Data and Data-intensive computing challenges in Earth and Universe Sciences

Jean-Pierre Vilotte
Institut des Sciences de l’Univers (CNRS-INSU)
Institut de Physique du Globe de Paris (IPGP)
Drive Scientific discoveries

- Observational data & simulation data
- High-end computational simulation
- Data inversion and Data assimilation
- Statistical data analytics

Across multiple disciplines

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

Socio-economical applications

- Climate evolution and forecasting
- Natural hazards (earthquakes, volcanoes, tsunamis, landslides, floods …)
- New energetic resources
- Environmental changes
An increasing wealth of data

**Ubiquitous data explosion: 100 PBs era**

- **Managing data**: streaming data processing, archiving, curation, metadata, provenance, distribution
- **Data analytics**: statistical streaming data analysis, machine learning methods of high-dimension data
- **Data-intensive simulation**: scalable, resilient large-scale, multi-physics, multi-scales simulations
- **Data-driven inversion and assimilation**: high-dimensional “Bayesian” inference methods
- **Statistics and stochastic methods**: direct-inverse uncertainties, extreme events statistics

**Next generation discoveries:**

- **Managing data**: streaming data processing, archiving, curation, metadata, provenance, distribution
- **Data analytics**: statistical streaming data analysis, machine learning methods of high-dimension data
- **Data-intensive simulation**: scalable, resilient large-scale, multi-physics, multi-scales simulations
- **Data-driven inversion and assimilation**: high-dimensional “Bayesian” inference methods
- **Statistics and stochastic methods**: direct-inverse uncertainties, extreme events statistics
Not a single dimensional challenge

Data Generation

Data-intensive Challenges

Data processing, transformation

Data analytics, Mining, unsupervised learning

Discovery, Insights, Prediction

Data management

data reduction query

data visualisation

data and method sharing

Adapted from Choudhary

Data -> Extraction/cleaning -> Integration/aggregation -> Learning models -> trigger / question -> predict
Big Data statistical analysis
Seismology: data-intensive analysis

Ocean seismic sources

Earthquake-induced property changes

Brenguier et al.

Landes et al.
Waveform analysis: data stream workflow

Data ingestion / quality control
- N-dimensional time (frequency) series
- Binary large objects: > ~100 TBs
- fine granularity (GBs)
- Partitioning, indexing, replication

Data processing
- Low level data access pattern
- Linear complexity
- fine-grained streaming data workflow
- Provenance and metadata management

Correlators (time-frequency)
- Cross-correlation and higher order statistics
- Quadratic complexity
- Thread-blocks GPUs / MIC
- Secondary data: ~ 6 * N^2 * N_t (N_f)
- Provenance and metadata management

Imagers (space-time-frequency)
- Convolution / projection
- High-order complexity
- Gridding
- Clustering - classification - machine learning
- Provenance and metadata management

From Shapiro, Vilotte et al.
LOFAR Epoch of Reionisation processing (> 100 MHz)

All-sky monitoring: detecting transient events and phase response-time

PB scale

Extraction of EoR signal:

Foregrounds
Jelic et al. 2008
Harker et al. 2010
Chapman et al. 2012

EoR

Noise
Challenges

What we value and experience

Researchers are part of the archiving process. They know what is relevant to understand their results.
Automated system should provide support for a consistent and effective acquisition of provenance metadata - Selective and extensible Provenance.
[A. Misra] [I. Foster]

Data stream processing engines

Data Intensive computation, present expensive requirements for provenance collection, either in terms of size or I/O [W. D. Pauw]
Enable active researchers to invent, refine scalable, statistical data-intensive methods

Support diversity of methods and implementation in a single data-intensive framework with data-handling services

Researchers remain in full control in their familiar community tools and libraries

Collaborative developments: from theoretical research to proof of concept to sustained use

Python library used to describe abstract workflows for distributed data-intensive applications.

Support for composition: Processing Elements defined with their own internal workflows.

Abstract streaming data flows: can be map and automatically executed in a variety of parallel environments.

Fine-grained provenance system: analyse and understand data relationship with triggered actions

Deployed on local Clouds (MAP-REDUCE streaming model)
Data-driven computing applications
Data inversion and assimilation
Data-driven applications: inversion and assimilation

**Exploration and marine geophysics**

- Seismology
  - Full Waveform Inversion
  - Extended Earthquake source

- Geomagnetism
  - Inversion of secular variation
  - Variational data assimilation

- Gravimetry
  - Inversion of gravity field
  - Geoid and Earth shape

**Global scale tomography**

- Toward Bayesian-inference reconstruction

**Geomagnetic secular variation**

**Earthquake source imaging**
Orchestrated workflow: data-intensive & HPC

Full Waveform Inversion (FWI)
- non-linear Bayesian inversion
- adjoint-based inversion

High-performance parallel codes
- forward and adjoint wave simulations
- billion of cores

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

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

Adapted from Tromp, Komatitsch et al.
FWI compute and data analysis

**Convergence**
- of data with computation

**Federating**
- autonomous diverse resources

**Handling**
- independent data sources

**Fluent**
- path from development to production

**Hiding** complexity

- **Federation of independent autonomous organisations**
  - data and computing infrastructures providers.

- **Services/access policies**: data-transfer, job control, task-oriented workflows

- **Transient storage** for users’ work in progress and intermediate data.

- **Shared persistent and caching storage**: optimise costs of data movement, assembly, processing, distilling and simulations over multiple investigations
Challenges

**What we want:** flexibility and reactive systems and users

- **Provides Run-Time feedback** on the process with **tuneable metadata** and controlled data movements
- **Avoids** useless waits for long and unfruitful runs
- **Fosters Dynamic Steering, Diagnostics**, saving computing cycles, storage ($$) and energy!
Turning large simulations into numerical laboratories
Mantle convection is recognized now to be the driving mechanism of plate tectonics. It governs the Earth's thermal and chemical evolution and involves both thermal and compositional transport. Mantle-convection processes act over multiple scales because of the many nonlinearities in the system, such as from rheology and sharp compositional gradients.

Some grand challenges in mantle convection of immediate relevance are (1) three-dimensional convection with increasingly high Rayleigh number (Vincent and Yuen, 2000; Yuen et al., 1999; Yuen et al., 2000), (2) thermal-chemical convection with multi-components (Gerya and Yuen, 2003; Gerya et al., 2004), and (3) fault-zone dynamics along plate margins with realistic rheologies (Regenauer-Lieb and Yuen, 2003).

Figure 8 shows the temperature and streamlines of high Rayleigh number (Ra = 10^8) convection carried out with a grid resolution of 400 x 400 x 400 grid points. Going up to a high-enough Rayleigh number to observe a transition in convection style is a grand challenge. We need to go to a Rayleigh number between 10^10 and 10^12 to detect such a transition, as has been found in high-resolution, 2-D simulations. Such a high Rayleigh number requires a grid of at least 2500 x 2500 x 2500 grid points. Attacking this problem requires the combined expertise of geophysics, information technology, and fluid dynamics. Visualization and advanced data analysis is crucial for interpreting the results.

Thermal-chemical plumes from subducting slabs or from the core-mantle boundary must be treated as multi-component systems. One efficient method for handling this situation is to employ tracers for describing the evolution of the many chemical constituents being carried by the convective velocity field. Up to one billion tracers have been employed to study the dynamics of thermal-chemical plumes at subducting slabs in two dimensions down to resolution of a football field (Figure 9); at least 100...
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

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

An end-to-end workflow!
Numerical laboratory: Shared Data Analysis

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

In-situ data reduction

On-the-Fly
- MPI-based power spectrum
- MPI-based parallel Halos finder
- Halos properties

Shared data analysis

Services on top of the data
- Higher-order statistics for matter field and Halos
- Topological analysis
- Dynamical analysis
- Visualisation
... 

Data life-cycle: persistent storage, provenance, publication
Climate and weather modelling

A continuum of time and space scales

From days to months, years, decades, and millennia

From local to regional, continental and global

Detection, attribution and prediction of extreme events and modes of climate variability
The volume of worldwide climate data is expanding, creating challenges for both physical archiving and sharing, as well as for ease of access and finding what is needed particularly if you are not a climate scientist.

Overpack et al., Science (2011)
The IPSL-CM application

Large number of models with a number of configurations a number of experiences an ensemble of realisations

Large number of variables, and files

Large volume of secondary data

~ 10 PBs scale
A resilient and flexible runtime environment

- Providing Run-Time feedback on the process with tuneable metadata and provenance-driven controlled data movement
- Avoiding useless waits for long and unfruitful run
- Fostering Dynamic Steering, Diagnostics, saving computing cycles, storage and energy ($$)

A flexible, resilient and reactive provenance-driven system
Numerical laboratory: Earth System Grid Federation

~ 10 PBs scale

Web processing service (WPS)

- Identity providers: OpenID, OAuth, LDAP
- Data sources: ESGF, Thredds
- Local replicas: CMIP5, obs4MIPS...

- ESGF search
- Access to local IPSL replicas

Malleefowl

- Workflow engine
- Esmvalwps

- Pyramid Web GUI
- PyCSW Catalog Service

Celery (scheduler)

PyWPS

- Birdy: execute

Climate Model Assessment Framework (CLiMAF)

- Exploration and analysis of climate simulations
- Share data processing and analytic methods and tools
- Advanced management of simulations and analysis
- Induction of a broad community of researchers and users
- Accelerate the full path of data use from capture to delivery of information
- Web services on top of data analysis platforms
- Pervasive provenance system

IPSL, courtesy of S. DENVIL
From HPC simulations to data-intensive platforms

Largest simulations at the petabytes scale
- From regional to global scales (climate, seismology, magnetohydrodynamics, etc.)
- From supernovae to turbulence
- Need for community access/reuse of the best and the latest secondary data through numerical laboratories with pervasive provenance system

Create new challenges:
- How to move/output data during simulations (vertical re-use, I/Os, parallel storage)
- How to reduce data through in-situ analytics
- How to stag in and stag out data (high-speed transfer protocol, access policies)
- How to explore/visualise data (render on top of the data, immersive analysis)
- How to analyse/instruments data (data analytics, immersive analysis, value added services …)

Research-driven
Huge variations in Data lifecycle and commitments
- On-the-fly (in-situ) analysis and visualisation (immediate, do not keep)
- Collaborative reuse and analysis secondary data (short/mid term, local)
- Community services and analytic tools (mid /long term, community commitment)
- Archival, curation, provenance, trust of secondary data (long term, community commitment)

Different from today supercomputer usage and access policies
  *A variety of data and computing resource access patterns*

! Dump data into large HPC providers, move data out to analysis platforms
Compute and Data-analysis federated infrastructures

a research-driven strategy
A research-driven variety of infrastructures

Ashby’s Law of Requisite Variety

Only variety absorbs variety
Data-intensive analysis platform and HPC

- Caches and persistent caching storage close to data-intensive analyse platform
- Data-intensive computing architectures and HPC simulation architectures
- Render on top of the data together with value added services, data analytics
- Induction to a broad research and user community (access and security)

Bridges and Gateways

**System-building**: technology-based and research-driven services

**Technology transfer across domains and locations**: variations of original design and emergence of competing systems

**Gateways consolidation**: research-driven technical solution with social choice integrated within research communities of practice federation of dissimilar autonomous systems into research-driven networks

(from Edwards et al., 2007)
E-Infrastructure challenges and strategy

System and infrastructure Big data analysis

Where should the caches and persistent storage be?

- Caches and persistent caching storage: sharing large chunk of observations and simulated data, optimise costs of moving data
- Not directly at the supercomputer (too expansive storage)
- Analysis computations and visualisations on top of the data
- High-speed transfer protocols from/to data sources (HPC, large instruments, data archives)

Complex data movements scheduling

- Data and metadata bases (scalability)
- Provenance-driven triggering and management
- Extended file management systems and model
- Augmented services with added-value to large community

Data organisation

- Most of these data are not hard to partition (scale-out)
- Provenance management system and lineage metadata
- Fine-grained data streaming flows
- Tier of large memory systems (random access)
Challenges and strategies

★ Difficulty getting things to run in multiple providers contexts
  • explore new virtualisation technology and sandboxed environment
    • Linux Containers, docker, Google Kubernetes
  • support software developments and maintenance
    • prepare the new generation of HPC architectures (exascale challenge, in-situ data analytics)

★ Difficulty to provide uniform access, trust and security model
  • leveraging existing identification systems across infrastructures

★ Difficulty handling data and computation strategy
  ✤ reduce computational costs
    • well-matched architectures to each stage
  ✤ reduce data movement costs
    • in-situ analytics, persistent storage, caching strategy, compression
  ✤ re-use of calculations and data
    • effective metadata and provenance system information

★ Align methods with research infrastructures
  • balanced and aligned investment for the full path of data use
    • maximise overall value of generating, collecting, preserving, curating data

★ HPC and Data Infrastructures tailored by scientific use cases
  • a variety of access and usage patterns requirements

★ Interdisciplinary task forces
  • share mutual understanding of methods and technologies (Astrophysics, Climate, …)
  • Interdisciplinary task forces
  • Computer scientists must meet flexible federation challenges
Data-intensive analysis platforms

**A scientific e-science environment** capable of “observing” (explore, analyse and model) massive and complex data generated by large-scale instruments, observation and monitoring systems, and numerical simulations in the sciences of Universe.

- Innovative methods, software, ICTs for large scale data-intensive computation and massive data statistical analysis that ultimately **induce a broad base of researchers to new research practices**
- Emergence of **cross-disciplinary expertise** in data-intensive computing and data analytics across scientific domains, research informatics, HPC and Data system engineers
- Accelerates full data use path: **valorisation of massive data generated by large-scale instruments, observation and monitoring systems**
- **Training and `intellectual ramps”** to engage a **new generation of researchers** to harvest data capabilities in their research practices to address new research challenges
- Community building around **simulation and data analytics shared application-software** together with **provenance and services** for open research and application science
- Consider **full path of data use and data life cycles** -> federation of HPC and data-intensive analysis platforms

A flexible and scalable federation of autonomous infrastructure providers/organisations

**Data resources - Data-intensive analysis platforms - HPC infrastructures**
Data-intensive e-Infrastructure Action Theme 3

- Identify and fund interdisciplinary use-cases for federated data- and e-infrastructures in environmental and global change challenges.
- Identify and fund large-scale Data and Model Inter-comparison Projects (DMIP) that are relevant to global change research.
- Through the above outcomes, inform data- and e-infrastructure policy with case-proven best practices that respond to concrete issues.

Milestones

29-31 August 2016, Paris, France: 2 scoping workshops in Paris: cross- and trans-disciplinary data-intensive use cases analysis; data and model inter-comparison (DMI) use cases

10-16 September 2016, Denver, Colorado: International Data Week: 2017 Belmont Call finalisation for data-intensive cross and trans disciplinary use cases and DMI projects

Contact: Mark Asch (ANR, France), mark.asch@agencerecherche.fr

http://www.bfe-inf.org