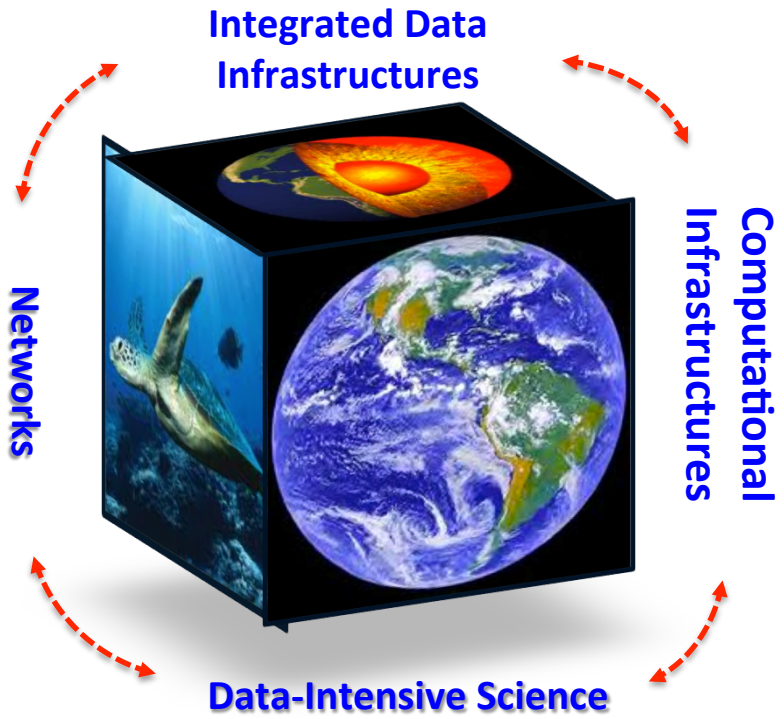
Big Data and Extreme Scale Computing challenges in Solid Earth Sciences

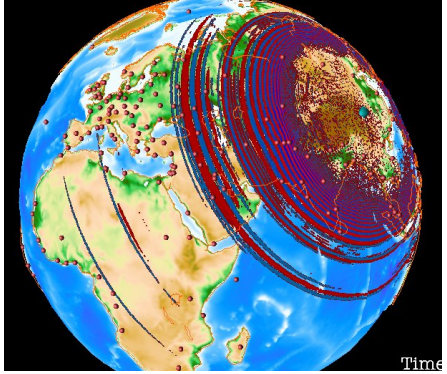


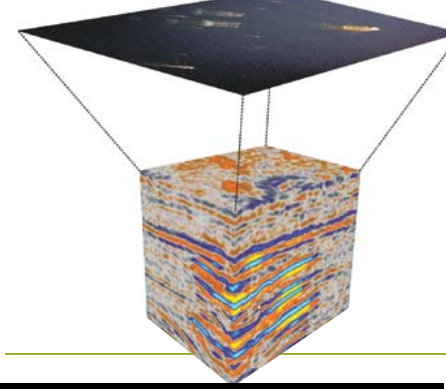
Jean-Pierre Vilotte, IPGP (CNRS-INSU)

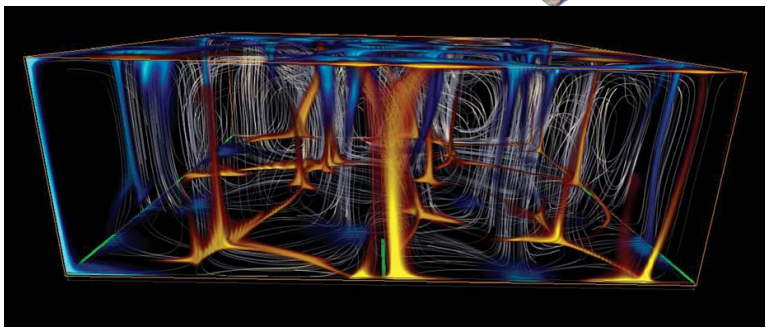
and the VERCE Team

Charleston, May 1, 2013

Data-intensive Research







International structuration

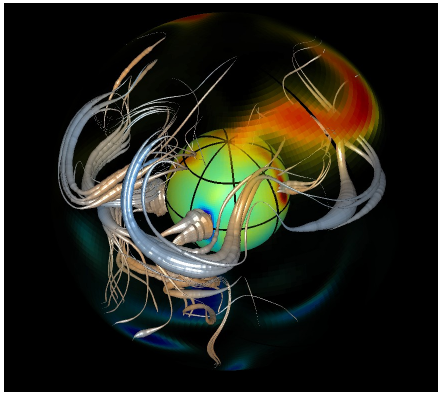
- Global observation and monitoring systems
- Integrated Distributed Data Archives
- Data and metadata format standards

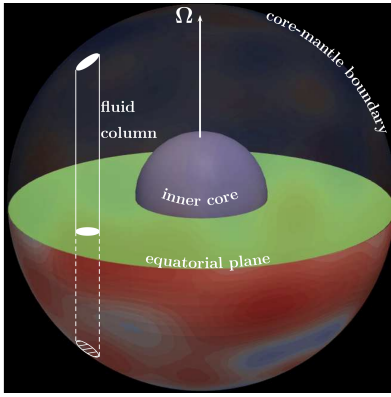
Scientific challenges

- Understanding Earth's dynamics and structures
- Imaging Earth's interior and seismic sources

Augmented societal applications

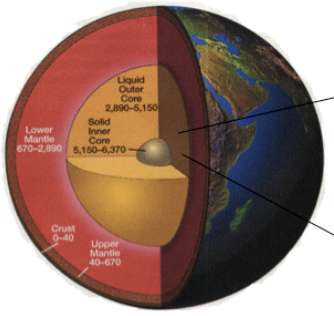
- Natural hazard and risk mitigation;
- Energy resources exploration and exploitation;
- Underground wastes and carbon sequestration;
- Nuclear test monitoring and treaty verification

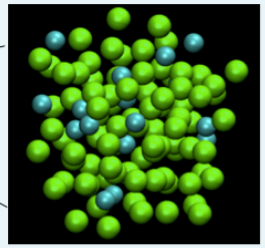


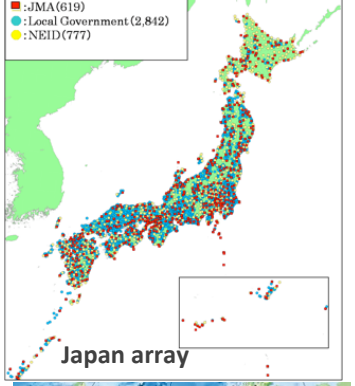


Data-intensive computing challenges

- Source detection and waveform data analysis
- High resolution inversion and data assimilation
- Quantification of forward/inverse uncertainties

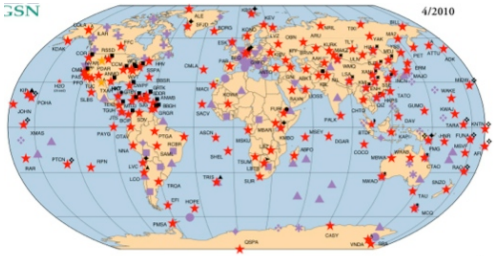




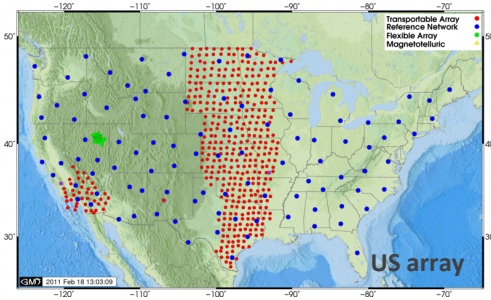


Integrated Seismology Data Infrastructure

- **J**apan **M**eteorological **A**gency
- **J**apan **N**ational **I**nstitute for **E**arthquake **D**isaster
- **E**uropean **I**ntegrated **D**ata **A**rchives
- US **I**ncorporated **R**esearch **I**nstitutions in **S**eismology
- . . .



FDSN Global array



Data-intensive e-science environment

Data-intensive Research

Data statistical analysis and detection

Data modelling: simulation, inversion, assimilation

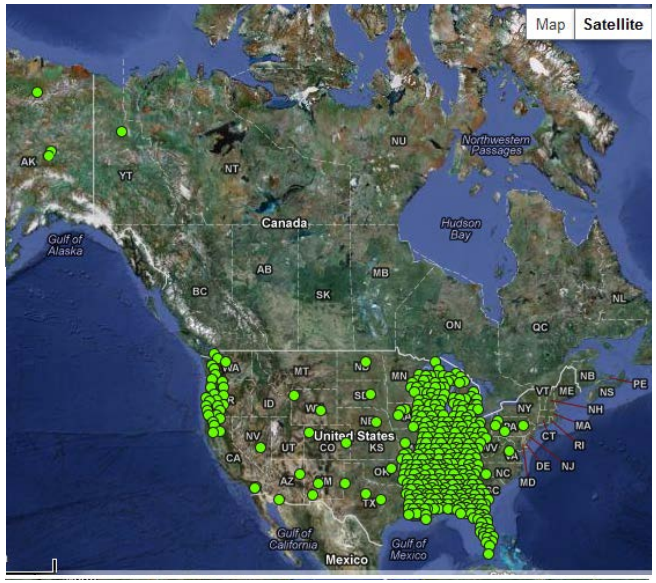
Data/Computing Infrastructure providers

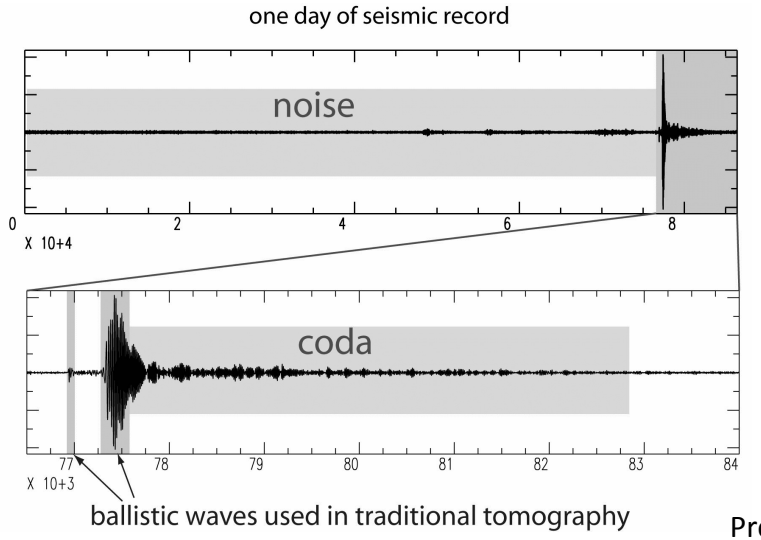
Public: HPC, Grid, Cloud

Private: Seismological Institutions and Data Centres …

Earth's interior imaging and dynamics: noise correlation, waveform analysis	**Natural hazards: new tools for monitoring earthquakes, volcanoes, and tsunami**	**Interaction of solid Earth with Ocean and Atmosphere: environment, climate changes**

Data-Intensive statistical analysis: Seismic noise correlation

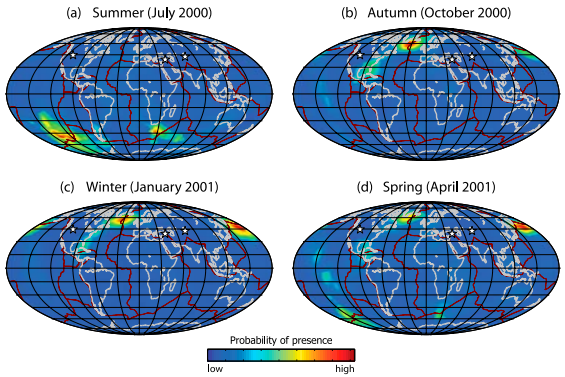




Property changes

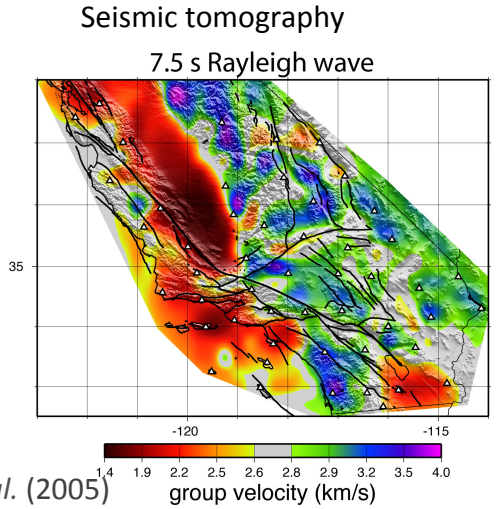
Exploiting the statistical coherence in space and time of continuous waveforms records from dense arrays of broadband and strong motion instruments

Deep oceanic wave sources

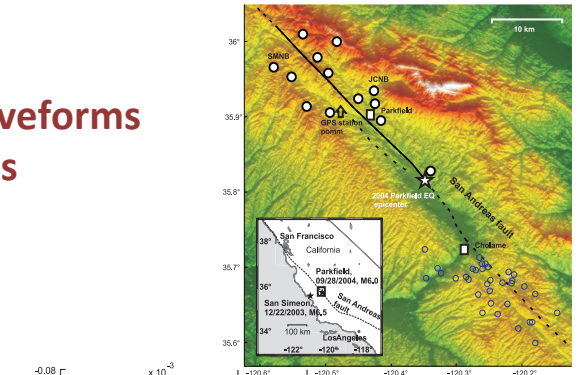


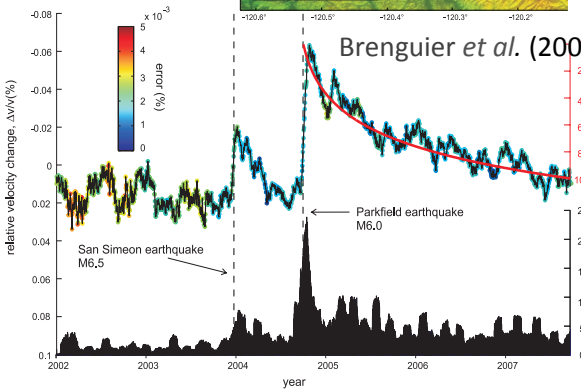
Landes *et al.* (2010)

Seismic tomography



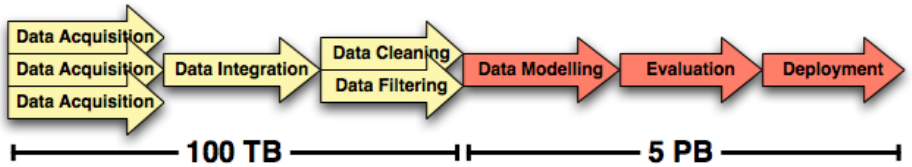
Shapiro *et al.* (2005)

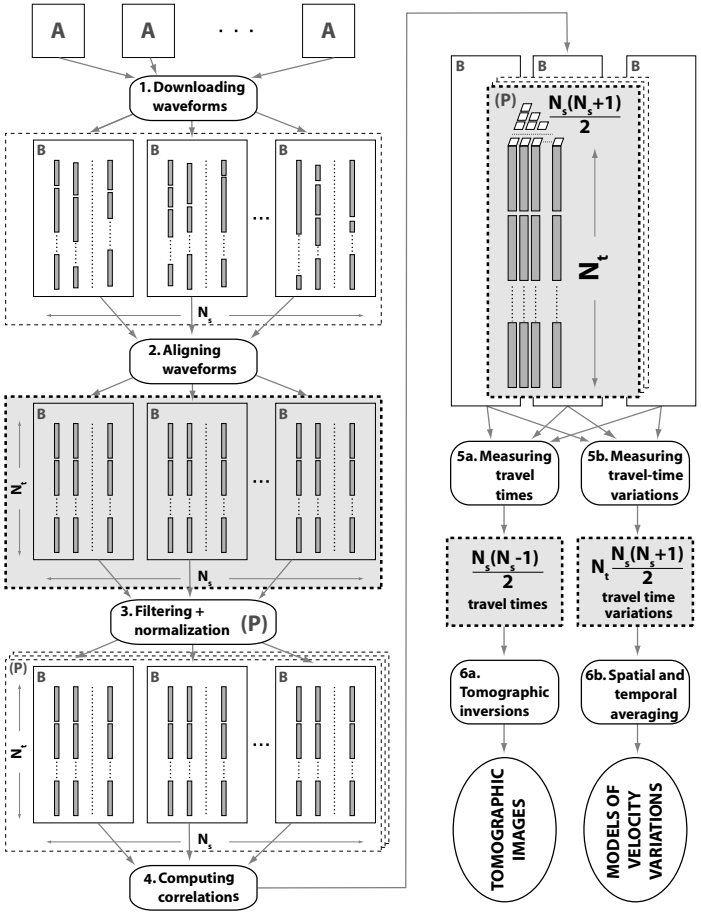




Brenguier *et al.* (2008)

Seismic noise correlation: Big Data





Data ingestion / quality control

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

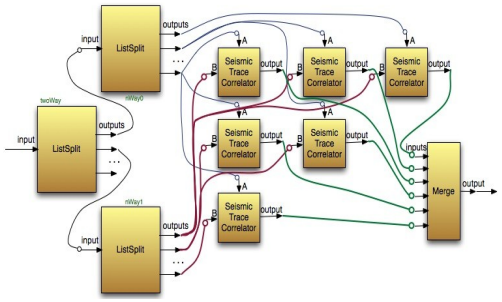
Data processing

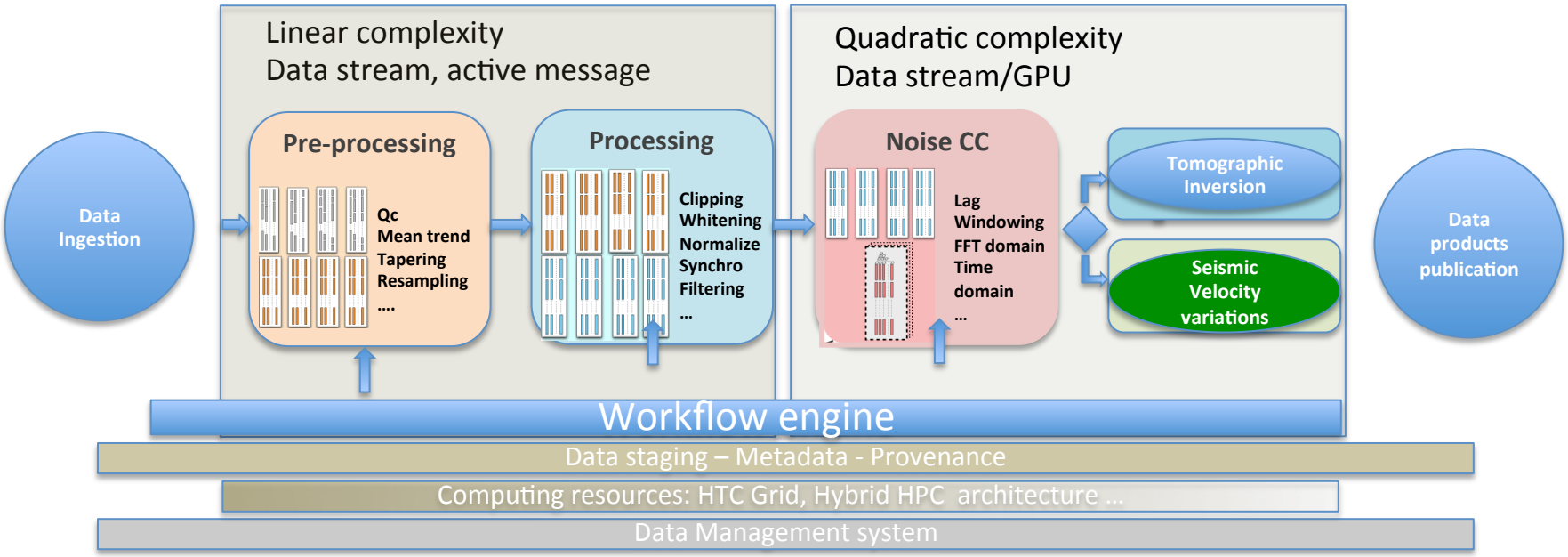
- **Low level data access pattern**
- **Linear complexity**
- Streaming data workflow
- Provenance and metadata management

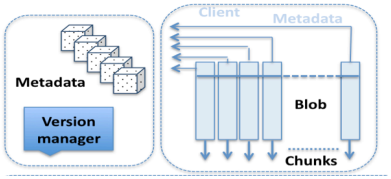
Data analysis

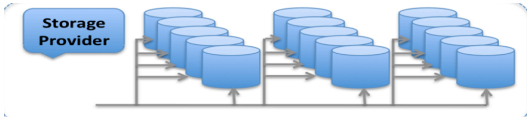
- **Cross-correlation** and higher order statistics
- **Quadratic complexity** and CPU intensive
- Thread-blocks CUDA and CSP
- **Secondary data : ~ 6 * N^2 * N_t**
- Provenance and metada management

Data-Intensive statistical analysis workflow







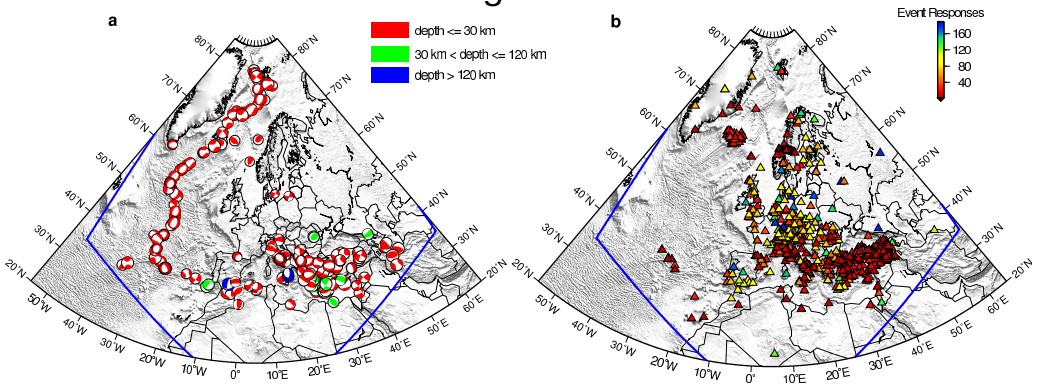


- Seismology PEs library and data streaming workflow (Dispel)
- Data management layer: PFS
- Data management layer integration with value added analytics: iRODS platform + MonetDB
- Data provenance layer integration

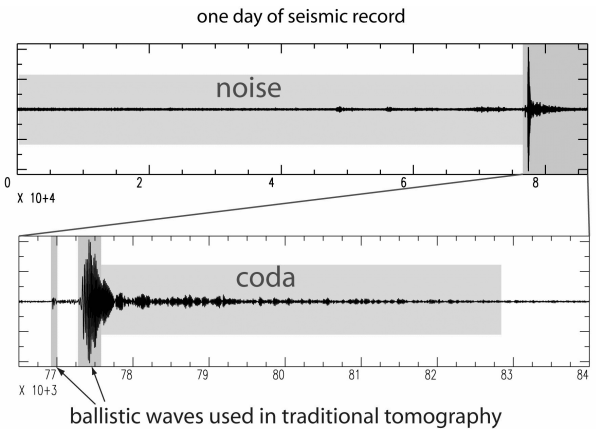
CPU-intensive modelling: seismic waveform inversion

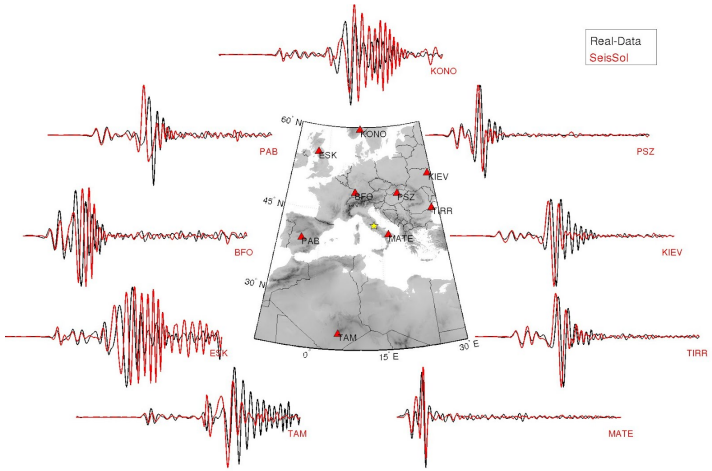
Adjoint Tomography of Europe

"Big Data"

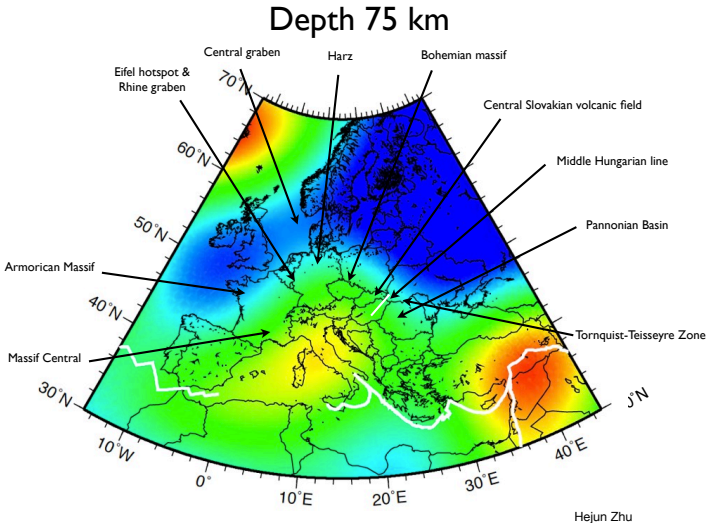


earthquakes	stations	iterations	simulations	CPU hours	measurements
190	745	30	17,100	2.3 million	123,205

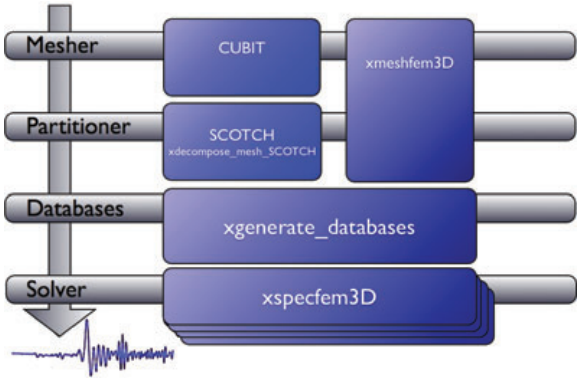


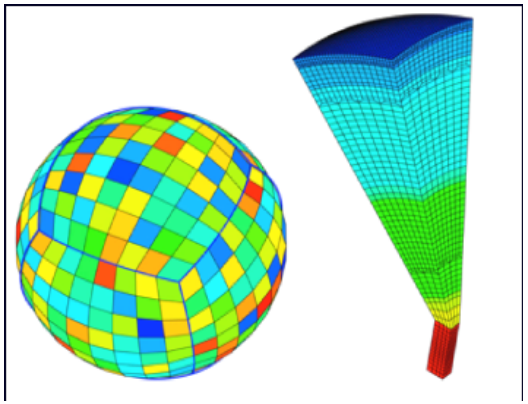


Krishnanet al (2012)

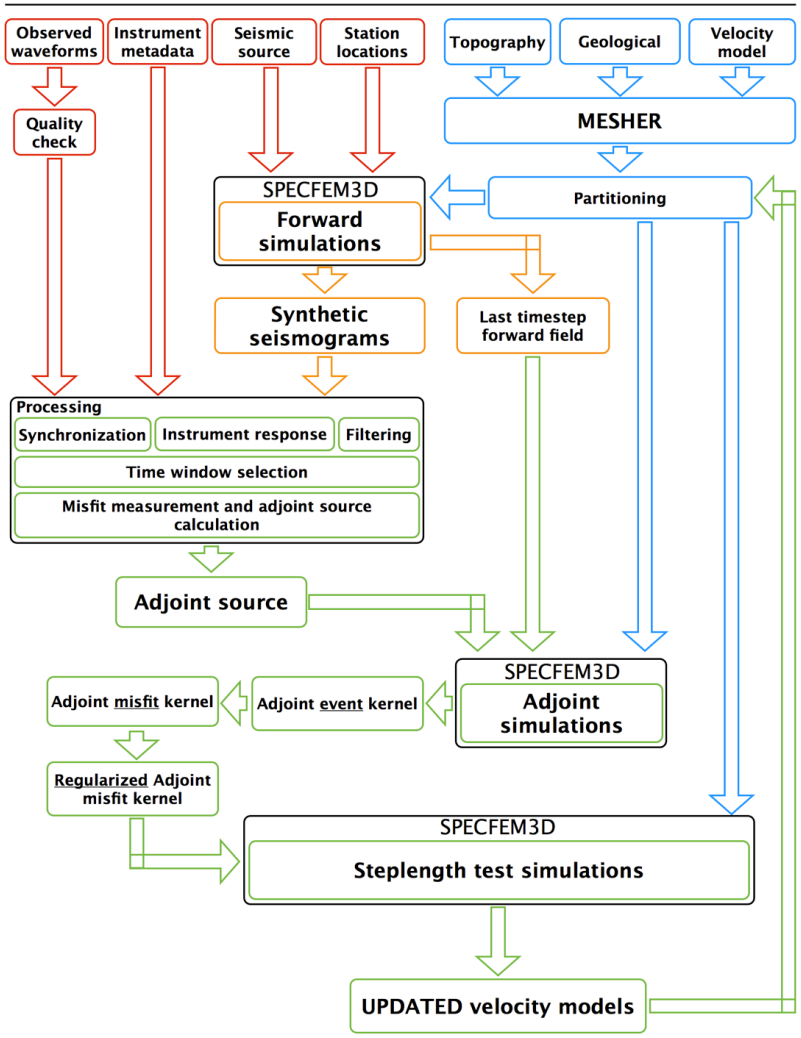


Tromp et al (2012)





CPU-intensive modelling: waveform inversion



High performance parallel codes

- Specfem3D, Seisol …

Waveform inversion

- Non-linear inversion
- Adjoint-based inversion methods: **->** one forward and one adjoint simulations per iteration and per earthquake

Orchestrated workflow

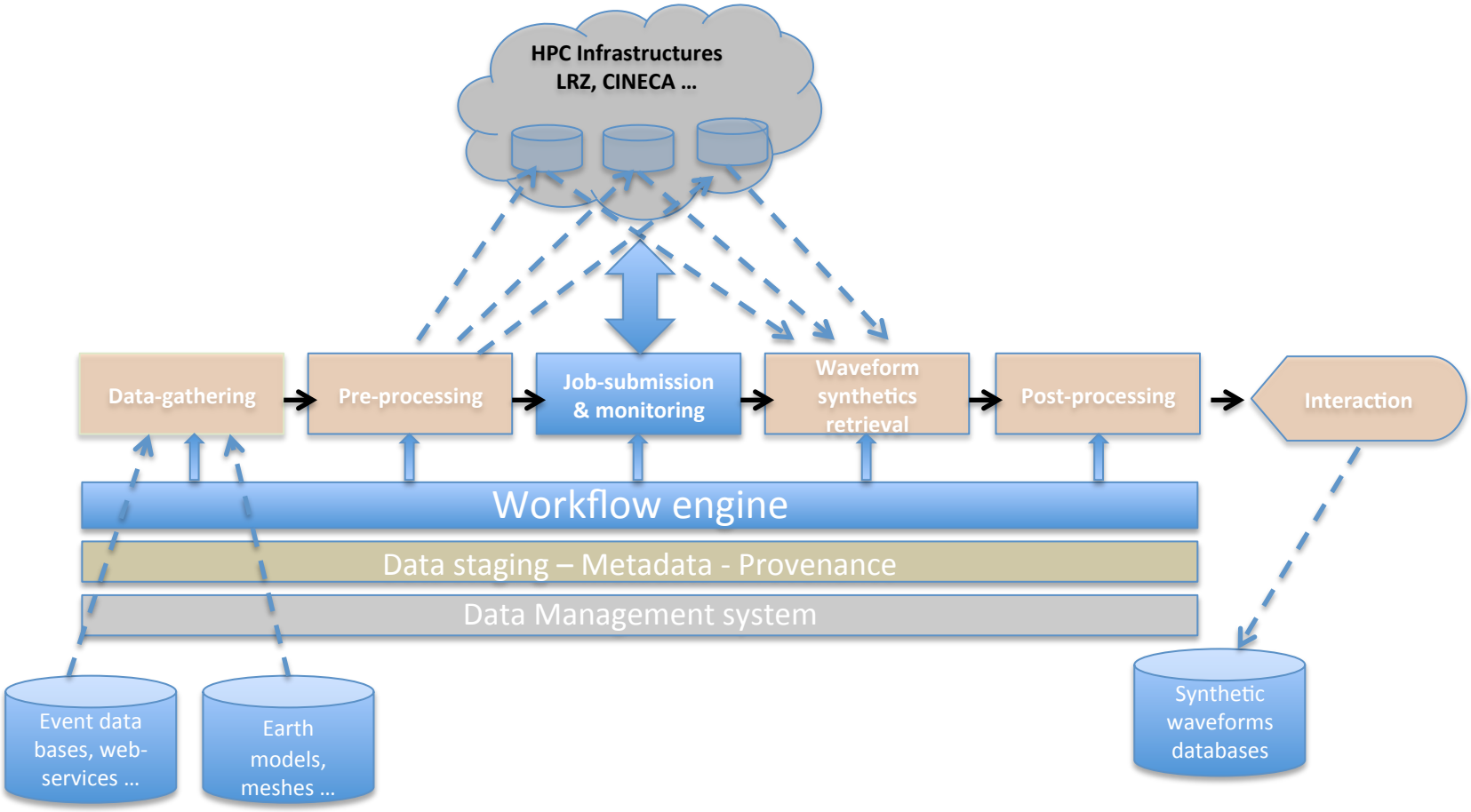
- Data Intensive analysis and High Performance computing
- Across Public HPC and Private data and computing infrastructures

Big Data

- Earthquake event waveforms: synthetics and observed
- State of the systems: x,y,z,t -> v, σ

Mesh generation

Data-intensive HPC workflow

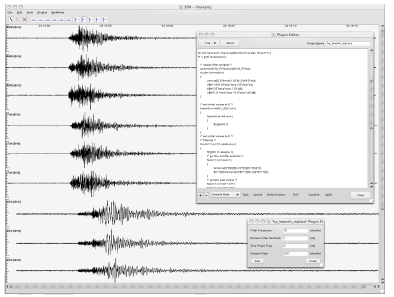


- Orchestrated workflows and execution models
- Stream based data analysis and enabled CSP wave simulation codes (Specfem3D and Seisol)
- Job submission across Grid & HPC DCIs: AAA (X.509 proxies), JSAGA/DCI-Bridge
- Data streaming and files transfer orchestration across DCIs:
- GridFTP enabled data transfer PEs, iRODS

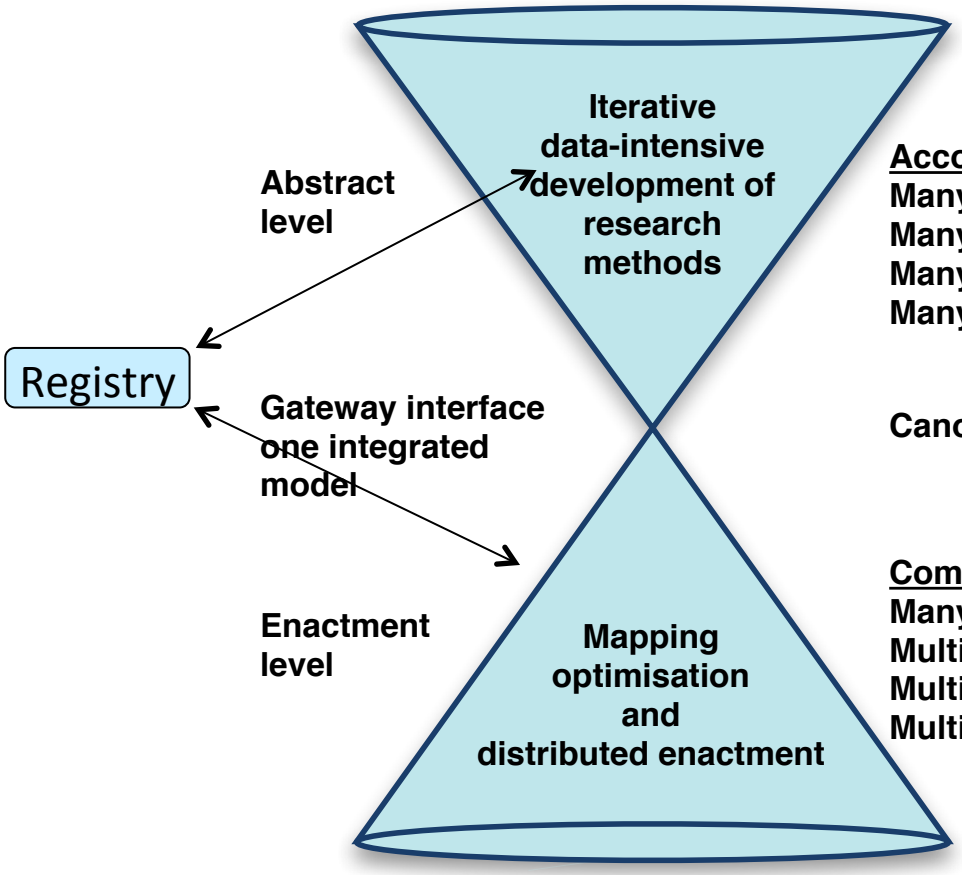
A service-oriented architecture

Seismology specialists → HPC and Data Analysis experts → Data-aware distributed computing engineers

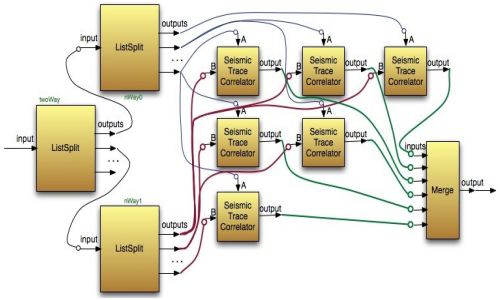
Separation of concerns



Workbenches for seismologists

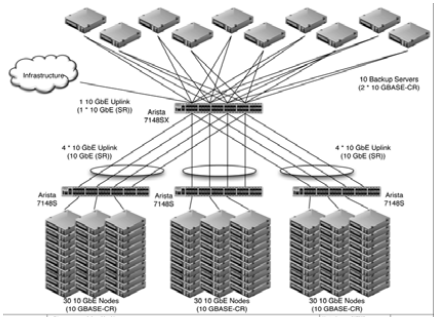


Accommodating
Many groups of researchers
Many tool sets
Many research strategies
Many working practices



Canonical representation

Composing or hiding
Many autonomous resources & services
Multiple enactment mechanisms
Multiple platform implementations
Multiple e-Infrastructures



System heterogeneity and complexity

Resilience toward "standards" evolution

Architecture

Architectural changes

- Tipping balance to data : data crawling architecture strategy;
- Support both Big Data DC architectures: data-intensive analysis – loosely coupled, data streaming on par with data throughput - and CPU-intensive architecture – tightly coupled;
- Compute in storage architecture and technology with added analytics;
- Augmented hierarchical object-based storage management, and heavy concurrent data access beyond POSIX;

What operational changes

- Supporting extended Data life-cycle within HPC infrastructures: data storage hierarchies and scientific gateways;
- Analytics platform must integrate Data-intensive HPC infrastructures and Data-intensive HTC infrastructures;
- Supporting orchestrated workflow – and data flow - across BD and EC DCIs and execution models: access policy, AAA mechanism, monitoring tools ….

Workflows

Forwarding looking workflow

- Seismic waveform inversion workflow is a possible proxy: extreme computing + data-intensive ;
- Seismic noise correlation analysis another possible proxy
- Need an abstract description level to identify communality with other domains: astronomy/astrophysics, sensor analysis …
- The devil is in the details: cleaning imply a knowledge of the acquisition/transmission
- Need incentive to involve of the domain communities and an analysis of the community data organization

Software missing

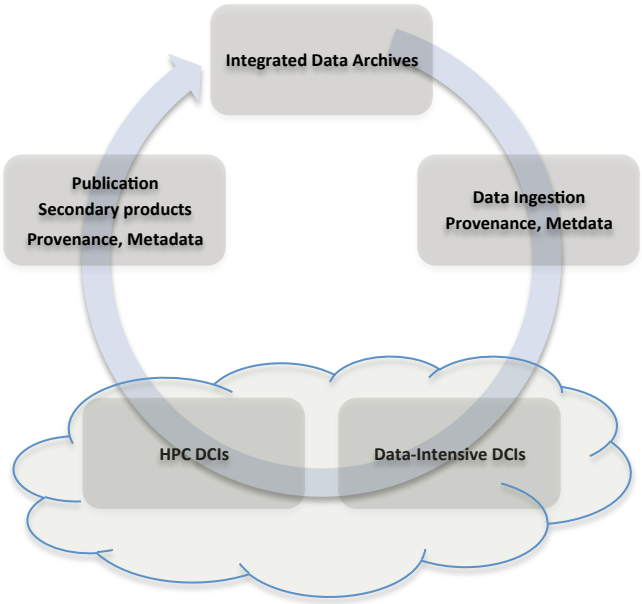
- Meta-workflows capability (ERFlow, SCI-BUS) and better enactment gateway support;
- Beside the software issue, it is a policy issue: policy-based IAA (credentials) and access protocols still desperately heterogeneous.
- Improved Scientific Gateway components: registries, Portlet and REST

Taxonomy

Big Data

Data Archives and Data infrastructure

Global observation systems: Integrated distributed data archives
Long term observatories: raw data preservation, data curation, data annotation
Data and Metadata standards
Data management and data exchange standards



Data-intensive research

- Increasingly large data sets (> 100-500 TBs each)
- Data-intensive: HPC modelling (inversion/assimilation); statistical analysis
- Different data life cycle:
  - Long-term (years) with shared services;
  - Mid-term (1-2 years), for research group analysis/modelling;
  - Short-term (few months) for massive processing (on demand ?) pipelines.
- Hierarchy of distributed storage -> vertical reuse optimization
- Orchestrated workflow across HPC infrastructures and Grid-like private/public infrastructures
- Secondary products publish in the Data archives with provenance and metadata

Software

Data management/exploration

- PFSs, iRODS, Scientific data bases (MonentDB)
- Data archives: Data and Metadata structure (<- acquisition/transmission & data exchange format)

Software library and tools

- Analysis domain specific libraries: ObsPy, Python, NumPy, SciPy, SeisHub, C/C++, Matlab
- Scalability and performance of cross-correlation and higher order statistics
- 3D wave simulation codes (Specfem3D and Seisol) continuous optimization. Good strong and weak scaling up to ~30-40 K cores.

Data management system needs

- Beyond Posix : n-dimensional objects, Blobs with dynamical adjustable chunk size, storage; concurrent access, versioning-based concurrent access
- Explore self-describing formats: HDFS, NetCDF, ADIOS

Software missing

- Fault tolerance: workflow & HPC codes

Interoperabilité

Data provenance

- PIDs and Handle System (EUDAT); see also on going discussion EarthCube/EPOS/EUDAT
- Semantics: OpenGeospatialConsortium (OGC) …
- Being implemented in coordination between VERCE, EPOS and EUDAT

VERCE

EPOS
EUROPEANPLATEOBSERVINGSYSTEM

EUDAT
European Data Infrastructure