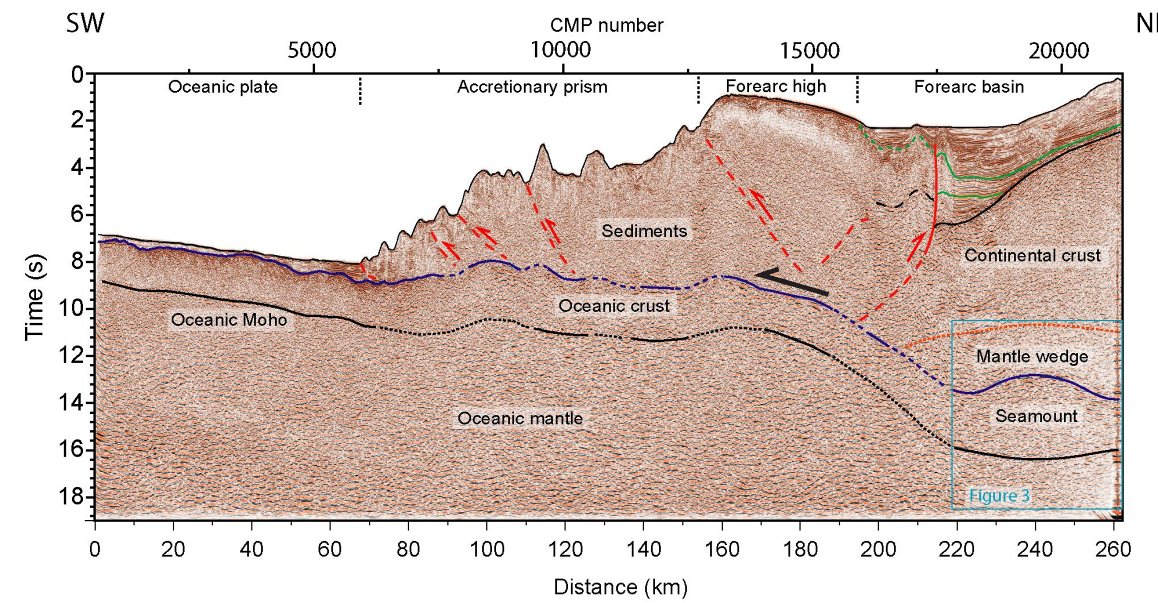
Agenda CMIP6 : **beginning 2016 - mid 2018**
Compute (12 000 coeurs/400 coeurs): x30 env. (CMIP5)
Storage volume (25 Po/2): x12 env. (CMIP5)
Number of files : (5 Mi) idem (but accumulation)
I/O throughput (25 Po/2ans) : x12 env. ie 50 To/day
Distributed data (/250 To): x10 env.
Cache Storage (data post-process: mixing, reduction) —> CMOR

Data-driven application: inversion and assimilation

Seismology full-waveform inversion example

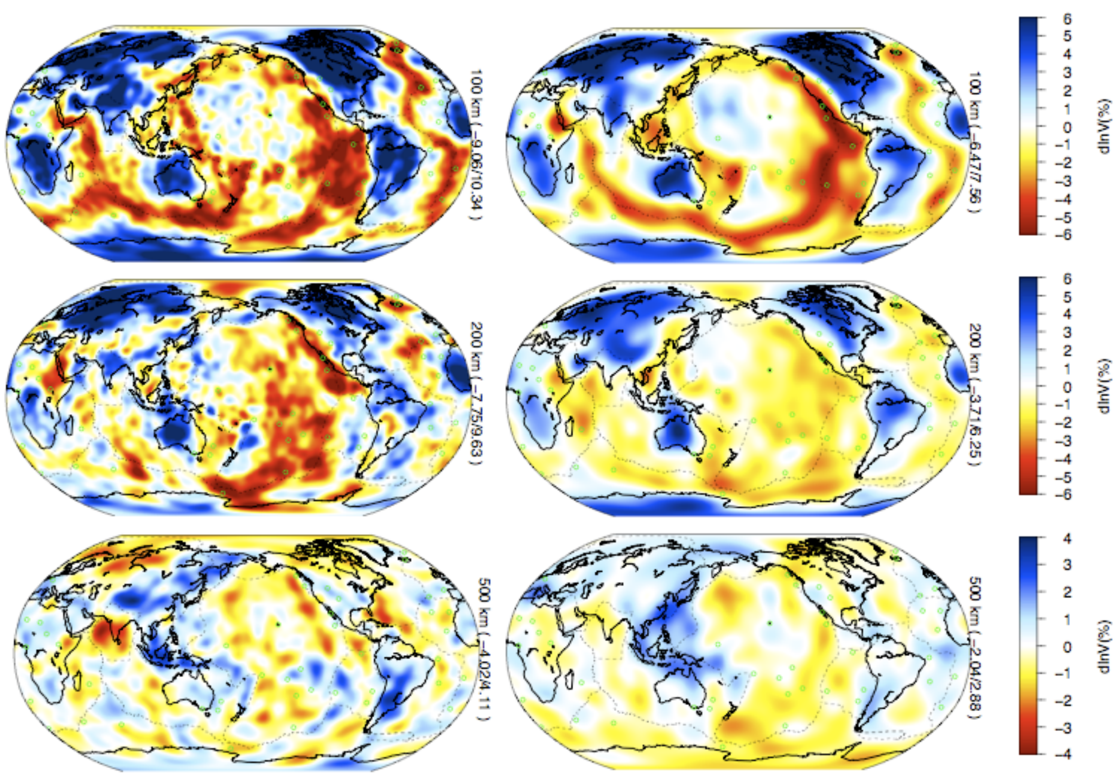
Exploration and marine geophysics

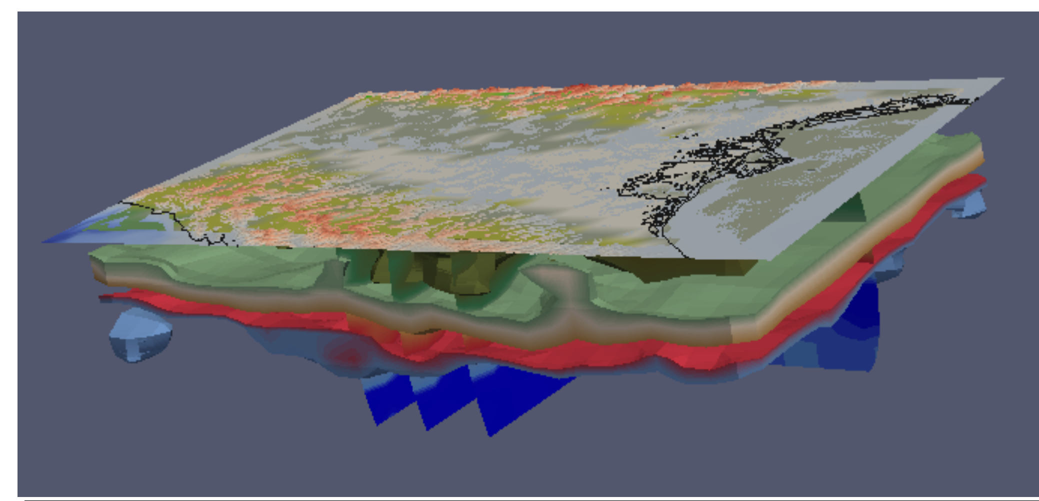


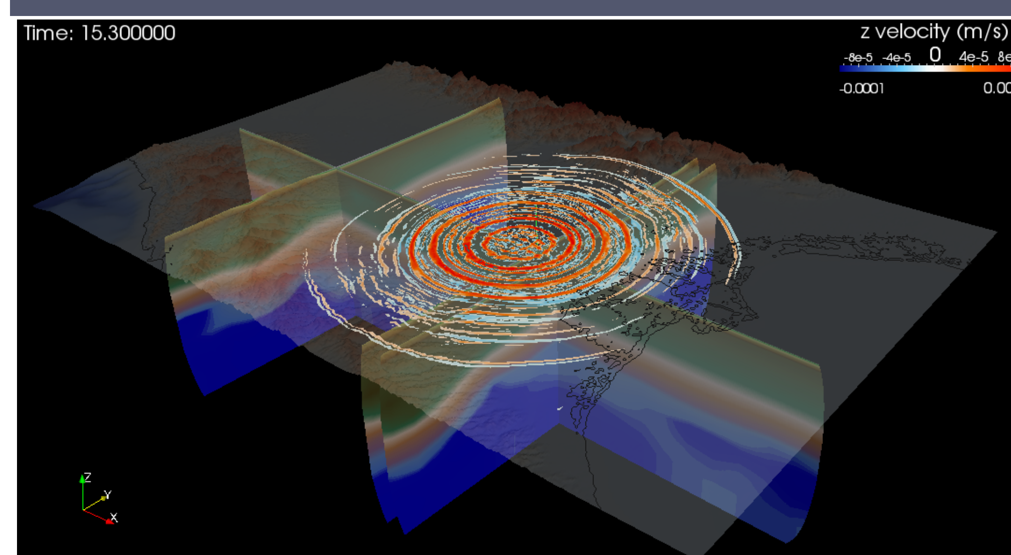


Regional tomography and earthquake imaging

Global scale tomography



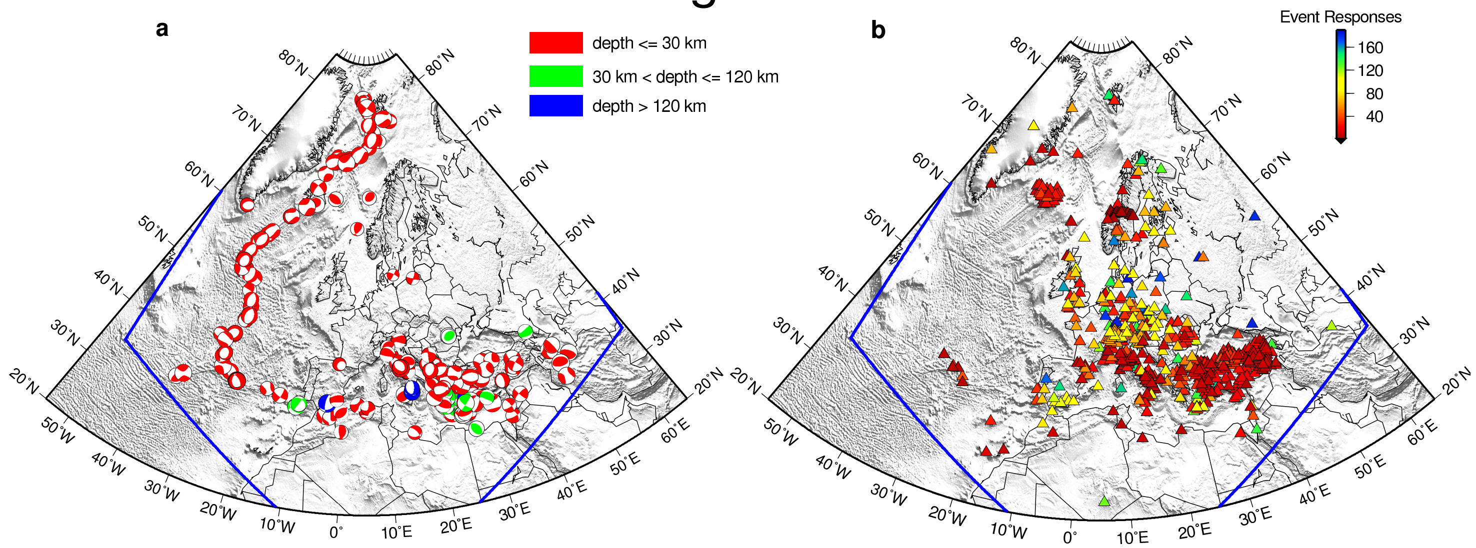


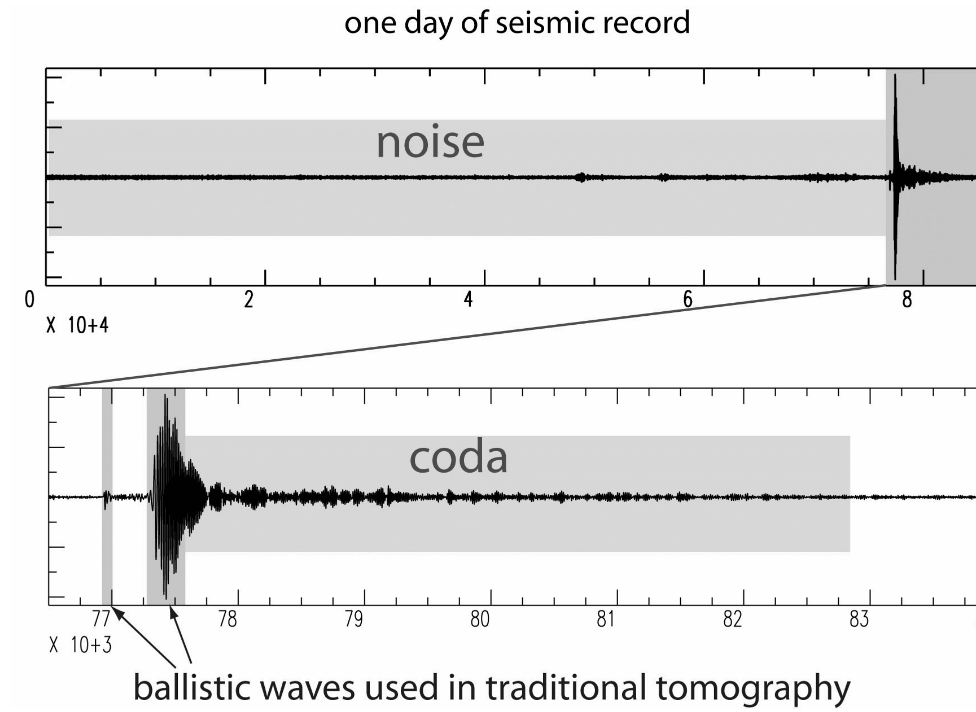


Full-wave form inversion: big data HPC problem

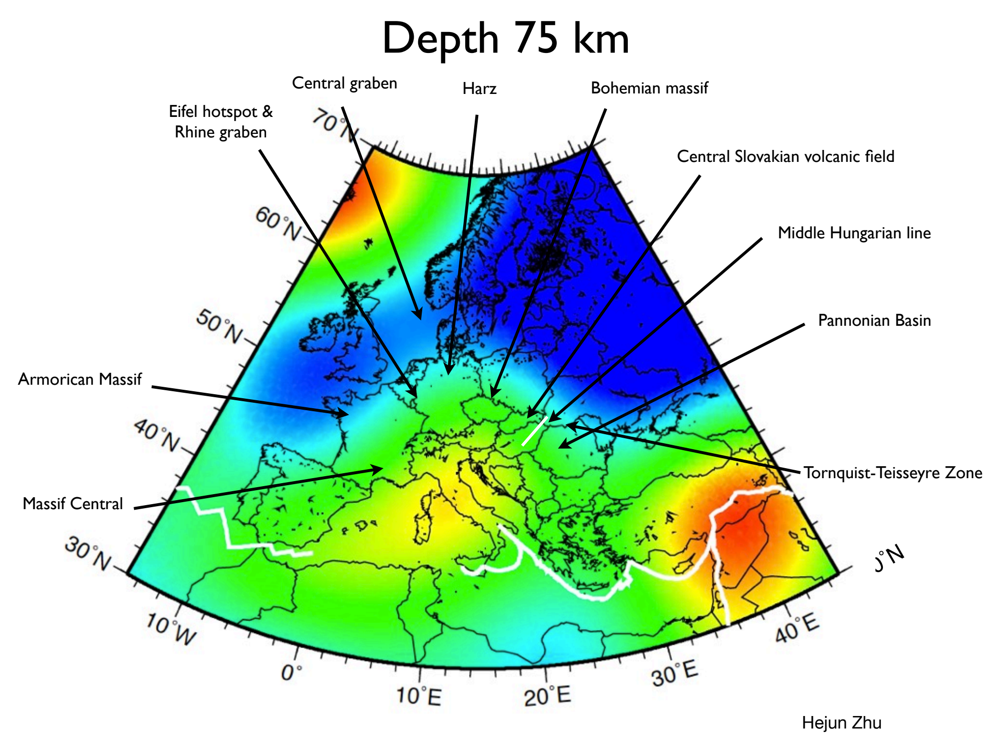
Adjoint Tomography of Europe

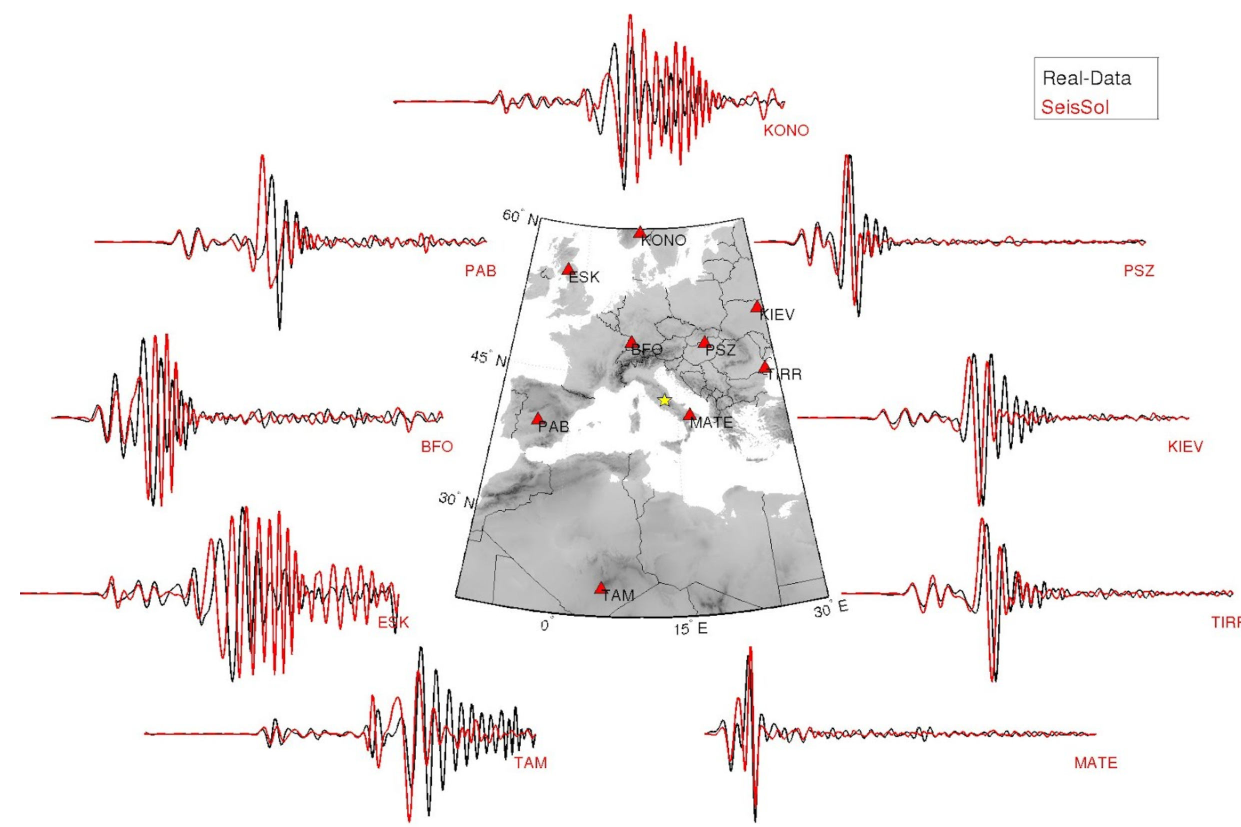
“Big Data”



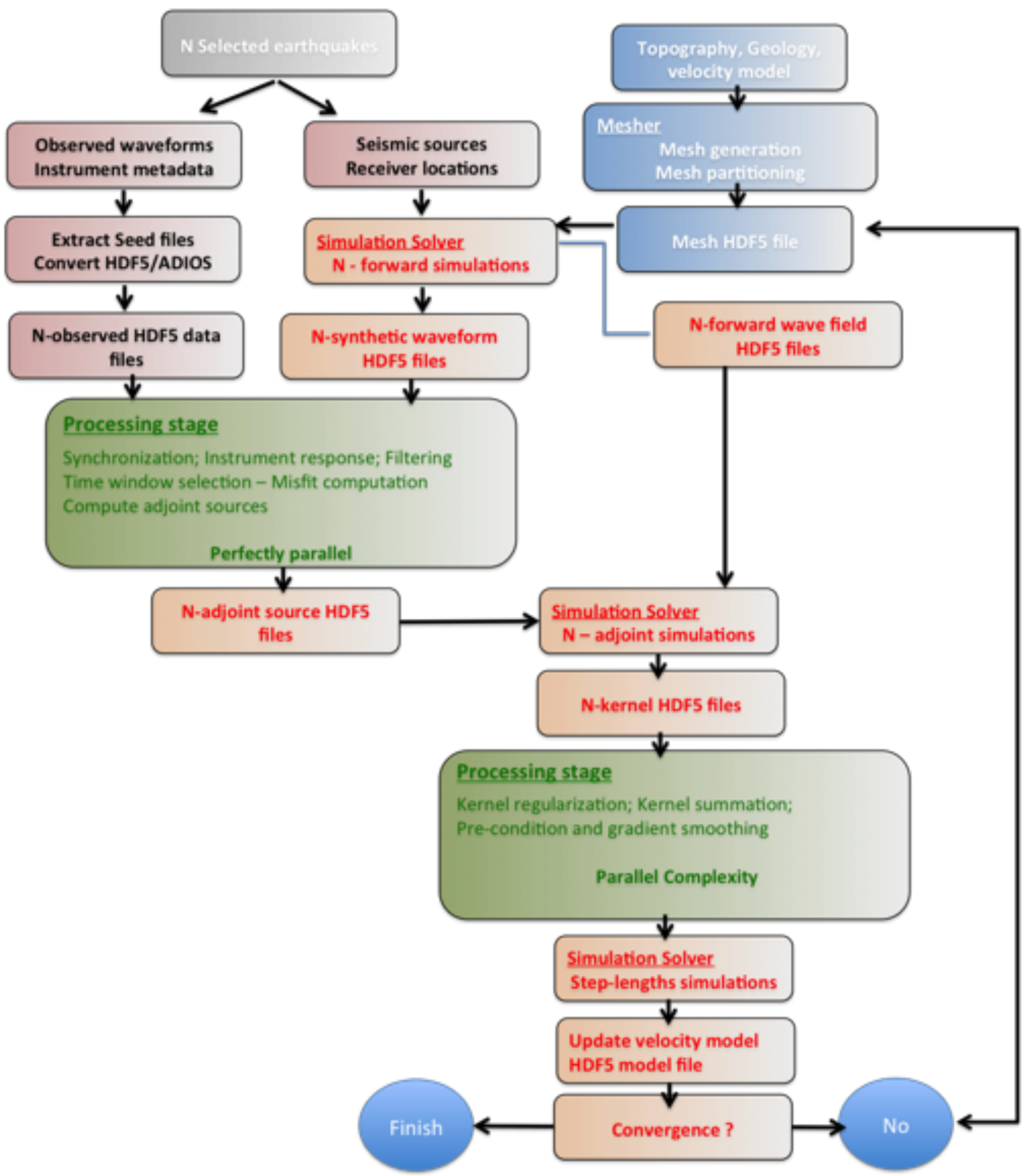


Tromp et al
Komatitsch et al





Orchestrated workflow: data-intensive & HPC



High-performance parallel codes

- forward and adjoint wave simulations

Waveform inversion

- non-linear Bayesian inversion
- adjoint-based inversion

Orchestrated workflow

- data-intensive analysis and HPC
- CPU and Data-intensive architecture

Big Data

- synthetics and observed wave forms
- Earth model and wave propagation
- I/O and CPU balance (~10s Gb/s, 100Tb per iteration)
- higher-oder abstract file format (HDF5)
- indexing and Data Bases

Parallel & adaptative mesh generation

- **~ billion of nodes**

HPC Data-intensive challenges

Large scale 3D simulation:

- multi-scale and multi-physics
- stochastic direct uncertainty evaluation

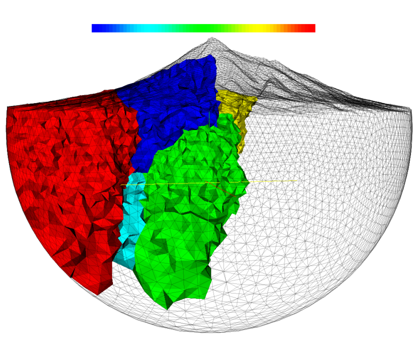
Inversion and Data assimilation:

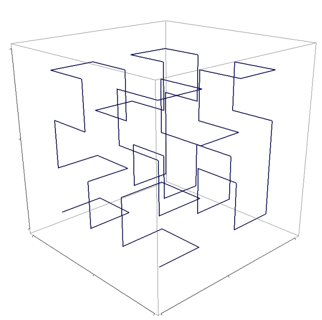
- adjoint-based methods: non linear iterations with large number of forward and adjoint simulations
- stochastic methods: inverse uncertainty quantification

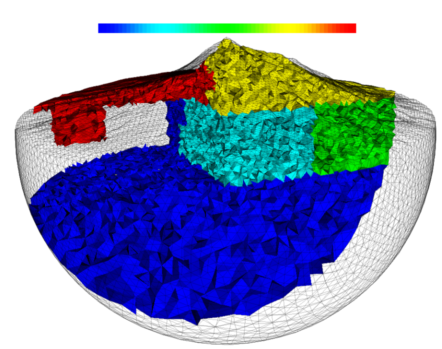
Orchestrated workflows:

- data analysis and modeling applications
- end-to-end applications

Hilbert SFC of level 2 and 64 sub-cubes







Domain decomposition by METIS (left) and SFC (right)

Scalability

Communication fabrics
Asynchronous time integration, vertical reuse
Explicit locality model (vertical/horizontal)
Parallel large system solver
Dynamic load balancing

Data-intensive HPC

Memory hierarchy and bandwith
Fast sequential IO
Hierarchy of storage HDD/SDD
Advanced data-structure and parallel filesystems

Multicore architectures

Mixed-hybrid parallel implementation
High-level task concurrency: asynchronous task parallelism; overlapping computation and communication
Self-scheduling at task level
Fault tolerance system

End-to-end analysis

Parallel unstructured mesh generation
Domain decomposition
Post-processing data-intensive data analysis
Data management

Data-life cycle: from 10 years to months

Data in HPC simulations and inversions

Largest simulations and inversions in Earth and Universe sciences approaches petabytes

- From supernovae to turbulence
- From climate evolution to oceanographic circulation
- From magneto-hydrodynamics to seismic wave simulation/inversion

Create new challenges:

- How to write enough output (speed of checkpointing)
- How to move the data (high-speed networking and transfer protocol)
- How to look at it (render on top of the data, drive remotely)
- How to analyse (value added services, big data analytics …)
- Architecture (supercomputers, DBs servers …)

Usage scenarios

Large variations in Data lifecycle and commitments

- On-the fly analysis (immediate, do not keep)
- Private reuse (short/mid term, local)
- Public service portal (mid /long term, community commitment)
- Archival and curation (long term, community commitment)

Different from today supercomputer usage patterns

Wide range of data access patterns, from high speed streams to large random access

Use cases: **turbulence, cosmology, climatology, seismic tomography**

Architectural challenges and strategy

System and an infrastructure for the posterior analysis

Where should the data be stored

- Not directly at the supercomputer (too expansive storage)
- Computations and visualisations must be on top of the data
- Need high bandwidth to the data source

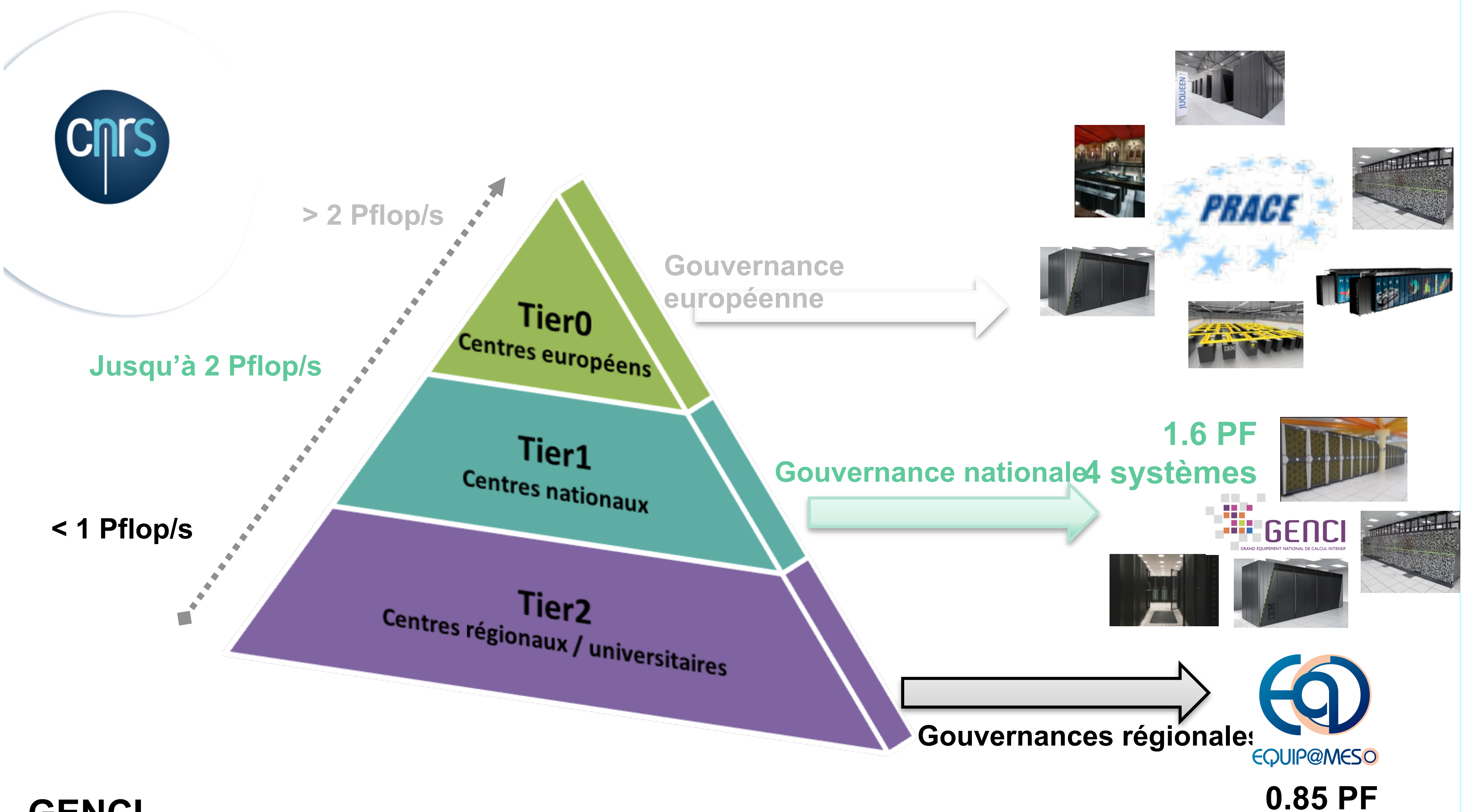
Scheduling of complex I/O access patterns

- Data bases - are they scalable?
- Extended file management systems and model ?
- Augmented with added-value analytics

Data organisation

- Most of those applications are not hard to partition (scale-out)
- Use fast, cheap storage for data streaming (sequential access)
- Tier of large memory systems (random access)

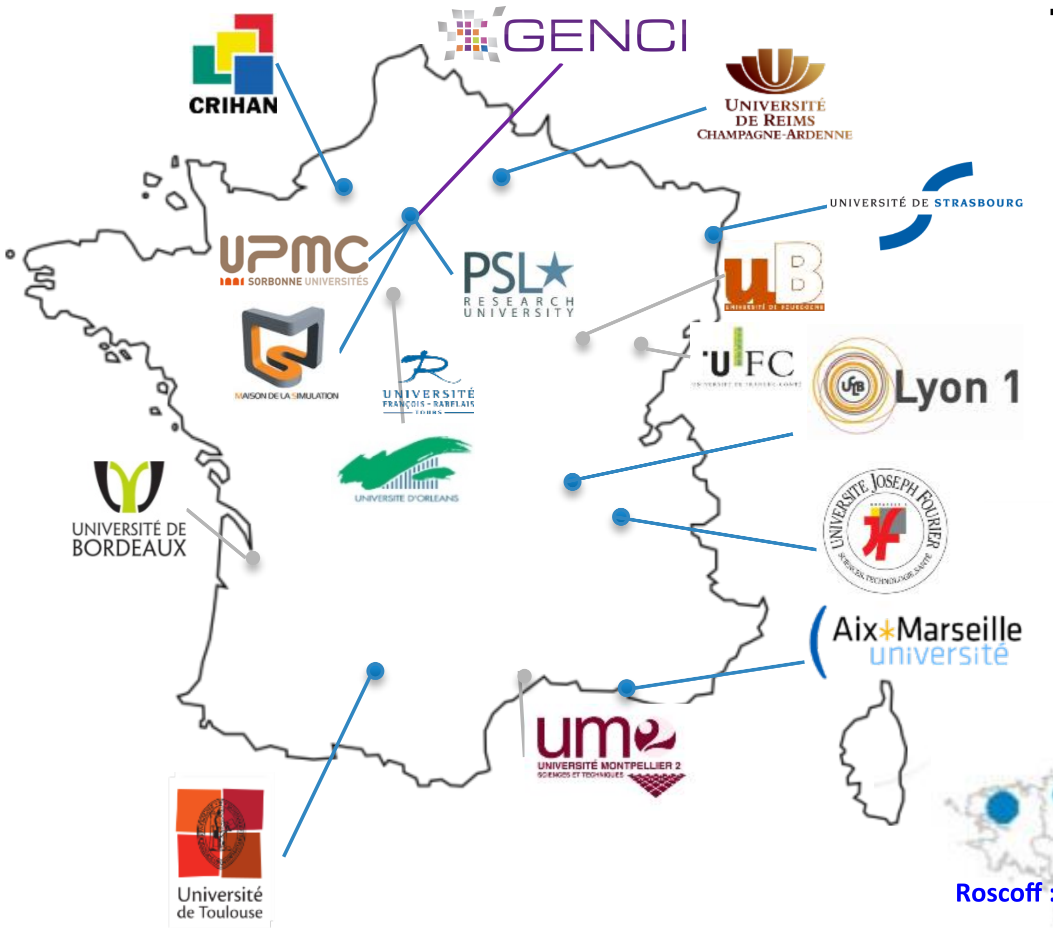
HPC ecosystem



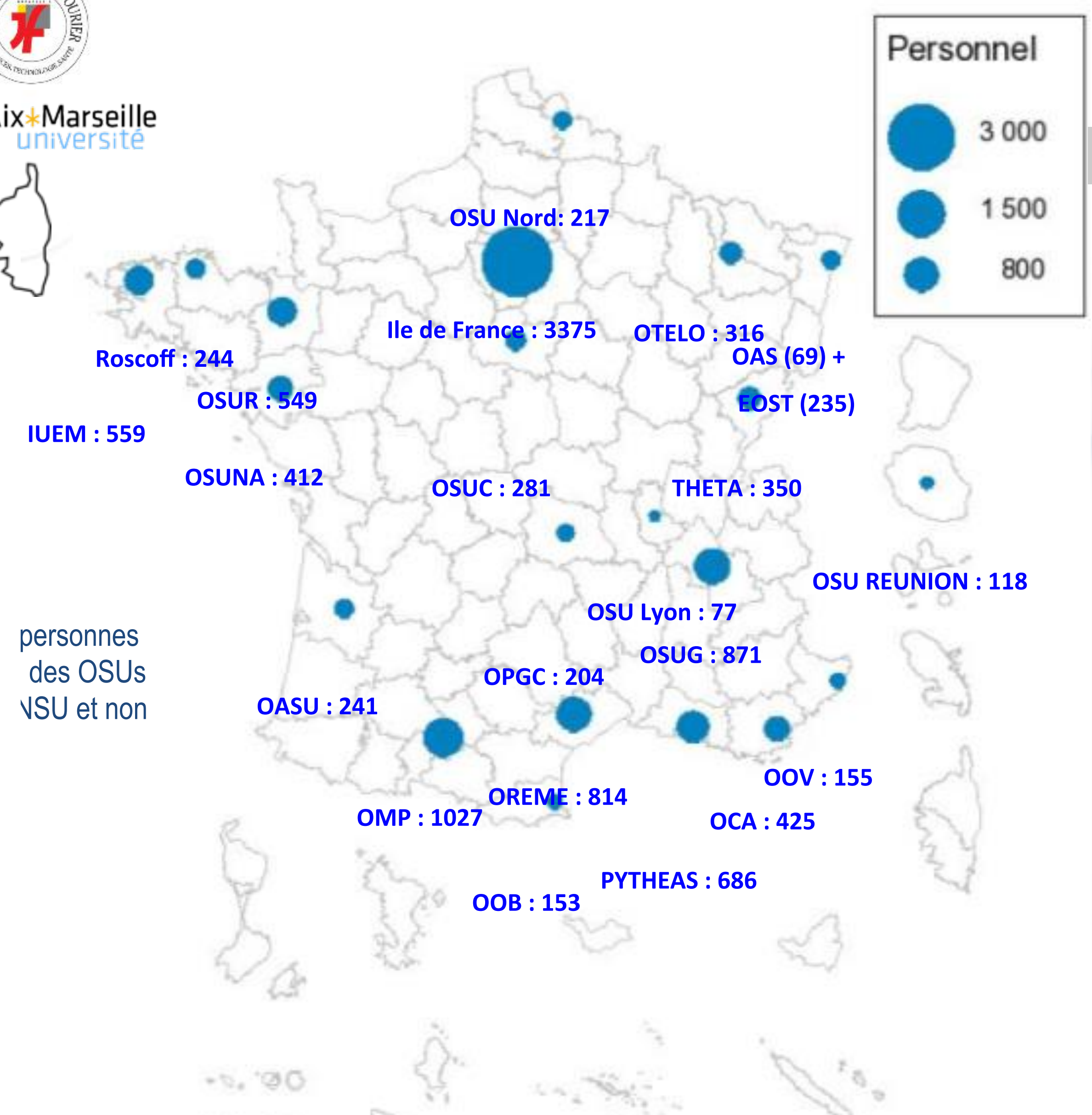
GENCI

- Société civile (créée en 2007)
- Partenaires: CNRS (IDRIS), CEA (TGCC), MENSR (CINES), CPU, INRIA
- Contribution française dans PRACE

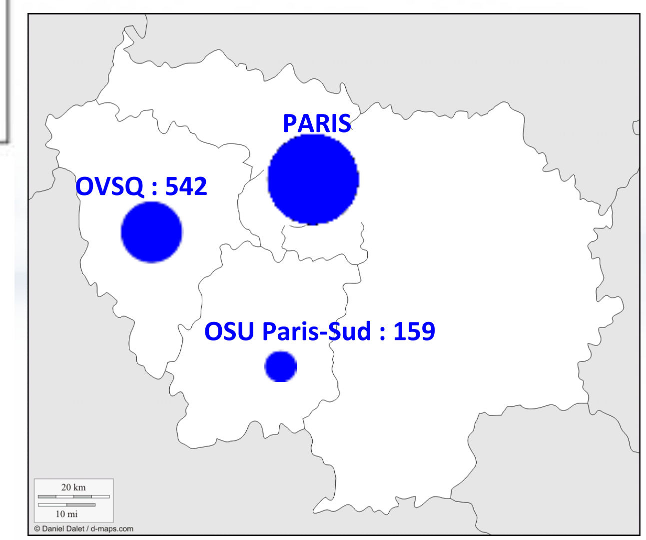
Synergy HPC and Data infrastructures



Tier-2: meso centres of Equipex@meso (GENCI)



Répartition des OSU en Ile de France



PARIS :
- Observatoire de Paris (1007p.)
- EFLUVE (251 p.)
- Ecce Terra (1085 p.)
- IAP (134 p.)
- IPGP (318 p.)

OSUs: Observatoires des Sciences de l'Univers (CNRS-INSU)

Data-centric HPC : an innovative scientific instrument

A Data-Scope *(A. Szalay)*

A scientific instrument capable of "observing" - explore, analyse and model - massive and complex data generated by large-scale instruments, observation and monitoring systems, and by numerical simulations in the Earth and Universe sciences.

- Innovative methods, software, technologies for large scale computation and massive data analysis that can ultimately be deployed on the national and European HPCs
- Emergence of a multi-disciplinary expertise in data-intensive computing and analysis across CNRS-INSU scientific teams and observatories
- Exploitation and valorisation of massive data generated by large-scale instruments, experiments and observation and monitoring systems in synergy with the "Observatoires des Sciences de l'Univers" (OSUs) and their Data Centers
- Formation and training in high-performance and data-intensive analysis of new generation of young researchers in Earth and Universe sciences through Masters and the Maison de la Simulation.
- Consortium behind applications development and generated data exploitation/distribution
- Data life-cycles are of primary importance: from > 10 years (preservation, curation, annotation), to ~months
- Emergence of multidisciplinary scientific and technology poles around BigData and data-intensive HPC

A new infrastructure urbanisation

- **Obervatories - Data centres - HPC infrastructures**