Laboratoire d'Informatique de Grenoble

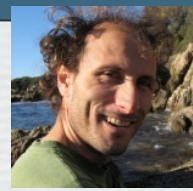
CORSE
Compiler Optimizations and Runtime SystEms

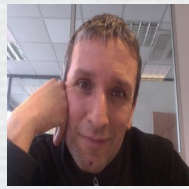
Corse Leader : Fabrice Rastello

https://team.inria.fr/corse/

6 Corse Members + 8 PhDs + 3 Post-Docs

- **Fabrice Rastello (CR1 INRIA)**
  - Loop transformations, register allocation, LLVM
- **Florent Bouchez Tichadou (MCF UJF)**
  - Static Single Assignment (SSA), compiler backend
- **François Broquedis (MCF Grenoble INP)**
  - Runtime Systems, OpenMP, Memory Management
- **Frédéric Desprez (DR1 INRIA)**
  - Parallel Algorithmic, Numerical Libraries
  -
- **Ylies Falcone (MCF UJF)**
  - Runtime Verfication, Enforcement, Monitoring
- **Jean-François Méhaut (PR UJF)**
  - Runtime Systems, Low Power Computing, Debugging



Corse Team Overview

- **Research Domain**
  - Compiler Optimization and Runtime Systems for Performance and Energy Efficiency
- **Issues**
  - **Scalability needs Heterogeneity/Complexity** : tradeoff between Specific Optimizations and Programmability/Portability
- **Target architectures**
  - Multi/Many cores, General Purpose Processors, Accelerators, GPU, Low Power Architectures
- **Applications**
  - Geophysics, Material Physics, Embedded Multimedia
- **Approach**
  - Combine Static/Dynamic & Compiler/Runtime

Corse Research Directions

- **Static/Dynamic Compiler Analysis**
  - Hybrid and extensible byte-code
  - Hybrid compilation, trace/static analysis
  - Instruction scheduling, computing and IO complexities
- **Runtime Systems and Interfaces with compilers**
  - Load balancing with compiler analysis information
  - Memory management (affinity)
  - Scheduling and load balancing directed by power management
- **Interactive debugging, properties monitoring**
  - Debugging with programming models
    - Gdb, OpenMP, OmpSs,
  - Dynamic monitoring of program properties

Compiler Optimizations : What and When ?

- **What work a compiler typically performs to optimize ?**
  - Analyze the source code (e.g. dependences, ranges, aliasing)
  - Emit the binary code from the Intermediate Representation (IR), Static Single Assignment (SSA)
  - Transform/Optimize the code (register allocation, loop unrolling)

- **When (analyze, emit, transform)**
  - Ahead-of-Time (static)
  - At install/load time
  - At runtime (with the runtime systems)

Compiler/static – Runtime/dynamic

- **Static/ahead-of-time**
  - Analysis: find (always) true facts (aliasing checking)
  - Optimizations: fine grain code transformation, regular code
- **Dynamic/runtime**
  - Runtime Systems: handle irregularities, data-dependent behaviors, execution environments, Load Balancing
  - Profiling/observation: find probable facts, perform statistics

The Runtime System may help the Compiler...

```
#pragma omp task
{
     for ( int  i = 0; i < n ; i++) d [I] = s [I] ;
}
```

Binary

```
#pragma omp task
{
 if (d+n <= s || s+n <= d  || d == s)
      par_for ( int i = 0; i < n ; i++) d [I] = s [I] ;
  else
      for ( int  i = 0; i < n ; i++) d [I] = s [I] ;
}
```

- **Code cloning**
  - Compiler: Aliasing Analysis for optimization
  - Generate optimized versions, one for each possible context, **binary code cloning**
  - Compiler has to generate **predicates**
  - Runtime system **checks the predicate** and **launches the task** with the adequate version

Compiler may also help the Runtime System...

```
/* b x b x b Blocked matrix-matrix multiply*/
for (I = 0, I < n; I += b)
 for (J = 0 ; J < n ; J+= b)
   for (K = 0 ; K < n ; K += b)
     C [I..I+b, J..J+b] += A [I..I+b, K..K+b] * B [K..K+b, J..J+b] ;
```

- **Static Complexity Analysis**
  - Arithmetic intensity
    - $F = 2n^3$
  - Memory access (cache size S)
    - $n^2 << S$: $Q = 3n^2$
    - $n \times b << S$: $Q = n^3/b$
    - $b \times b << S$: $Q = 2\ n^3/b$
    - $B << S$: $Q = n^3$
    - Otherwise : $Q = 2\ n^3$
  - F, Q metrics transmitted to the runtime system
  - Metrics help the load balancing decisions

Corse software

- **Compilers**
  - Plugins in existing compilers (mostly LLVM), new LLVM passes
  - Specification andTIREX toolbox (SSA Intermediate Representation)
- **Runtime Systems**
  - New load balancers (Charm++, OpenMP, Erlang,...)
  - Unconventional computing platforms : Kalray MPPA, FPGA, PIM (Processor in Memory)
- **Interactive debugging, properties monitoring**
  - Debugging with programming models
    - Gdb, OpenMP, OmpSs, Temanejo
  - Dynamic monitoring of program properties
  -
- **End-to-End Solutions for Applications**
  - BOAST : DSL for computing kernels (wavelets, stencils,...)
  - Geophysics (SPECFEM, Ondes3D), Material Physics (BigDFT)

Corse Collaborations with Brazil

- **UFRGS (P. Navaux), Porto Alegre**
  - LICIA (CNRS INRIA International Lab)
  - ExaSE (INRIA Associated Team)
  - Joint Laboratory on Exascale Computing (JLESC) : INRIA, BSC, UIUC, ANL, Juelich,
  - **HPC GA project** (FP7 Marie Curie IRSES)
  - HPC4E (H2020 EUB Project) : WP2 Disruptive Exascale Computer Architecture
- **UFMG (F. Pereira), Belo Horizonte**
  - Prospiel (INRIA Associated Team)
  - Diogo Sampaio's PhD Thesis
- **USP (A. Goldman), Sao Paulo**
  - Runtime Systems, Actor Programming
  - PhD students : E. Francesquini, R. Goncalves
- **PUC Minas (H. Freitas), Belo Horizonte**
- **UFSC (M. Castro, L. Pilla), Florianapolis**

Laboratoire d'Informatique de Grenoble

J-F. Méhaut
UJF-CEA-INRIA/LIG

High Performance Computing Geophysics Applications (HPC GA)

Seventh Framework Programme (FP7)
Marie Curie Actions, PEOPLE
International Research Staff Exchange Scheme (IRSES)

01/01/2012 12/31/2014

Agenda

- **Introduction**
  - FP7 PEOPLE IRSES programme
  - HPC GA project (PIRSES-GA-2011-295217)
- **Wave Propagation on manycore processors**
  - Wave propagation models
  - Kalray/MPPA architecture
  - Performance/Power Results
- **Conclusion**
  - HPC GA Results

FP7 PEOPLE IRSES Programme Objectives

- Attract/retain numerous, well-trained, motivated researchers in Europe
- Development of adequate and broad skills for both the **private and public sector**
- Strengthen international dimension
- No age limits for project participants (mobility):
  - Early stage researchers (**ESR**) : Post-graduate students <4Y
  - Experienced Researchers (**ER**) : Post-doc 4-10Y, senior post-doc >10Y
- Duration of exchange : **maximum of 12 months**
- EU contribution to mobility cost : **2100 Euros/month**

FP7 IRSES objectives

- To establish or deepen **partnerships** between research organisations within **Member States/Associated Countries** and organisations in countries covered by the **European Neighbourhood policy** and **Countries with S&T Agreements** through a joint programme of exchange of researchers

- **Countries with EU agreements on S&T**
  - Argentina, **Brazil**, Chile, **Mexico**...

HPC GA summary (01/2012 – 12/2014)

- **Challenges** of **designing high performance geophysics simulations** for **massively parallel architectures**
  - **Pluridisciplinary project**
    - Geophysics
    - Applied Mathematics
    - Computer Science
  - **2 EU Countries**
    - France and Spain
  - **2 Countries with S&T agreements**
    - Brazil and Mexico

HPC GA consortium

- **INRIA (France) : O. Aumage, J. Diaz, J-F. Méhaut**
  - Grenoble and Bordeaux : Computer Science, Runtime Systems
  - Pau : Applied Mathematics

- **UFRGS (Brazil, Porto Alegre) : P. Navaux, L. Farina**
  - Computer Science, Runtime Systems
  - Applied Mathematics, Numerical Methods

- **BCAM (Spain, Bilbao) : E. Zuazua, E. Akhmatskaya**
  - Applied Mathematics, Numerical Methods

- **UNAM (Mexico) : V. Atienza, J. Tago**
  - Geophysics

- **BRGM (France, Orléans) : F. Dupros, H. Aochi**
  - French Geoscience Institute

- **UJF-ISTerre (France, Grenoble) : E. Chaljub**
  - Grenoble Geoscience Lab

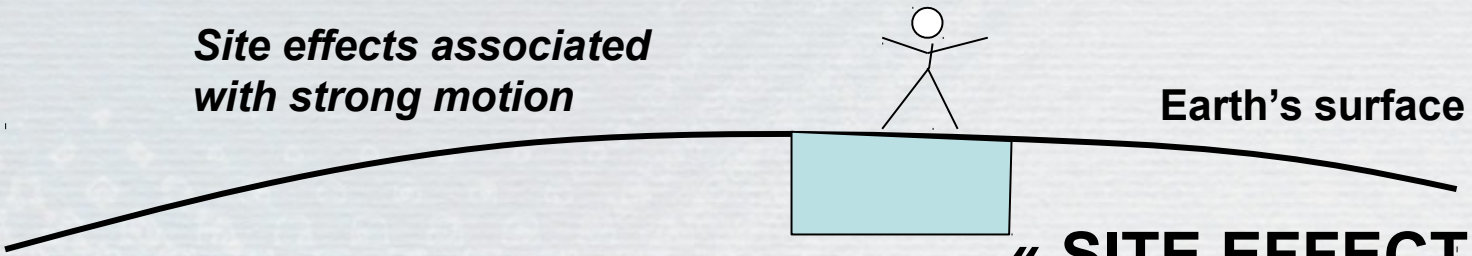
HPC GA workpackages

Workpackage	Title	Leader
WP0	Management	INRIA
WP1	Runtime for New HPC platforms	UFRGS
WP2	Numerical Models	BCAM
WP3	Adaptation of Seismic and Tsunami Applications for Manycore Architectures	BRGM
WP4	Wave Propagation in Complex Media	UNAM

HPC GA scientific workpackages

- **WP1 : Runtimes for new HPC platforms**
  - Participants : UFRGS, INRIA (Grenoble, Bordeaux)
  - Tasks : NUMA Hierachies, GPU programming, Low Power procesors, I/O and Data Management
- **WP2 : Numerical Methods**
  - Participants : BCAM, INRIA (Pau), UFRGS
  - Tasks : Elasto-acoustic coupling
- **WP3 : Geophysics Applications and Manycore Architectures**
  - Participants : BRGM, UJF-ISTerre, UNAM, INRIA
  - Tasks : Performance evaluation, code optimization
- **WP4 : Wave progagation in complex media**
  - Participants : UNAM, UJF-ISTerre, BRGM
  - Tasks : sedimentary basins, Site/city interaction (Mexico)

Seismic Wave Propagation



« SITE EFFECTS»

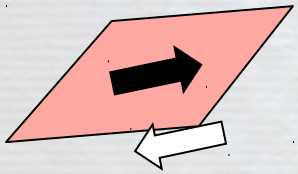
Local: few meters, 20-40 sec

Wave radiation

« PROPAGATION »

Regional : tens of km, 20-40 sec
National : hundreds of km > few mins
Global : Earth > several hrs

Seismic Fault



« SOURCE »

Magnitude 6 : average of 10 km < 10 seconds

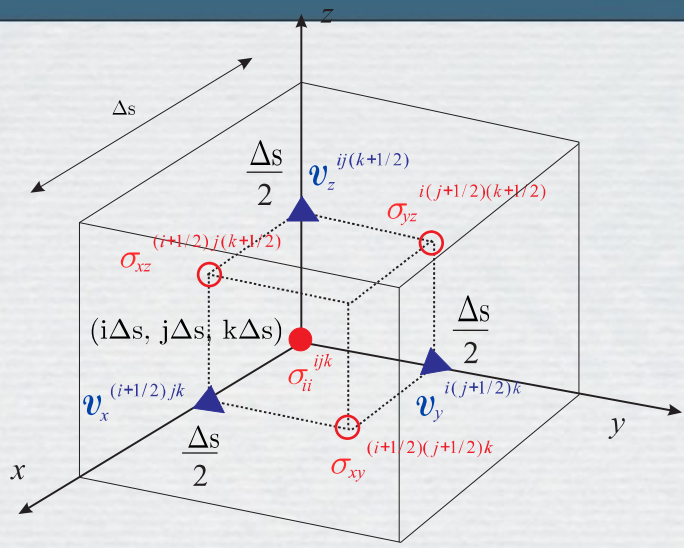
Seismic Wave Propagation Models

- Used to **predict the consequences** of future **earthquakes**

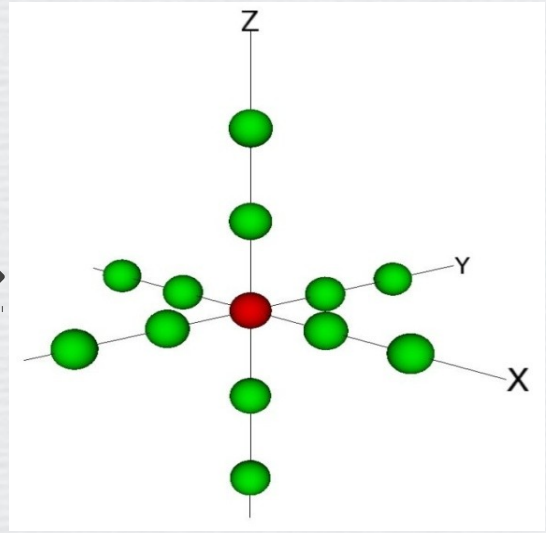
- **Seismic waves** are represented by a set of **elastodynamics equations**;
  - Solved by implementing the **explicit finite difference method**;

Seismic Wave Propagation (Ondes3D, BRGM)

- Simulation composed by **time steps**
- **In each time step (3D simulation)**
  - The first triple nested loop computes the **velocity** components
  - The second loop reuses velocity result of the previous time step to update the **stress** field




```
for x ← 1 to x_dimension
  do { for y ← 1 to y_dimension
       do { for z ← 1 to z_dimension
            do { COMPUTE_VELOCITY() }
for x ← 1 to x_dimension
  do { for y ← 1 to y_dimension
       do { for z ← 1 to z_dimension
            do { COMPUTE_STRESS() }
```



4th-order stencil

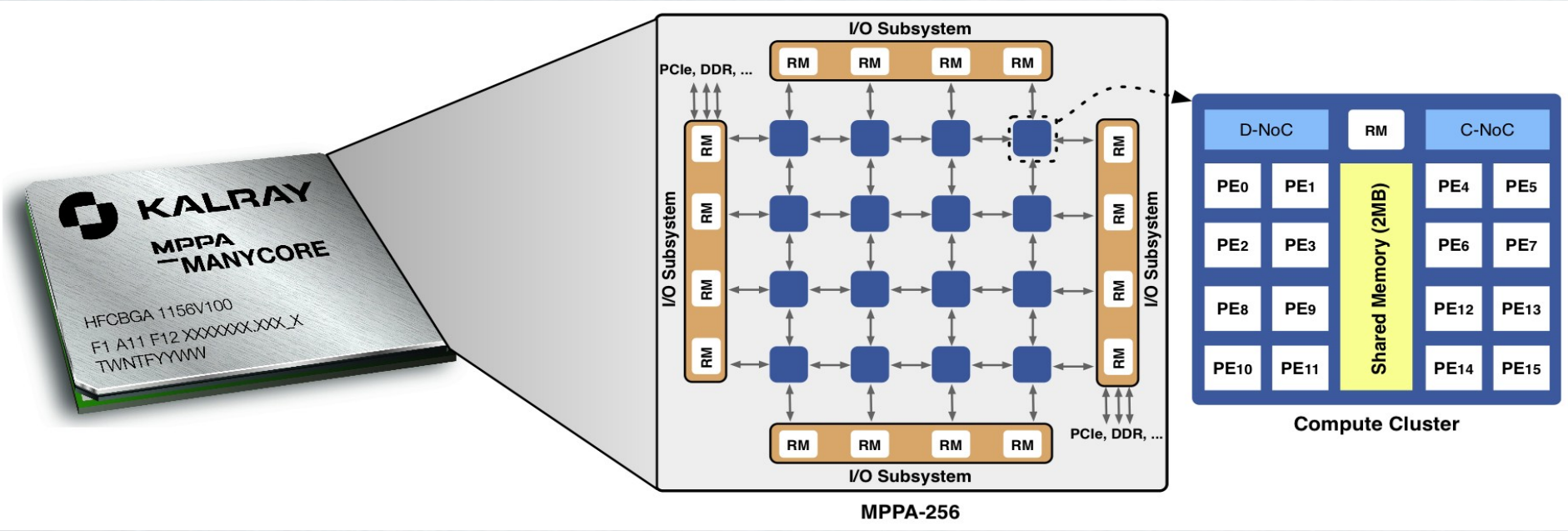
Manycore architectures

- **Industrial Collaboration with a SME**

  - **Kalray (http://www.kalray.eu)**
    - French fabless semiconductor and software compagny founded in 2008
    - France (Grenoble, Paris), USA (California), Japan (Tokyo)
    - Compagny developing and selling a new generation of manycore processors

  - **MPPA-256**
    - Multi-Purpose Processor Array (MPPA)
    - Manycore processor : 256 cores on a single chip
    - Low Power Consumption [5W - 11W]

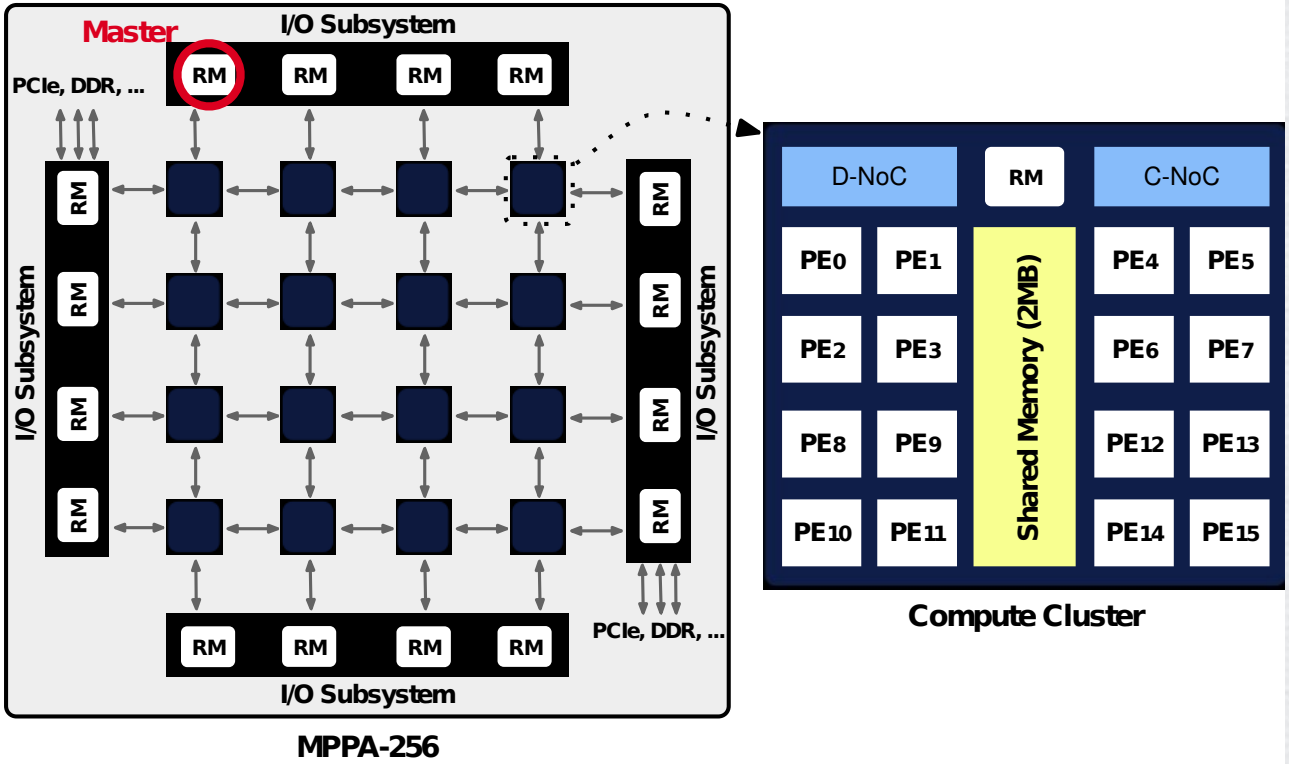
Kalray MPPA-256 architecture



- **256 cores (PEs) @ 400 MHz** : 16 clusters, 16 PEs per cluster
- PEs **share 2MB of memory**
- **Absence of cache coherence** protocol inside the cluster
- **Network-on-Chip (NoC)** : communication between clusters
- **4 I/O subsystems** : 2 connected to external memory

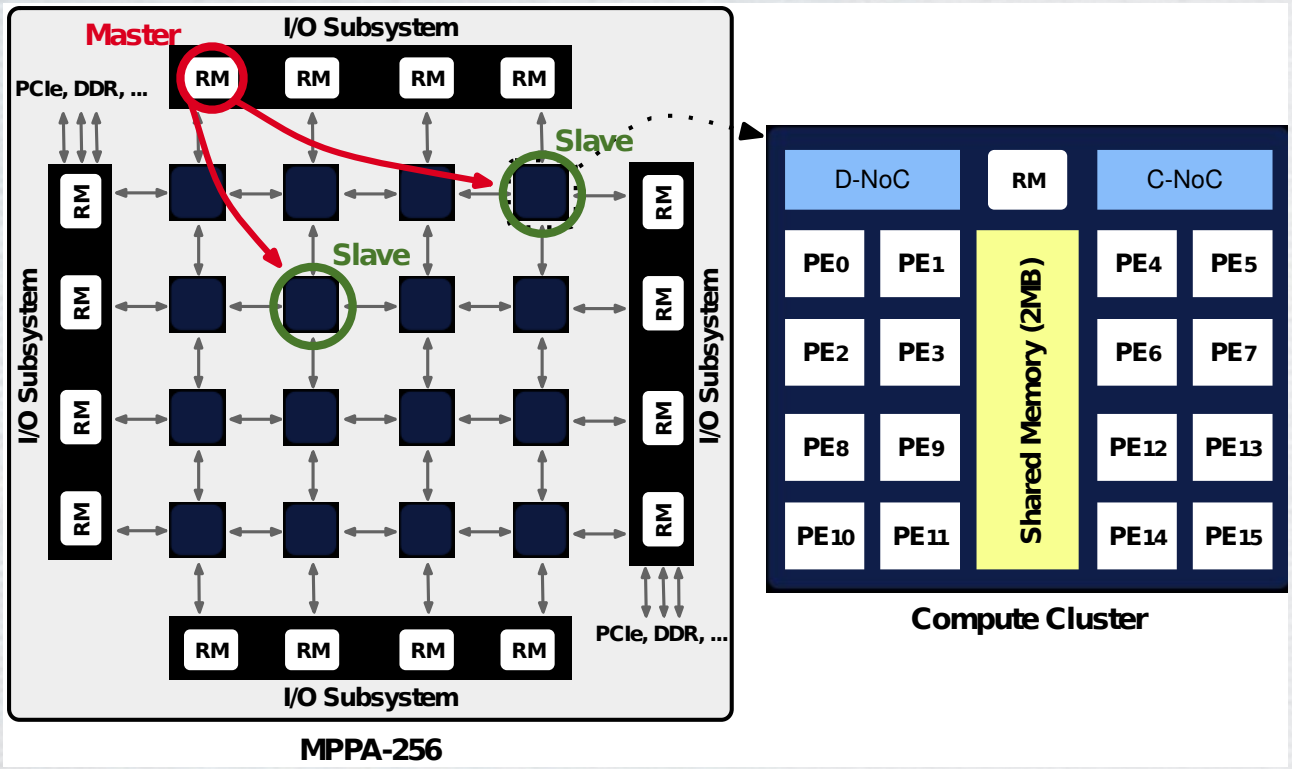
MPPA MANYCORE

Parallel Execution on MPPA



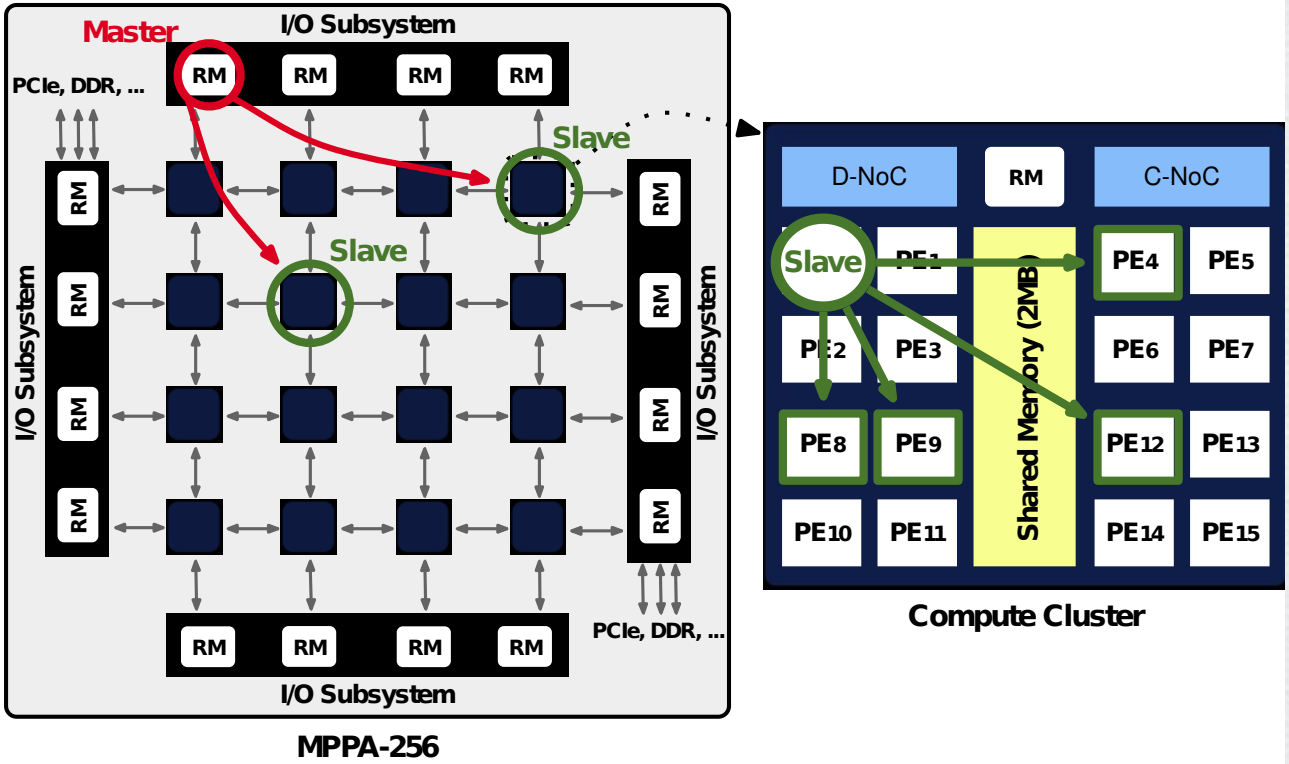
- A **master process** runs on an **RM** of one of the **I/O subsystems**

Parallel Execution on MPPA



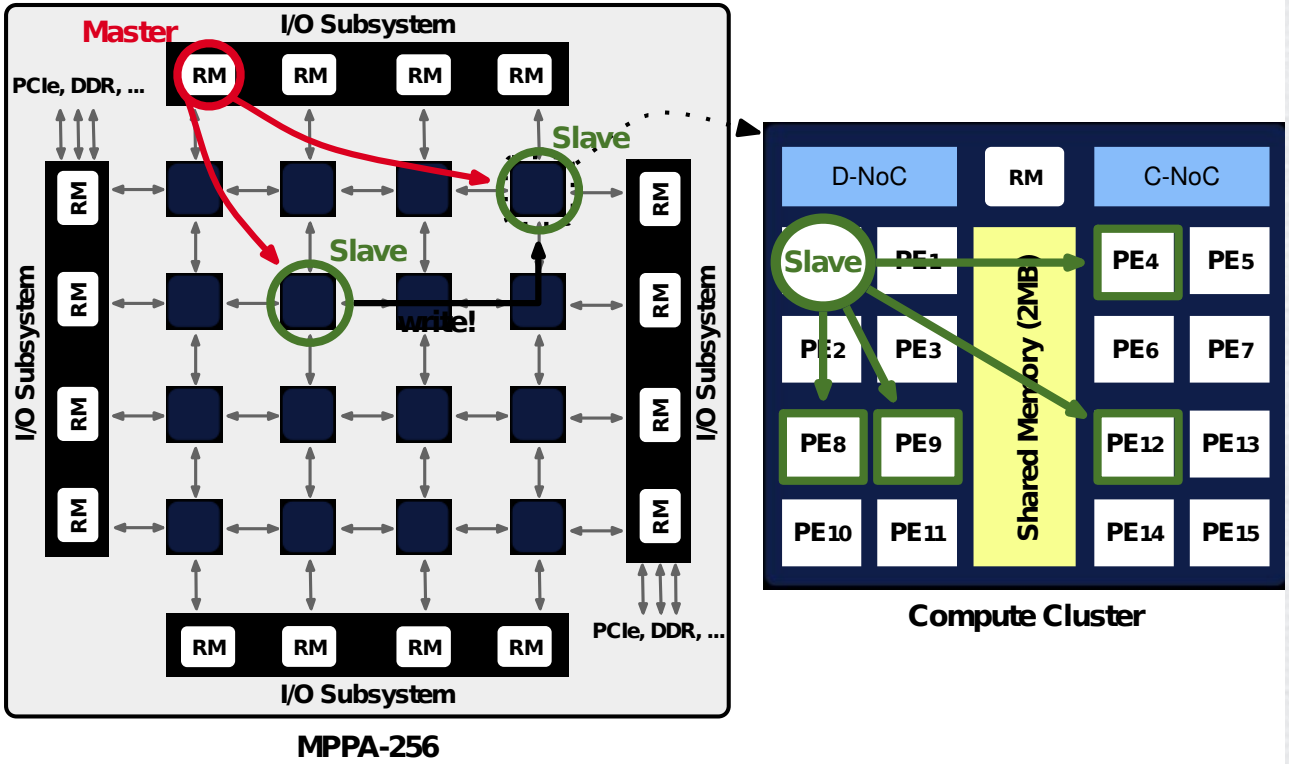
- The **master** process spawns **slave processes**
- One slave process per cluster

Parallel Execution on MPPA



- The **slave process** run on the **PE0** and may create up to **15 threads**, one for each PE
- Threads share 2 MB of memory

Parallel Execution on MPPA



- Communications take the form of **remote writes**
- Data travel through the **NoC**

Challenging issues with MPPA

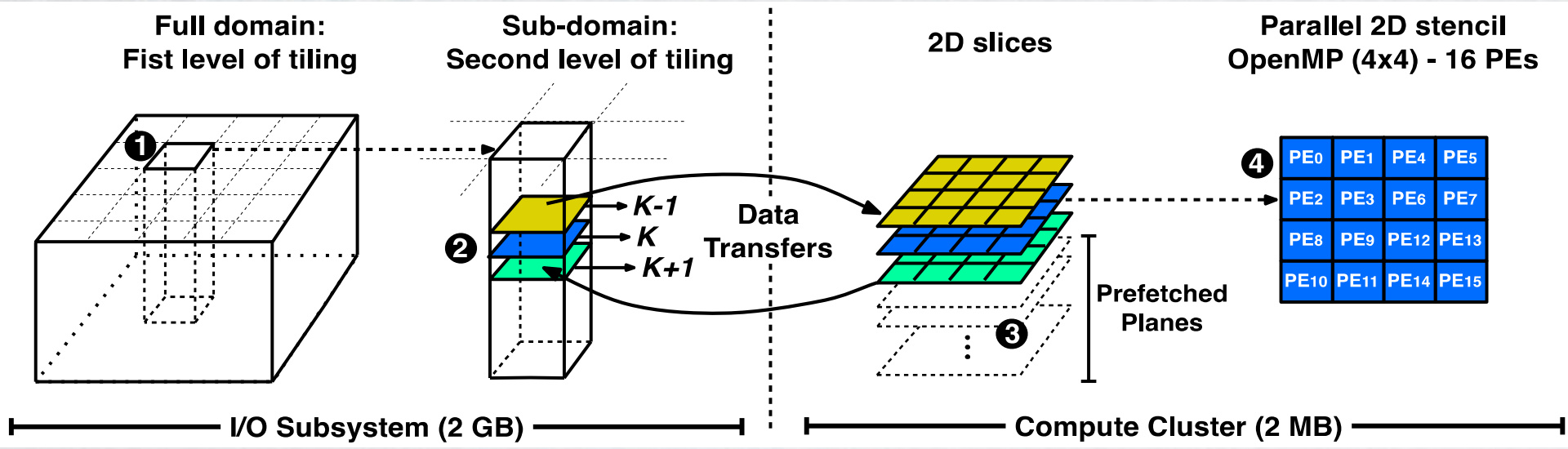
- **Memory**
  - Real seismic simulation data **don't fit in 2MB per cluster**
  - No cache coherency
  - Data transferts from/to the DRR **explicitly managed by the programmer**

- **Data transfers**
  - **Specific API** to perform data movements
  - Asynchronous data transfers to **overlap communications with computations**

- **NoC**
  - Data transfers should **match the NoC topology** to reduce communication costs
  - **Send few data transfers containing large amount of data** is better than several data transfers containing few data

Overview of Parallel Execution on MPPA-256

- **Two-level tiling scheme** to exploit the memory hierarchy of MPPA-256

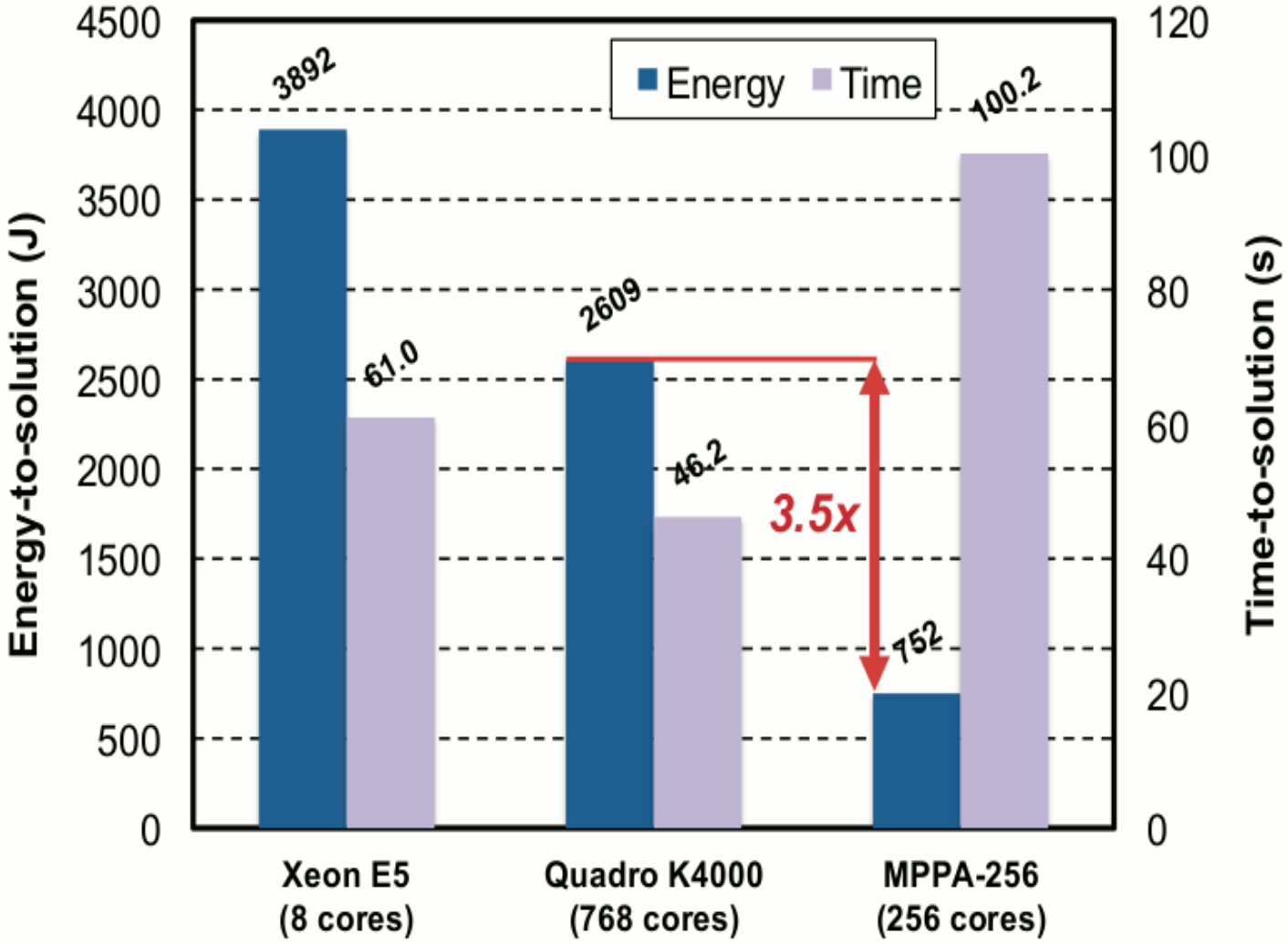


Results : MPPA vs Multicore and manycore platforms

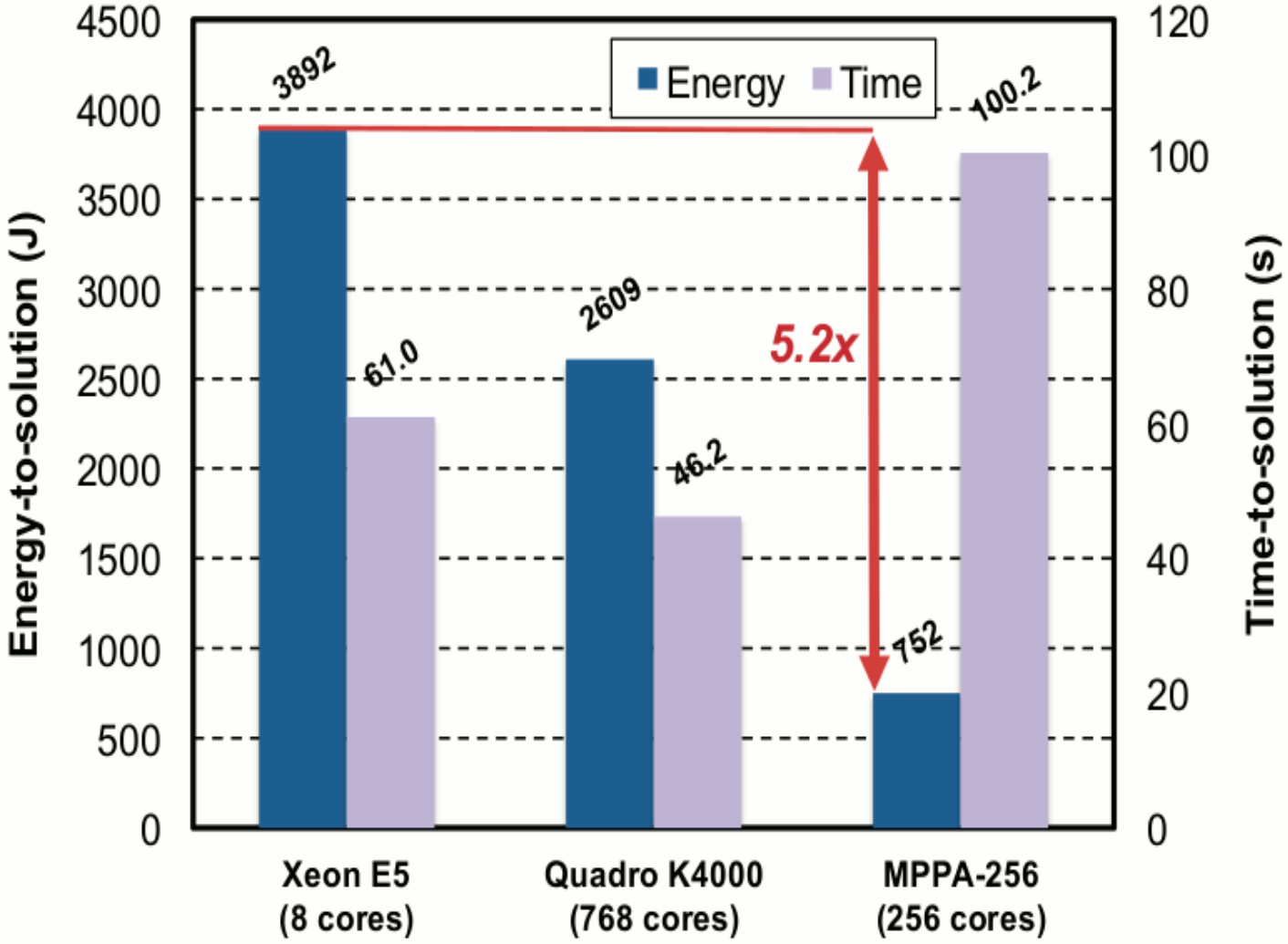
- **Intel Xeon E5**
  - Sandy Bridge-EP with 8 cores at 2.4 GHz
  - 32 GB of DDR3

- **NVIDIA Quadra K4000 GPU**
  - Kepler architecture
  - 768 CUDA cores at 800 MHz
  - 3GB of main memory

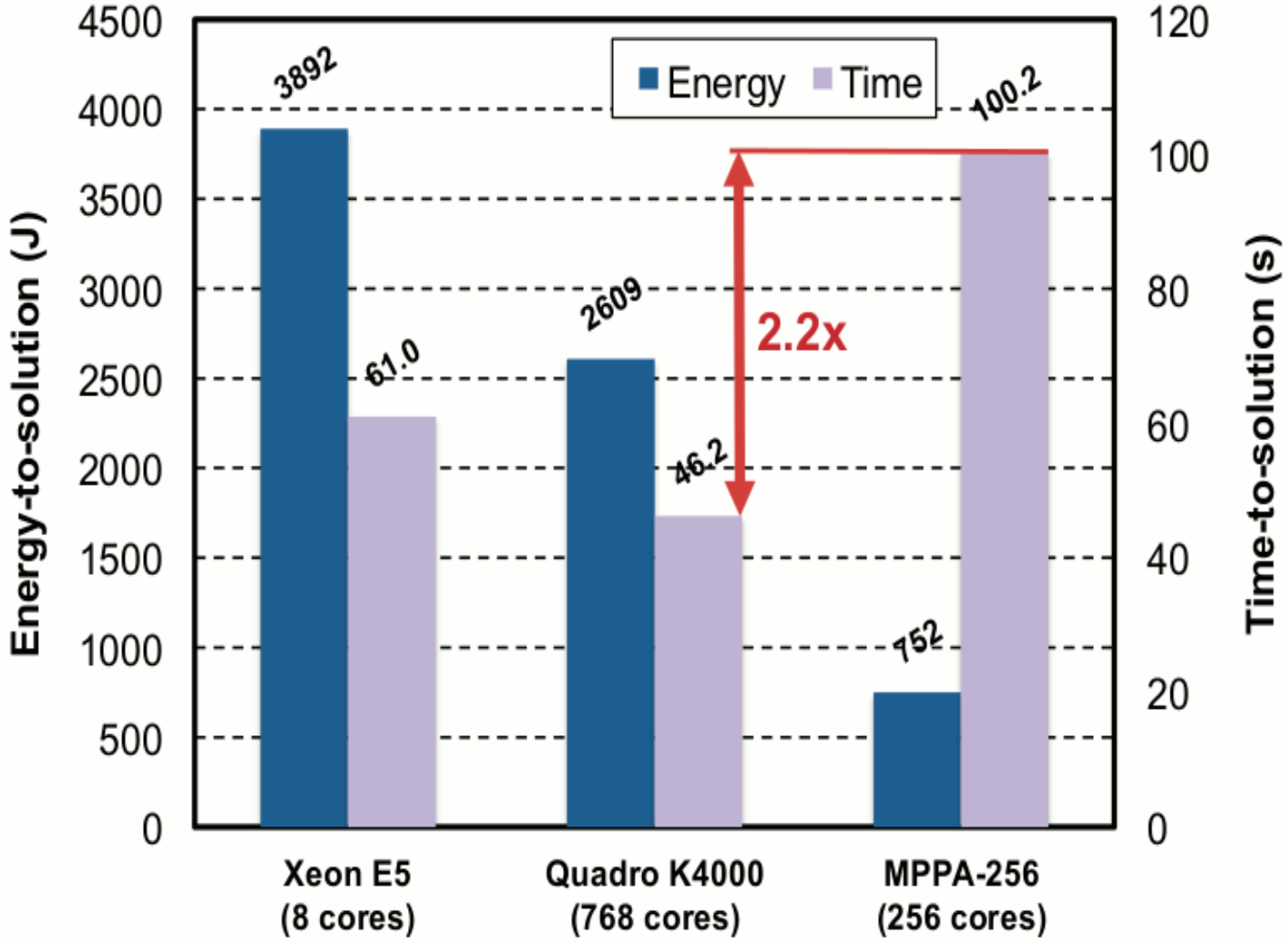
Results - Input problem size of 2 GB



Results - Input problem size of 2 GB



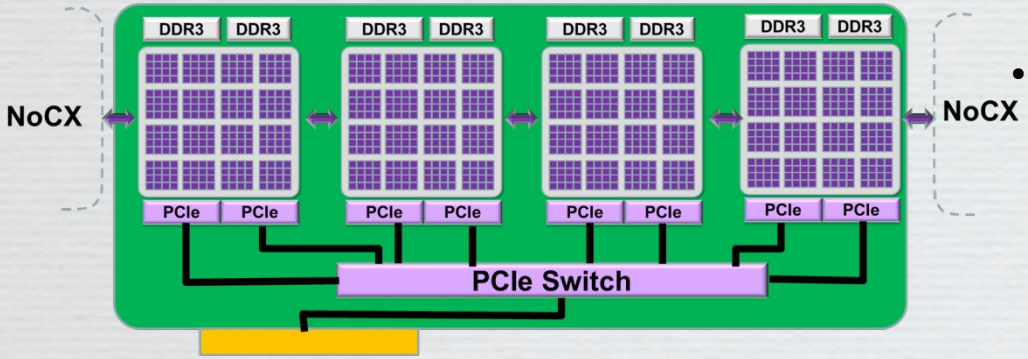
Results - Input problem size of 2 GB



Kalray/MPPA Roadmap

- **Multi-MPPA co-processor**
  - Kalray recently (SC) announced a multi-MPPA board that features four MPPA-256 processors with less than 50 W of power consumption

- **TURBOCARD2 : available now**
  - 4 MPPA-256@400MHz
    - 1 TFLOPS SP /300 GFLOPS DP
    - 32 bit architecture
    - 32 GB DDR3 1600 MT/s



- **TURBOCARD3 : <u>available 3Q'15</u>**
  - **4 MPPA-256@800MHz**
    - **3,2 TFLOPS SP / 1,6 TFLOPS DP**
    - **64 bit architecture**
    - **64 GB DDR3 2133 MT/s, 8 channels**

IEEE computer society

26th International Symposium on Computer Architecture and High Performance Computing

Paris, France

2014

Best Paper Award

to

"Márcio Castro, Fabrice Dupros, Emilio Francesquini, Jean-François Méhaut and Philippe Navaux"

for the paper

"Energy Efficient Seismic Wave Propagation Simulation on a Low-power Manycore Processor"

General chair

PC Chairs

Journal of Parallel and Distributed Computing

Available online 13 November 2014

In Press, Corrected Proof — Note to users

On the energy efficiency and performance of irregular application executions on multicore, NUMA and manycore platforms

Emilio Francesquini[a, b, g], Márcio Castro[c, d], Pedro H. Penna[e], Fabrice Dupros[f], Henrique C. Freitas[e], Philippe O.A. Navaux[c], Jean-François Méhaut[g]

Show more

doi:10.1016/j.jpdc.2014.11.002

Highlights

- Programming for a manycore is challenging.
- Limited memory and NoC are among the most important constraints of manycores.
- For CPU-bound and mixed workloads, MPPA-256 achieves better performance than Xeon.
- MPPA-256 consumes up to 13 × less energy than embedded and general-purpose multicores

HPC GA Results

- **About 30 joint publications**
  - Journals (15) : JPDC, FCGS, ParCo, IJPP, Cluster Computing, Journal of Geophysical Research, Seismological Research Letters, Journal of Computational and Applied Mathematics, Mathematical Modeling,…
  - Conferences (15), 1 Best Paper Award (SBAC PAD 2014)
- **2 Best PhD Thesis Awards (in Computer Architecture and HPC)**
  - 2013 Marcio Castro (INRIA)
  - 2014 Laercio Pilla (UFRGS/INRIA)
- **Cotutelle PhD (UFRGS/University of Grenoble)**
  - 3 UFRGS and INRIA
    - Laercio Pilla,
    - Joao Lima
    - Francieli Boito

HPC GA Results (2)

- **ESR students became Associate Professor**
  - Laercio Pilla (UFRGS/INRIA) → University Federal of Santa Catarina (UFSC)
  - Marcio Castro (INRIA/UFRGS) → University Federal of Santa Catarina (UFSC)
  - Joao Lima (UFRGS/INRIA) → University Federal of Santa Maria (UFSM)
  - Josue Tago (UNAM/UJF) → University National Autonoma of Mexico (UNAM)

HPC GA Mobility

- **UFRGS → INRIA**
  - 24 missions (5 ER, 19 ESR)
- **INRIA → UFRGS**
  - 7 missions (4 ER, 3ESR)
- **UNAM → INRIA/UJF**
  - 6 missions (4 ER, 2 ESR)

Thank you !