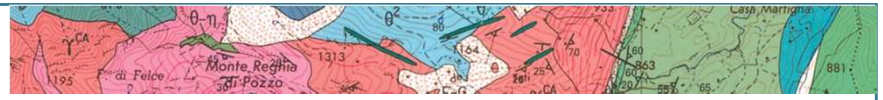


Modélisation du risque sismique sur architecture multicoeurs

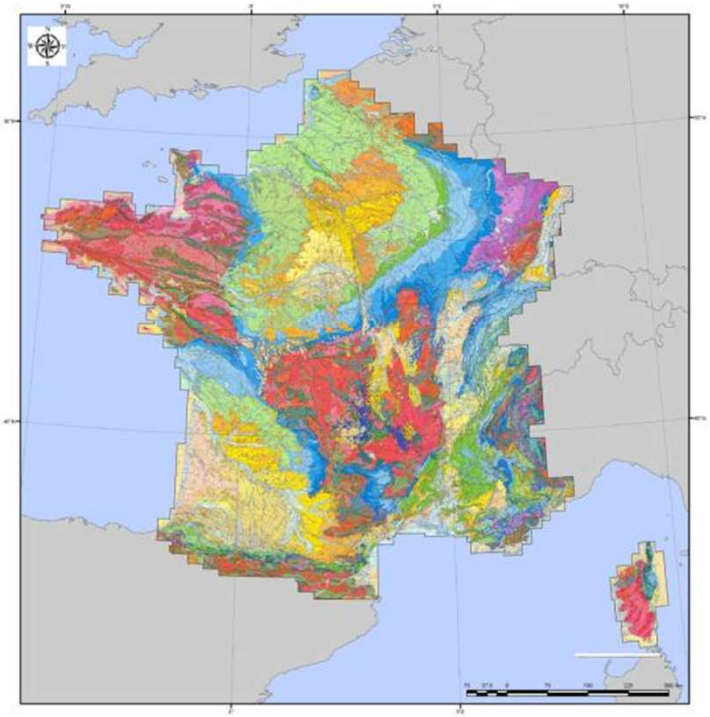
Fabrice Dupros

Collaborations with Inria and UFRGS (Porto Alegre)

IDEAS

Géosciences pour une Terre durable
brgm

French Geological survey

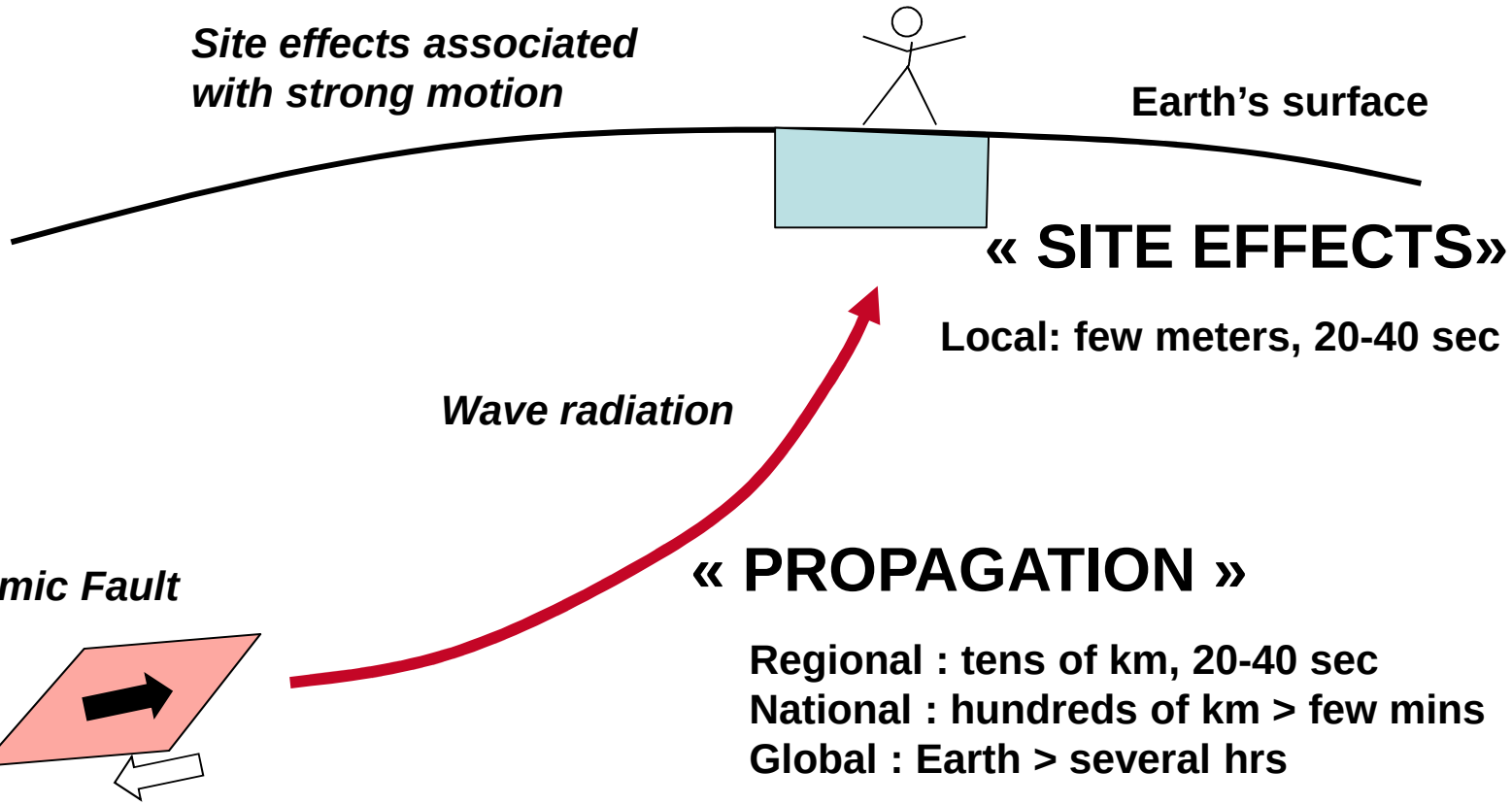


The « Referentiel Géologique de la France »

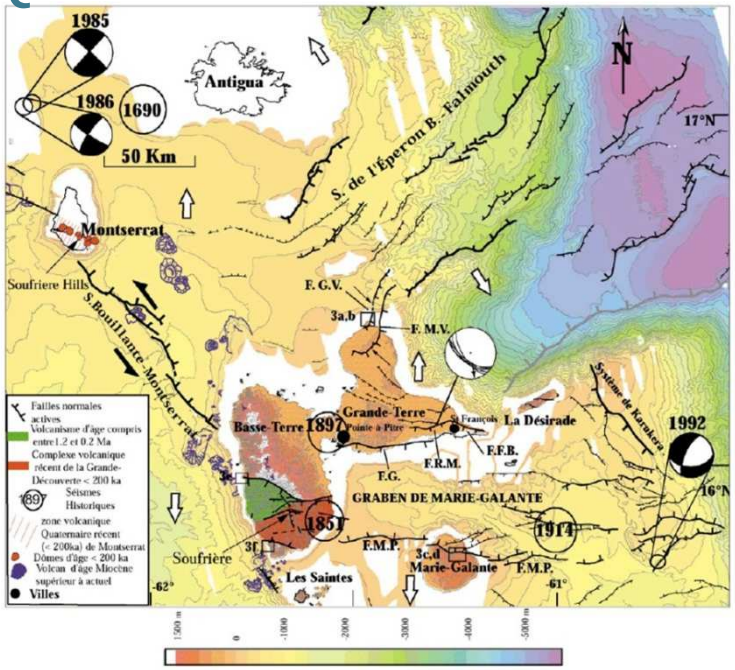
- Meet expectation of public authorities, engineering companies and scientific communiy
- Anticipate and answer to new societal needs (energy, natural hazards, ressources, ..)

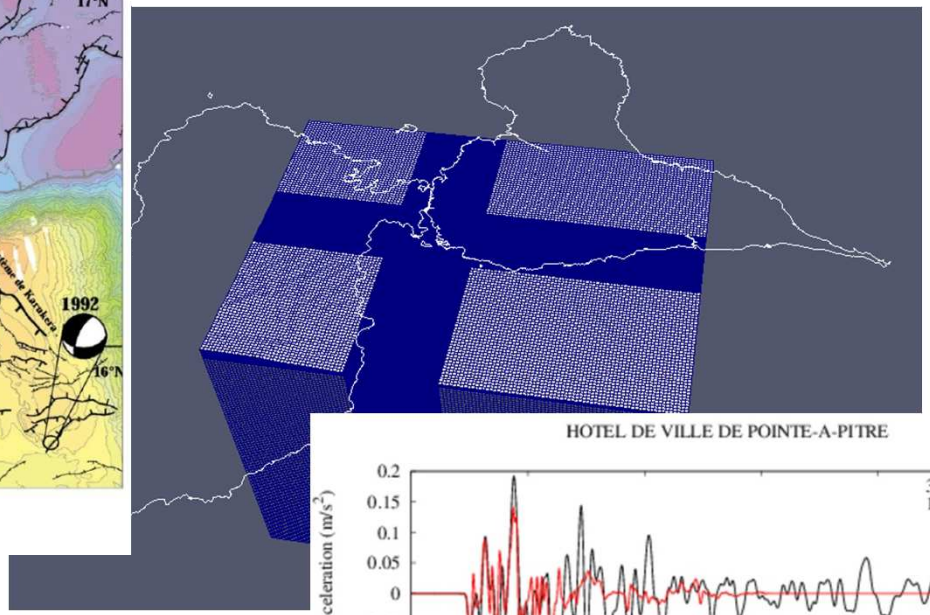
Natural hazards

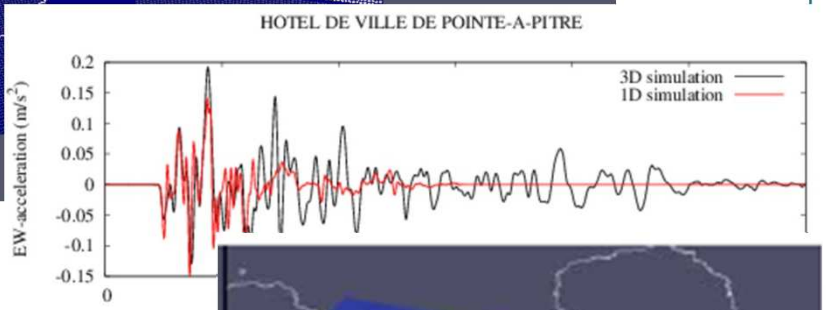
Earthquake risk assessment

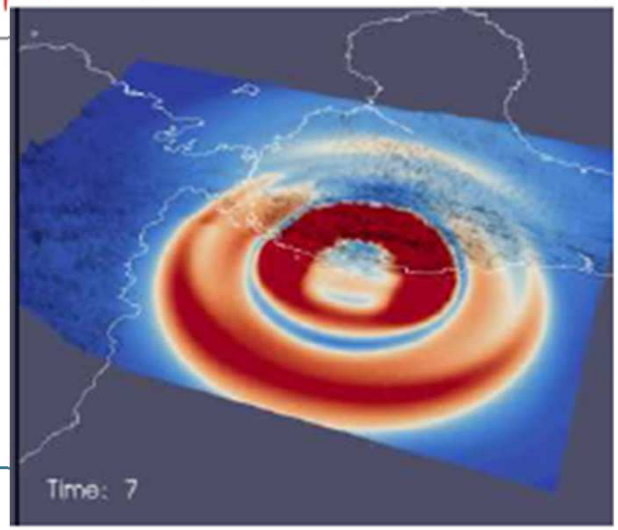


Quantitative seismic hazard assessment









REGION GUADELOUPE

> *Simulations in the FWI (Guadeloupe)*

> *Regional scale (tens of kms)*

⟶ **I/O → several tens of gigabytes**

Challenges

- 3D Full Wave Inversion
  - Seismic imaging
  - Risk associated with underground cavities
- Reliable 3D geological model
  - *RGF framework*
  - *Hetereogeneity and availability of data*
- Near real-time modeling
  - *Shakemaps after seismic events*
- Uncertainty
  - *Provide some robust criteria for risk assessment*

Challenges

> *Lack of quantitative seismological data based observations*

> *Strong impact on risk analysis*

> Virtual seismic world

- Generate a physically realistic earthquake catalogue
- Simulate eathquake ground motion
- Analysis of virtual data

> Some key fingers

- *Regional scale 100 x 100 x 30 Km / 100 000 earthquakes*

→**Computing resources > 6 millions of CPU.hours**

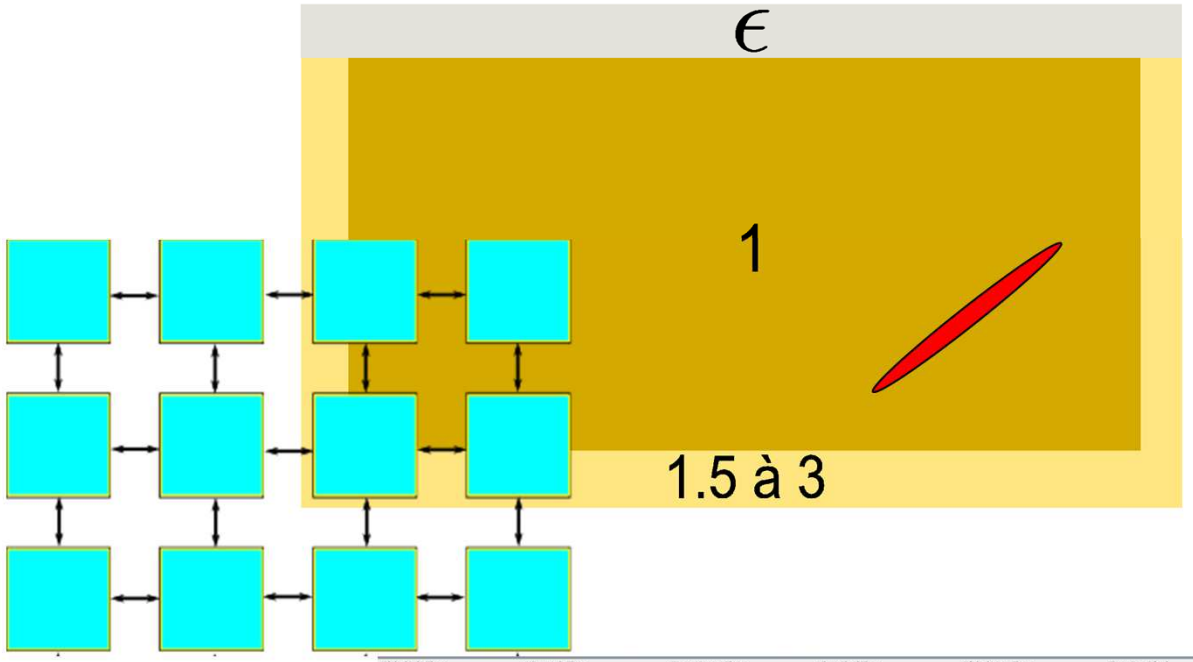
→**Storage resources > several tens of Terabytes**

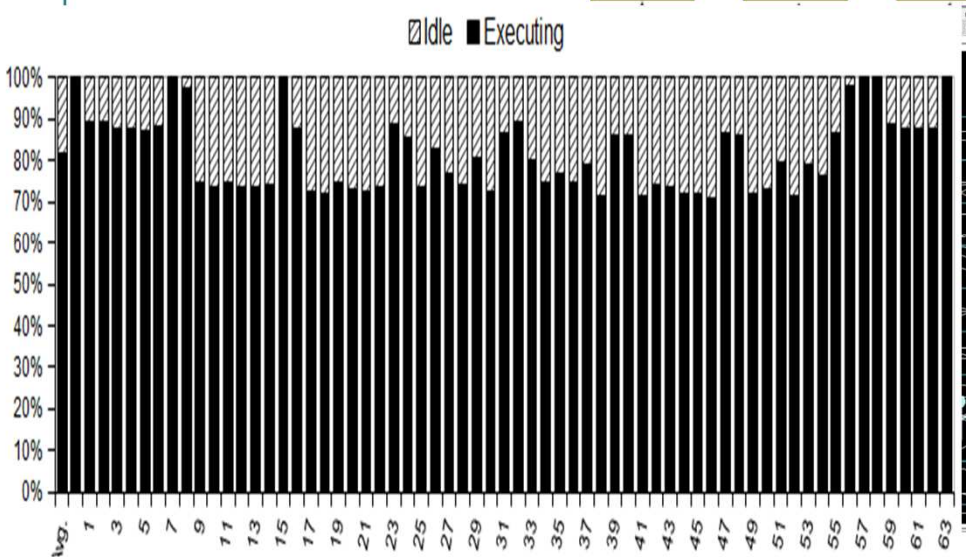
- *Availability /Efficiency of architectures / tools ?*

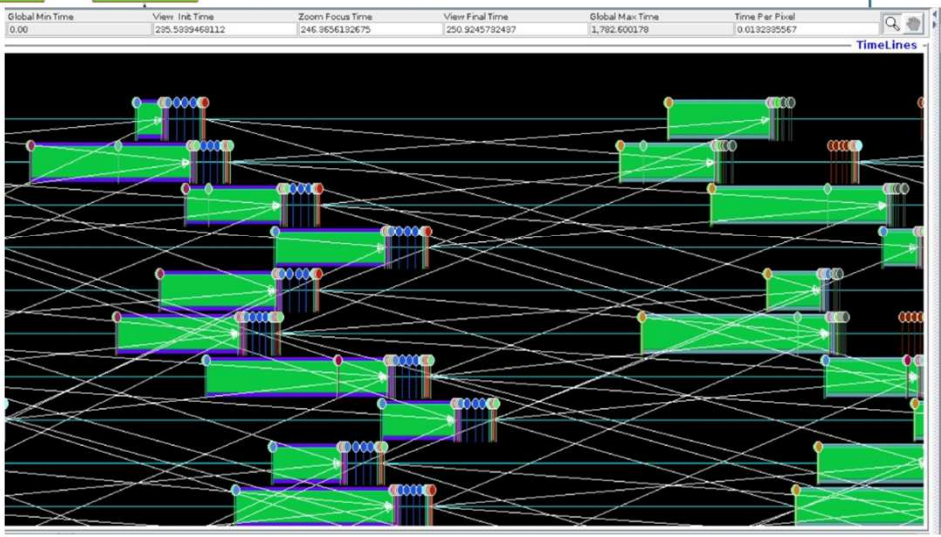
Data transfer

Load imbalance

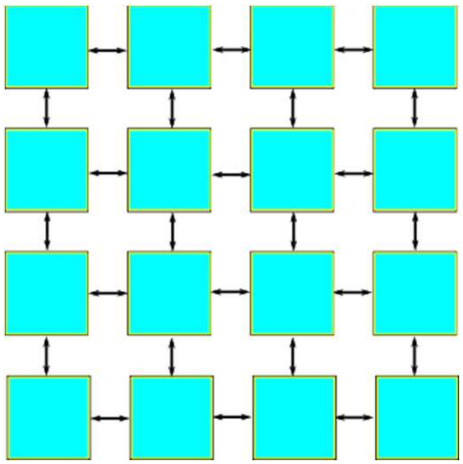
- From ABC
- From I/O
- From seismic sources





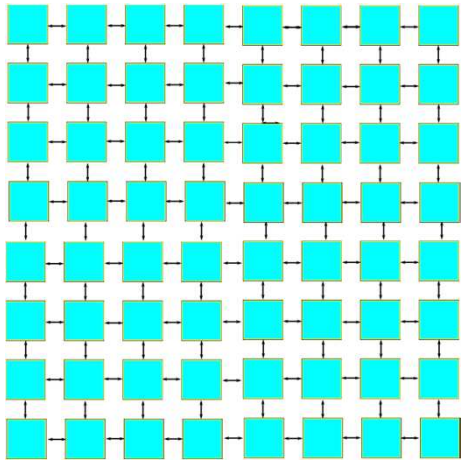


Data transfer



AMPI (Charm++) - *(Kale et al., 2001)*

- Over-decomposition
- Virtual processor
- Tailor load-balancer *(Pilla et al., 2012)*



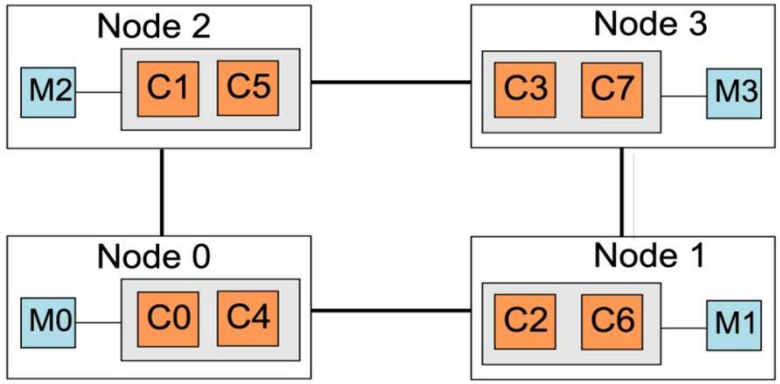
Data transfer

	0	1	2	3	4	5	6	7
0	0	8.97	8.20	18.67	1.14	8.74	7.43	18.73
1	8.01	0	18.87	8.24	9.26	0.04	18.39	7.92
2	7.62	18.96	0.62	8.43	9.07	18.91	0	7.88
3	19.12	9.10	9.56	0	20.88	9.06	8.62	0.38
4	0	9.14	8.21	18.85	1.37	9.04	7.55	18.53
5	9.09	0	20.12	9.45	10.47	1.60	19.61	8.89
6	7.65	19.33	0.17	8.64	8.43	19.35	0	8.46
7	19.67	9.78	9.34	0.75	20.70	9.50	9.13	0

Architecture with four NUMA nodes

> **Sequential**

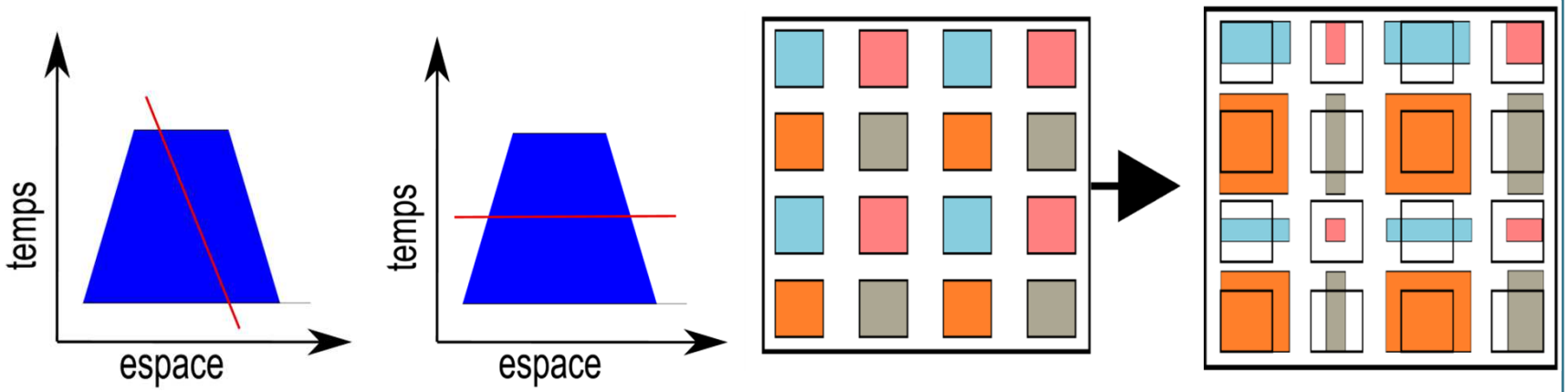
- Penalty ≈ 10% → one link
- Penalty ≈ 20% → two links



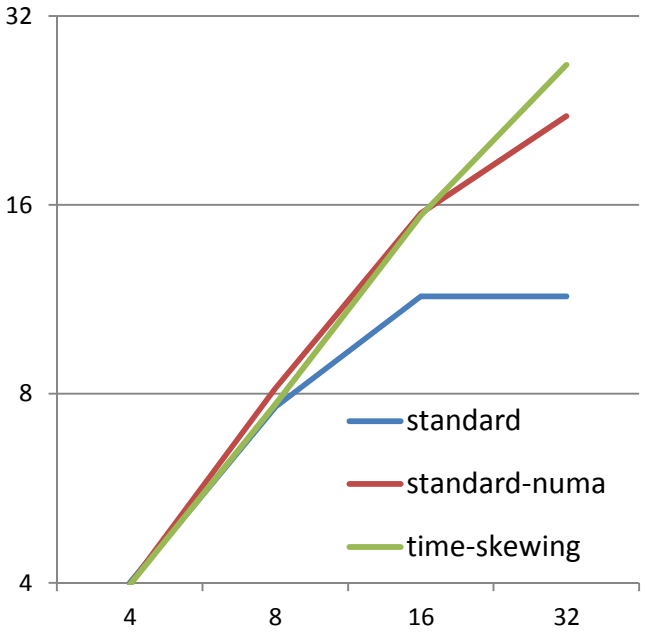
> **Sequential**

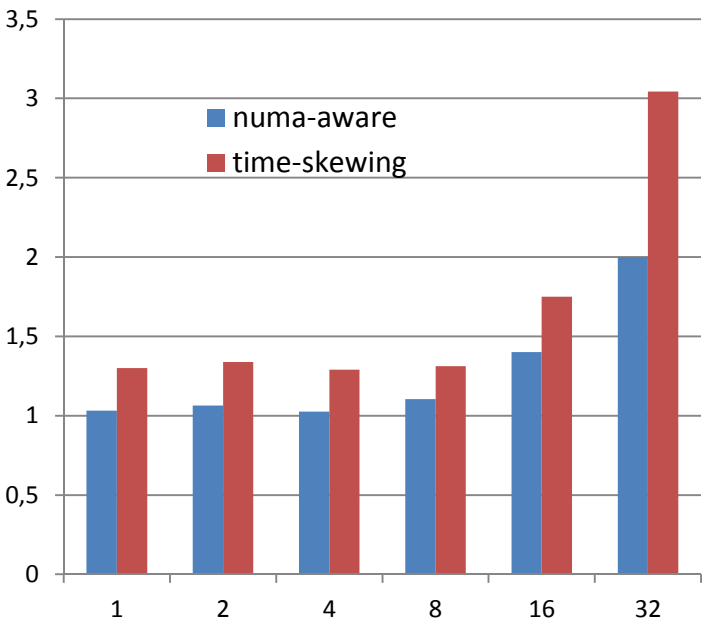
- from 54% to 22% depending on data size

Data transfer

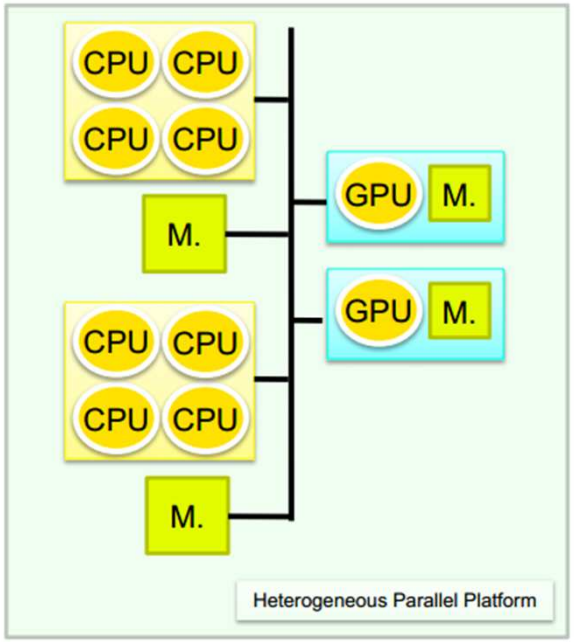


4-way Nehalem





Ongoing work – heterogeneous architecture

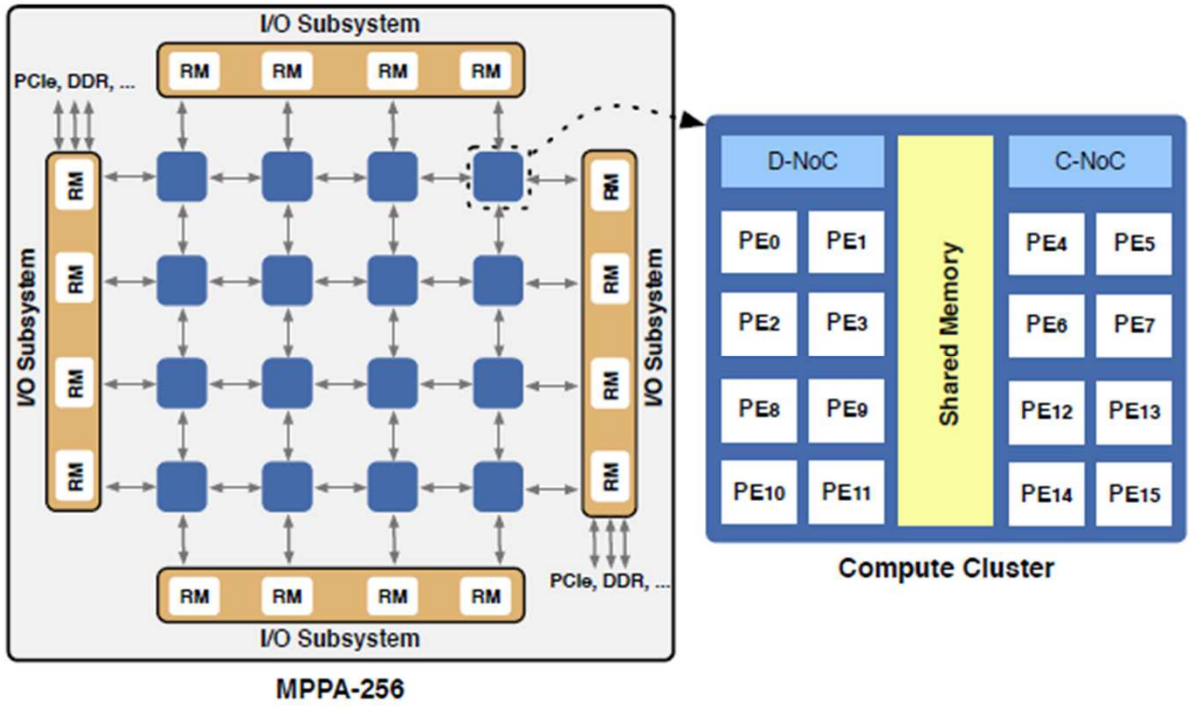


- Exploit existing CPU and GPU versions of the code
- Rely on top of **StarPU** runtime system (Augonnet 2011)
- Express task parallelism.

- **What could be the benefits ?**
  - as MPI + cuda exhibits good scaling (Michea et al. 2011)
  - as the performance is almost 40x between a GPU and x86 core for this kernel

- **Rational behind this experiment**
  - Node-to-node performance between GPU and CPU version is an average of ~4x (see also Komatitsch et al., 2012 for performance for Specfem3D)
  - Use the relevant architecture for a given task (I/O – ABC – etc..)
  - Run larger model by exploiting the memory available at the node level.
  - Energy consumption and portability

Ongoing work – Manycore embedded architecture



Kalray MPPA-256

- 2 MB/compute cluster
- 4 GB on I/O node
- 230 Gflops – 5 W

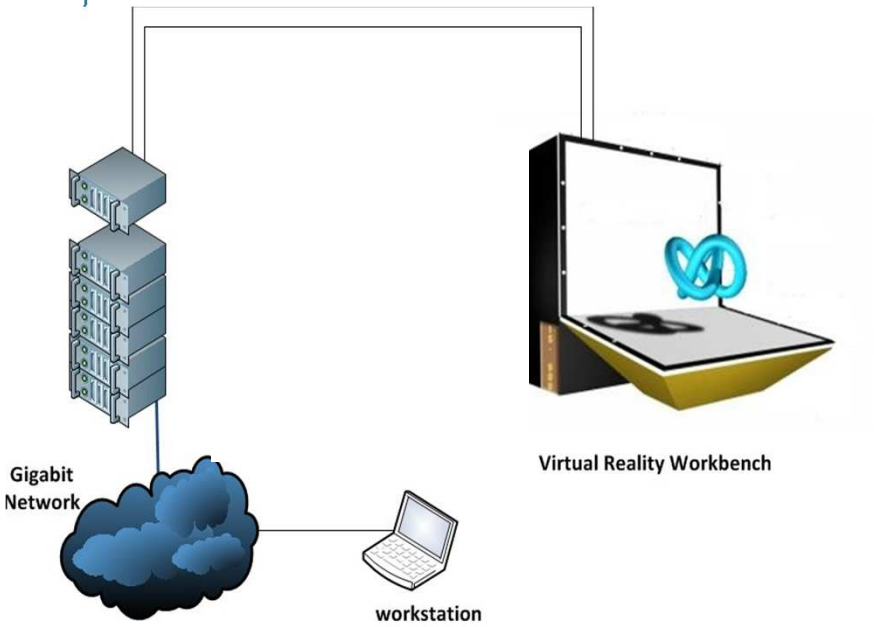
Travelling-salesman (vs Sandy-bridge)

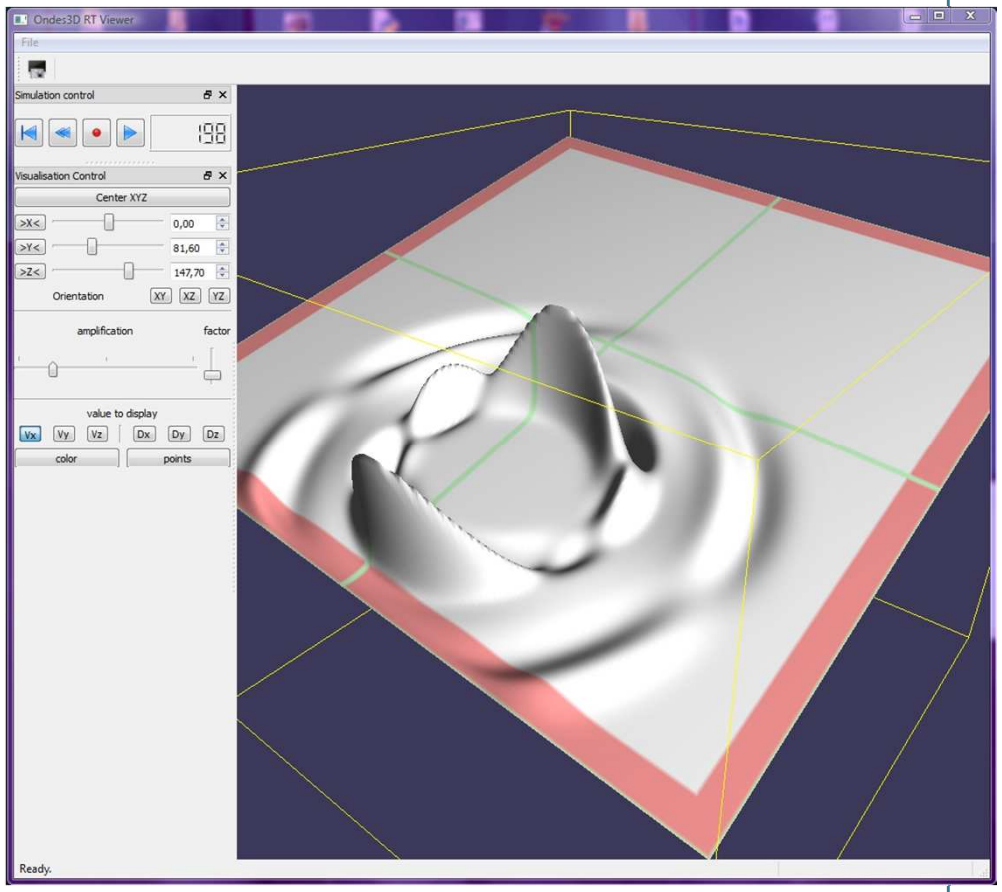
- Time to solution (1,6x)
- Energy to solution (~10x)
- *(Castro et al. 2013)*

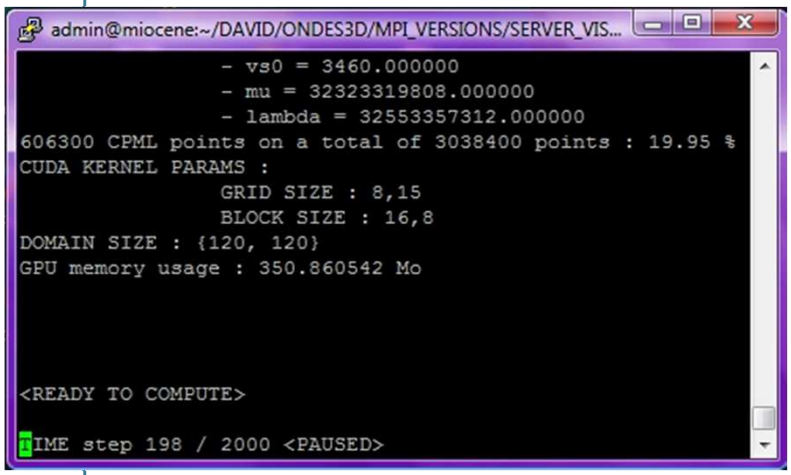
Seismic kernel

- Ongoing work
- Nearly perfect scalability on one compute cluster with OpenMP
- Optimized sliding-window algorithm to be implemented

Ongoing work – In-situ vizualisation







- Prototype with four MPI nodes *(Michea et al.,2012)*
- Efficient environment required for large model
- Very convenient for :
  - debugging
  - desktop computing/vizualisation

Conclusion

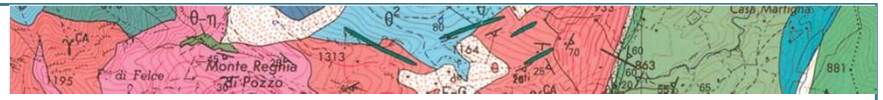
> Minimize data movement at various scales

- At Pre/Post processing level
- At the machine/node/multicore processor level

→ Key point for Exascale and energy consumption

> Need to :

- Rely on advanced runtime to handle efficient data migration
- Improve/Rewrite algorithm to tackle emerging platforms
- Provide a strong effort on post-processing tools/architecture



Thanks for your attention

Géosciences pour une Terre durable
brgm