Re-engineering **astrophysical** and **material science** codes

Exa-Scale joint workshop

HiPeac, Manchester, Jan 2018

Luca Tornatore, G. Taffoni, D. Goz – *INAF*

P. Gorlani, G.P. Brandino, S. Cozzini – *Exact-LAB*

L. Lavagno – *Politecnico di Torino*

Outline

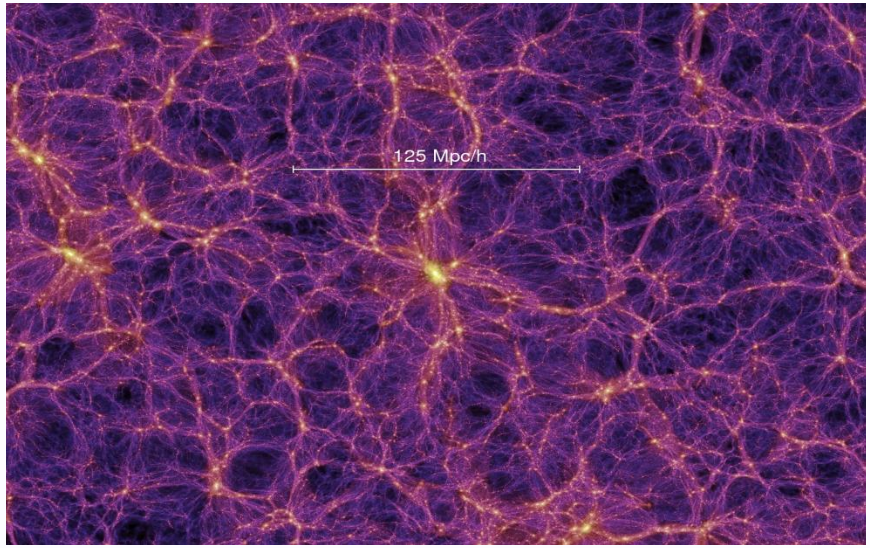
Summary report of the achieved results in **3 different projects** in the frame of **ExaNeSt** collaboration:

- *INAF*
  2 codes from Astrophysics re-engineering FPGA
- *Exact-LAB*
  1 code from Material science FPGA

First code

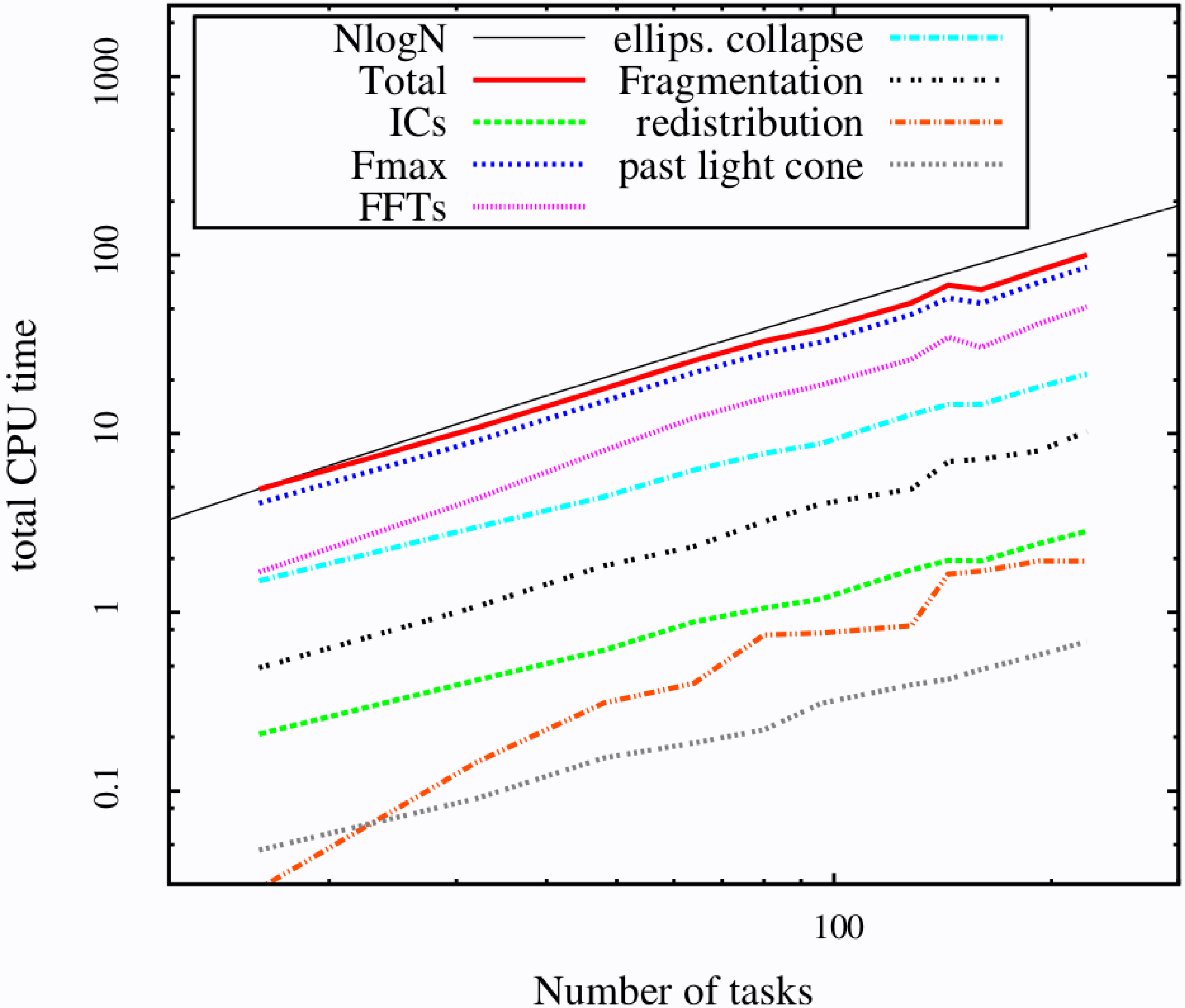
- Astrophysics: PINOCCHIO
  Luca Tornatore, G. Taffoni – INAF
- Astrophysics: HiGPUs
- Material science: LAMMPS

PINOCCHIO and the Cosmic Web



- **Very fast** (~10^3 to N-body codes) and approximate (~ few %) solution to the full non-linear gravitational problem, to obtain positions and mass of hirarchically collapsed objects in the Universe. Based on **Lagrangian Perturbation Theory (LPT)** and **FFTs**
- Currently used in the ESA's EUCLID project to build ***thousands*** mock realizations of the Universe at the ***highest possible resolution*** to calculate a covariance matrix

PINOCCHIO: already fast and well-behaved



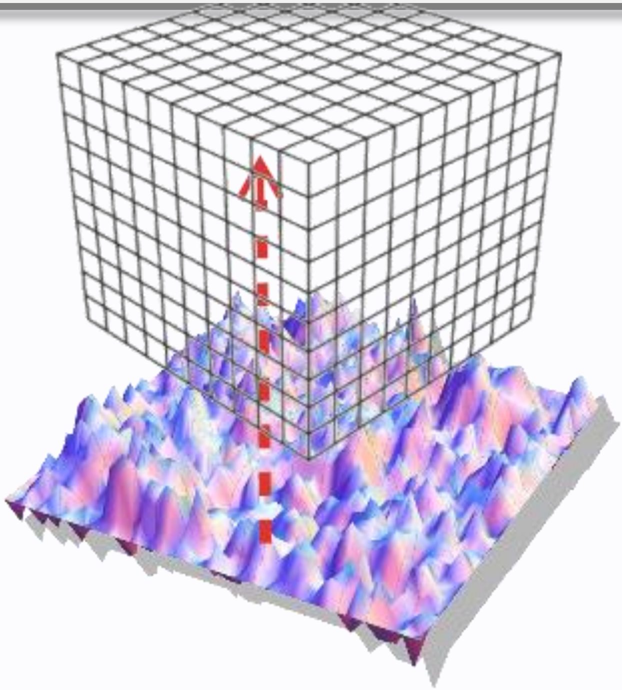
We concentrate on:

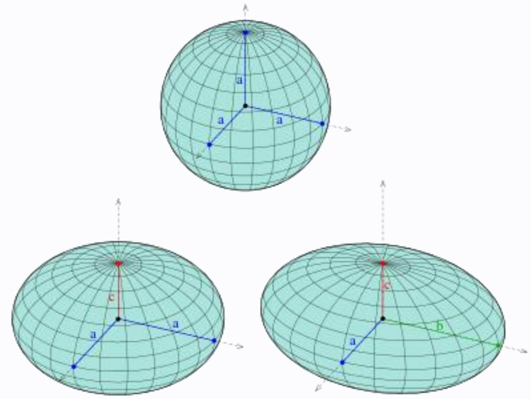
Fmax =
FFTs +
Ellipsoidal collapse

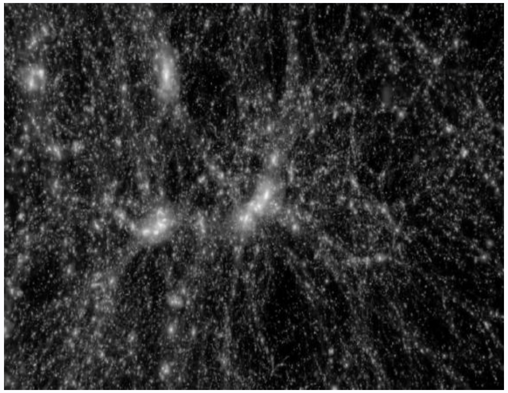
Initial Condition Generatios

i.e. ~ 60-80% of the total time

PINOCCHIO: 3 major steps







Generation of **density** in Fourier Space

Calculation of **ellipsoidal collapse**

Hierachical assembly of objects

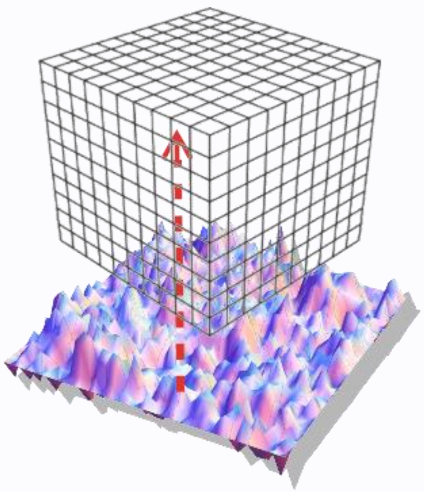
~70% of wall-clock time

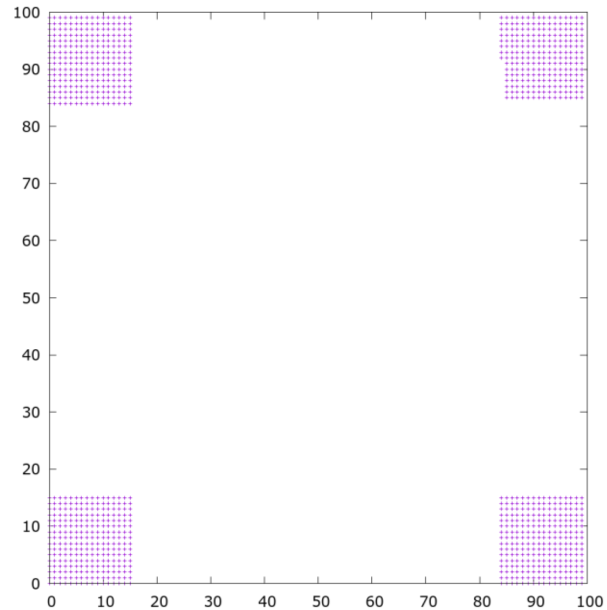
Memory issues

Computation intensive

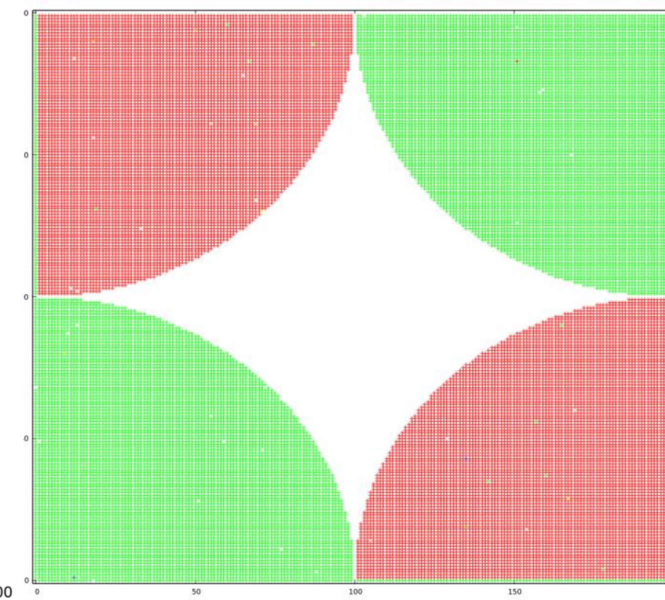
Memory - and Communication-intensive

1 – Refactoring the i.c. generation





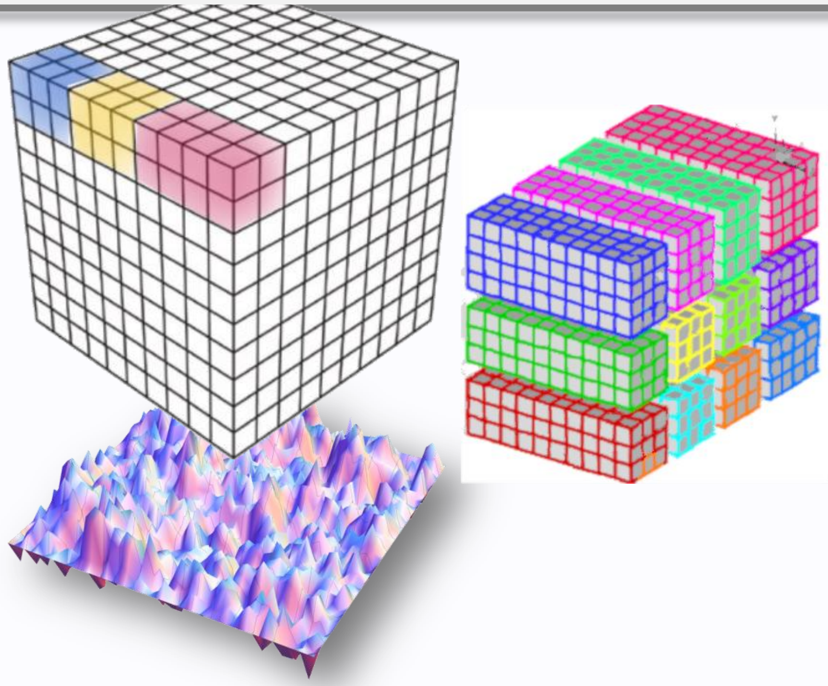
Subsequently generated pseudo random-nums, from corners inwards

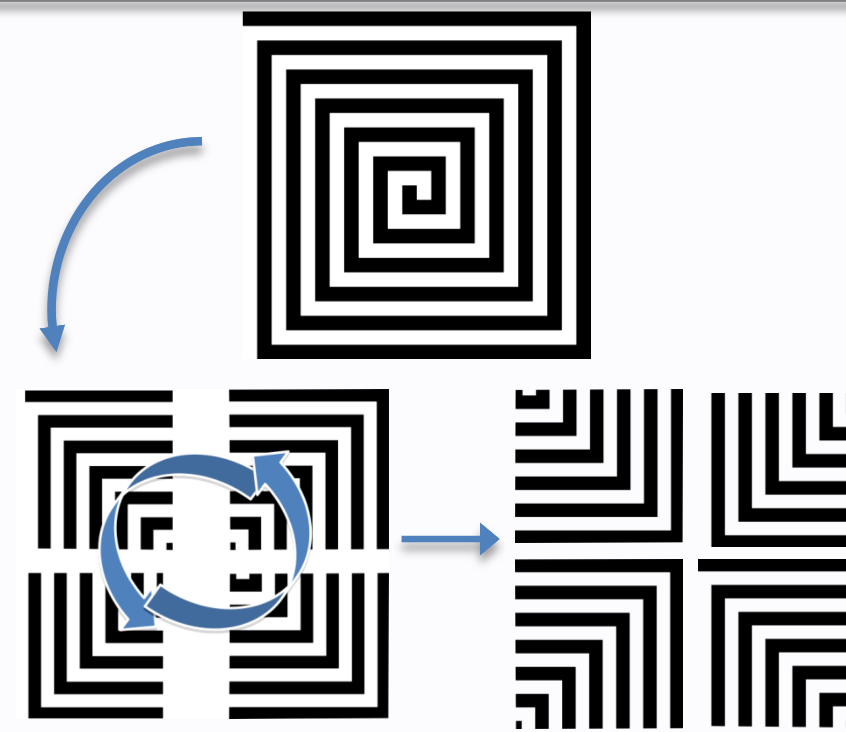


Plane used with some special symmetries: --> all in memory

Initial power spectrum has physical properties and symmetries; to exploit them, it's built from a random field that must be entirely resident to all MPI tasks, posing **severe memory limitations**, since we aim to N → several 10^4 or more

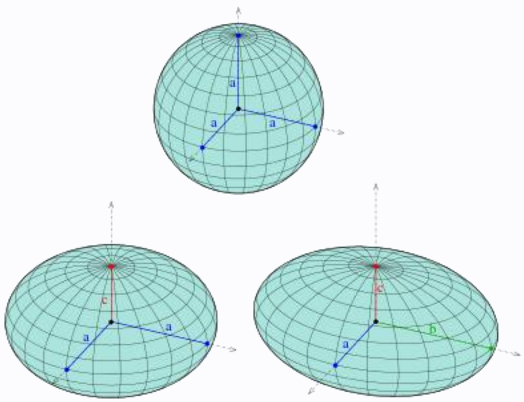
1 – Refactoring the i.c. generation





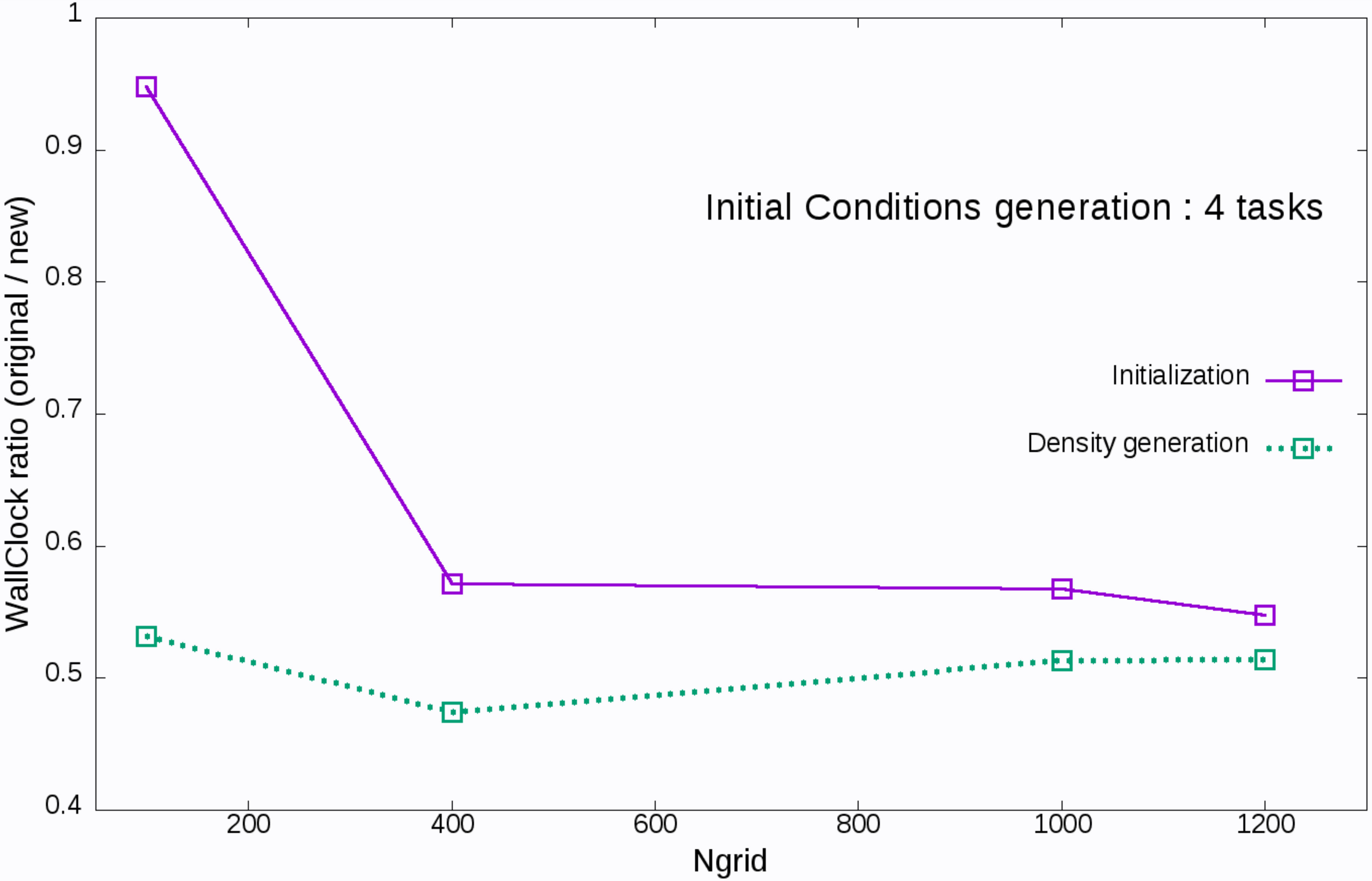
- 2D/3D decomposition for FFT, instead of 1D
- completely re-designed algorithm to generate power-spectrum
  - ✓ has same symmetries and properties
  - ✓ each MPI task must have only its portion of the initial random field
- re-engineering of memory patterns

2 – Refactoring Ellipsoidal Collapse

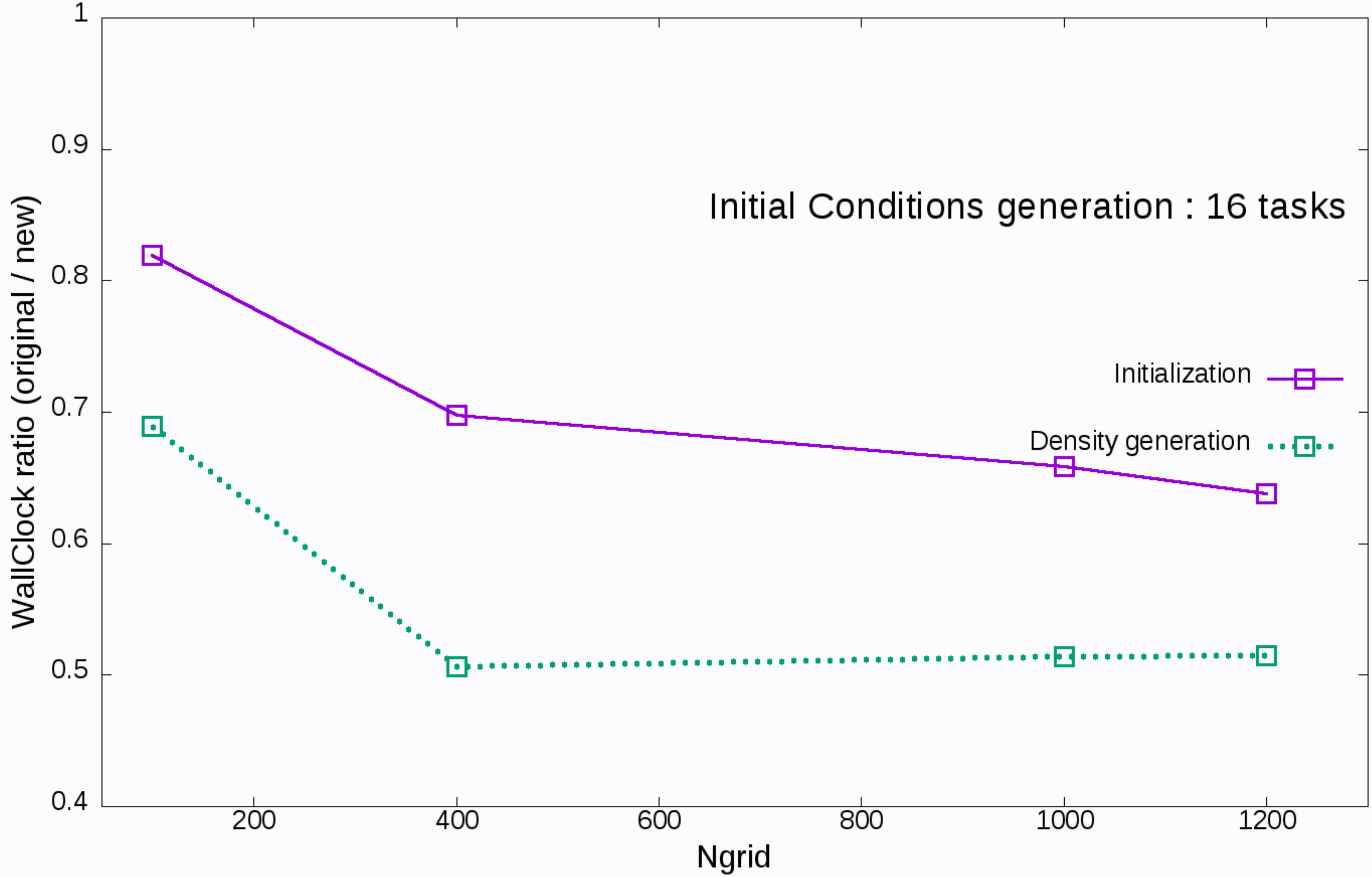


- **Re-engineering** of loops and floating point operations
- **Vectorization** through **AVX/AVX2** SIMD instructions
- Twice as fast
- Some intrinsic conditionals that may be masked efficiently with **AVX-512** set

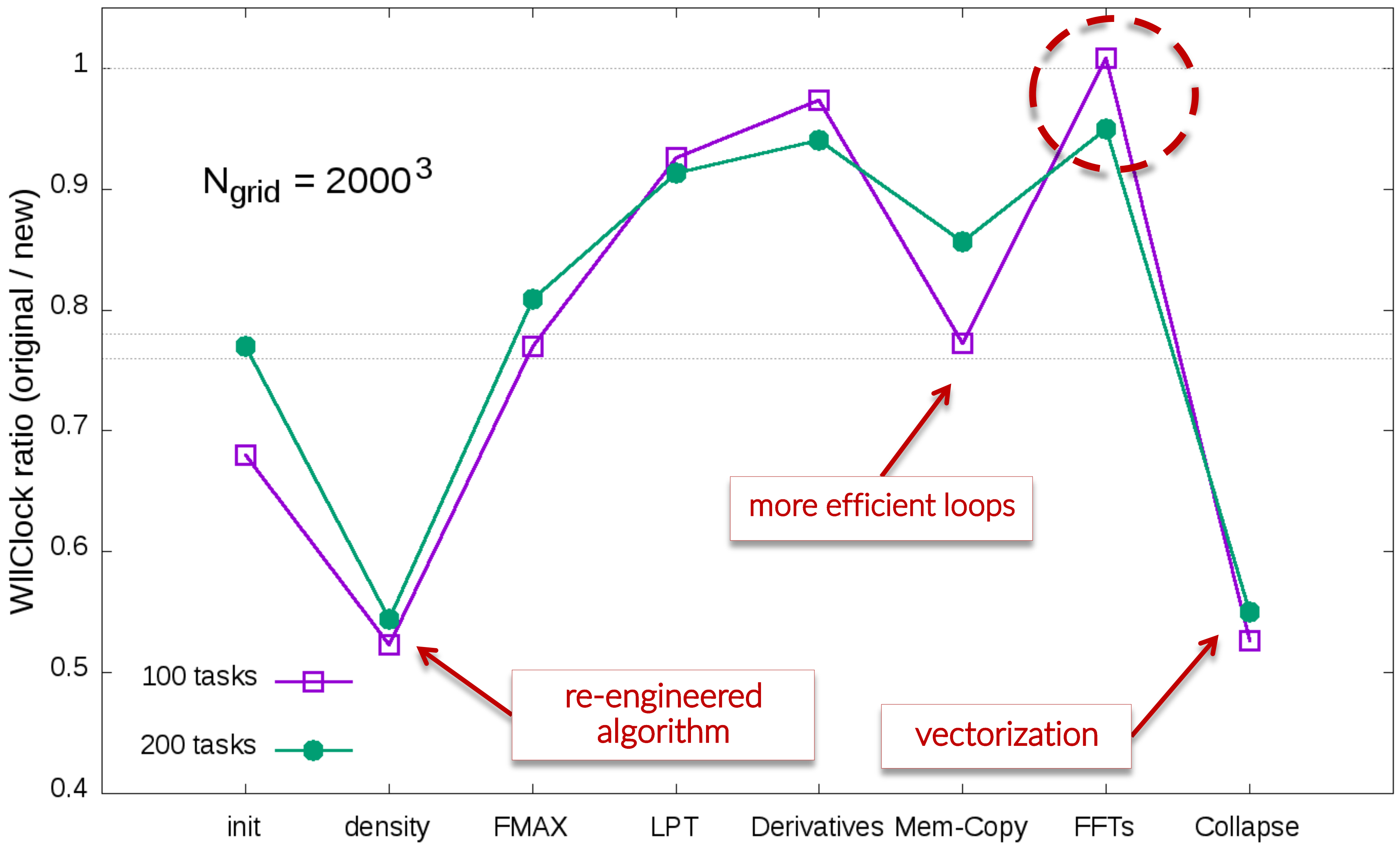
I.C. generation : results



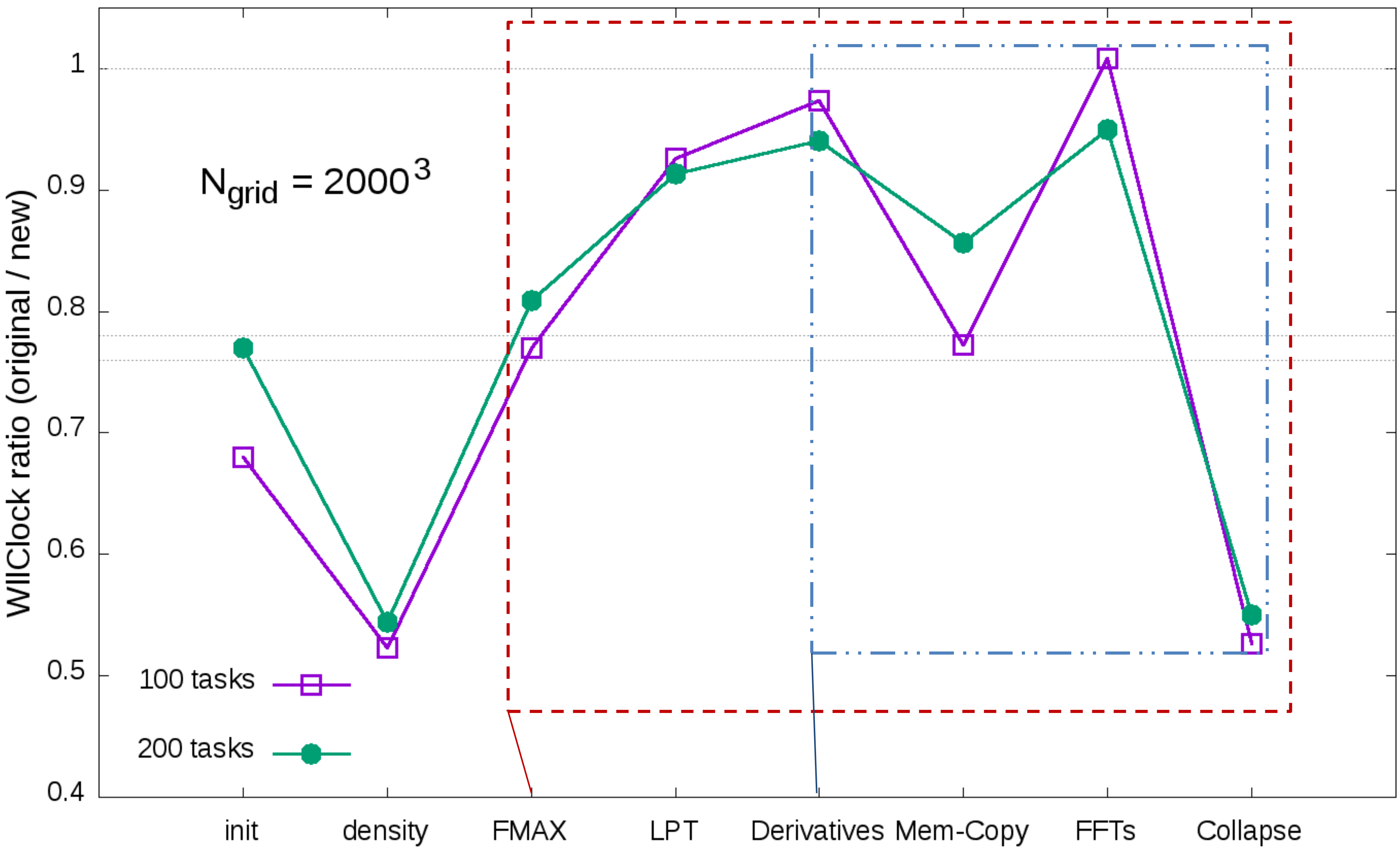
I.C. generation : results



Current status



Current status

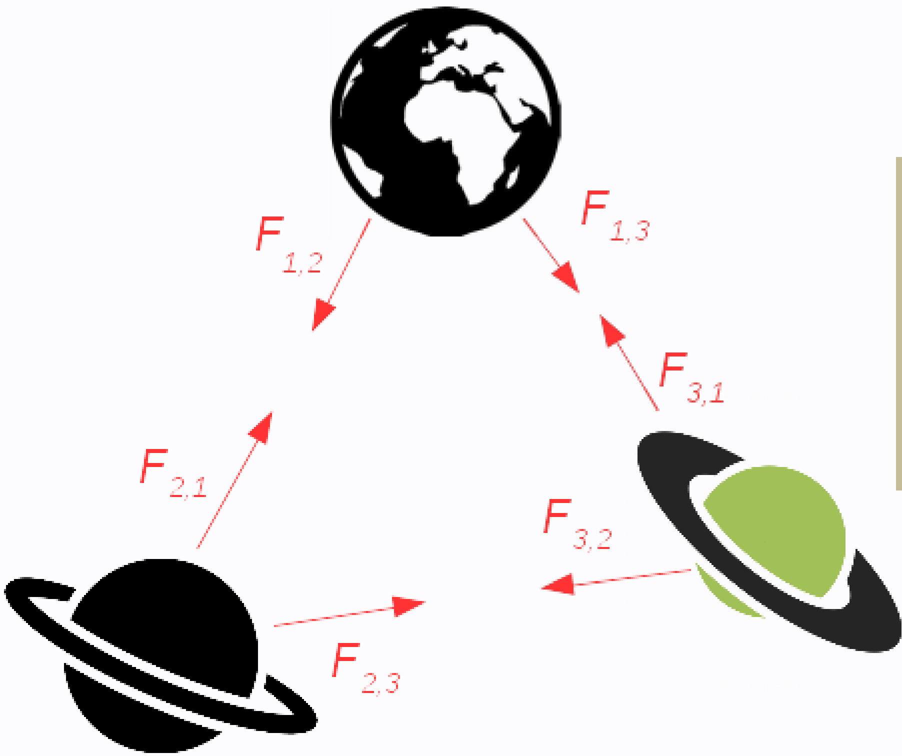


Second code

- Astrophysics: PINOCCHIO
- Astrophysics: HiGPUs
  David Goz – INAF
  Luciano Lavagno – Politecnico di Torino
- Material science: LAMMPS

ExaNeSt

Direct *N*-Body problem



ALL PAIRS	ALL BODIES
COMPUTE FORCES	*UPDATE VELOCITIES AND POSITIONS*
O(N²) → $O(N^2)$	*O(N)*

Tree-, **PM-** and Tree+PM- based methods are much faster (~*O(NlogN)*) but are approximate, not suited for ***real close encounters***

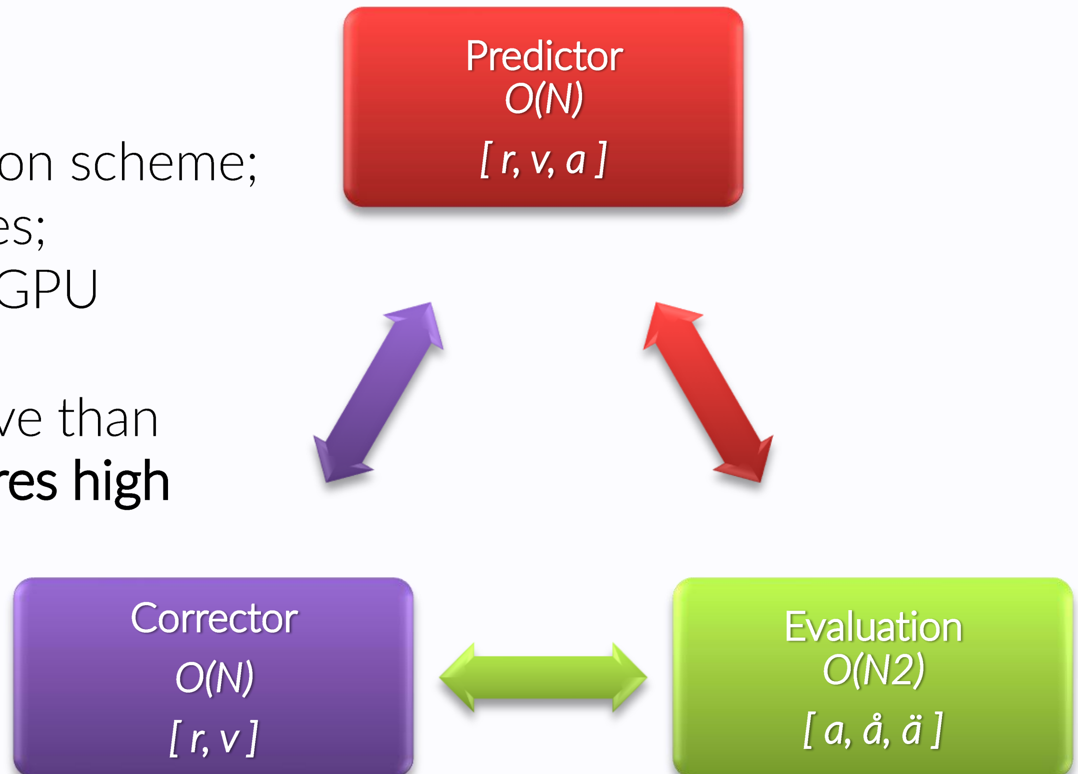
HiGPUs

HiGPUs (R. Capuzzo-Dolcetta, M. Spera, D. Punzo 2013) is a **direct *N*-body code** suitable for studying the dynamical evolution of stellar systems composed up to **10 millions** of stars.

It features:

- ***Hermite 6th order*** time integration scheme;
- individual time step for particles;
- implementation optimized for GPU

Hermite scheme is more expensive than lower-order integrators but **ensures high orbit accuracy**



Porting *N*-body codes on FPGA

Proposed solutions available in literature (in **single precision)**:

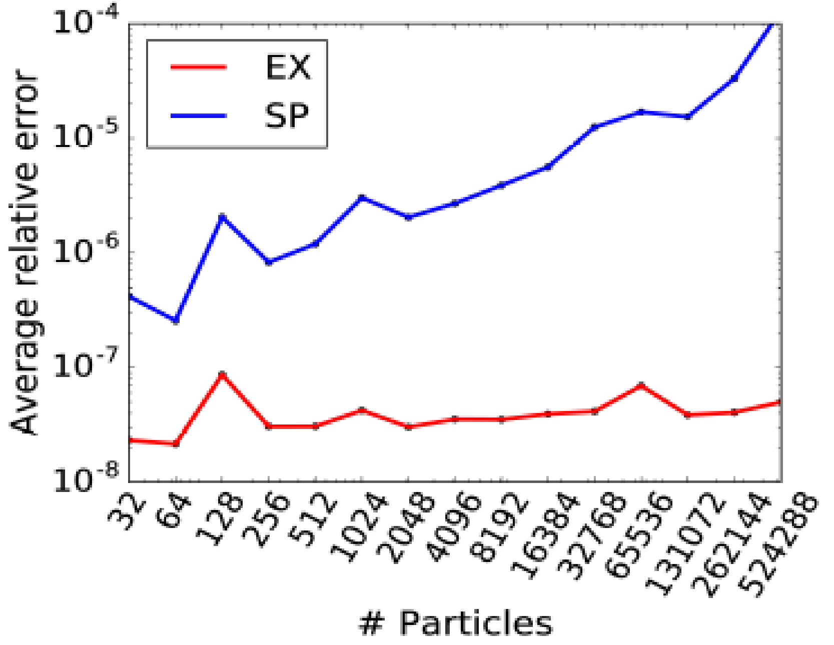
- based on hierarchical Tree algorithm (Kawai et al. 2006);
- based on first order Simple Euler method (Peng et al. 2016, Del Sozzo et al. 2017).

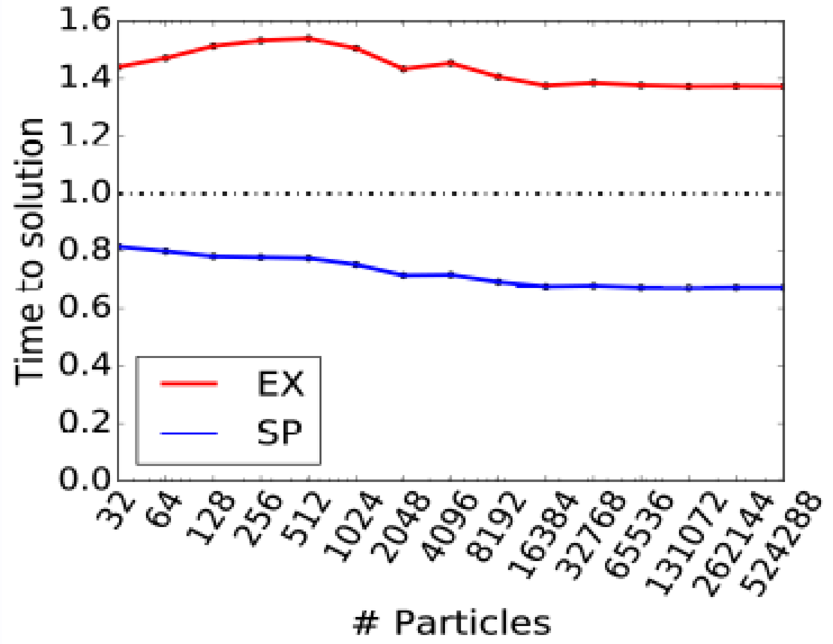
Our aim:

Porting the **full double-precision Hermite scheme**, starting from the most computationally-intensive kernel

Strategy	Motivation
Vectorization	Smaller amount of ops
Use of ***local memory***	Memory burst mode
Extended-precision arithmetic	Trade-off btw accuracy and resources usage

Results (on GPUs testebed)





- *EX*-arithmetic ensures to keep control over the accumulation of the round-off error during the simulation;
- time to solution reveals some overhead to handling *EX*-arithmetic.

→ VIVADO HLS fails in generating the correct RTL when using *vector types*

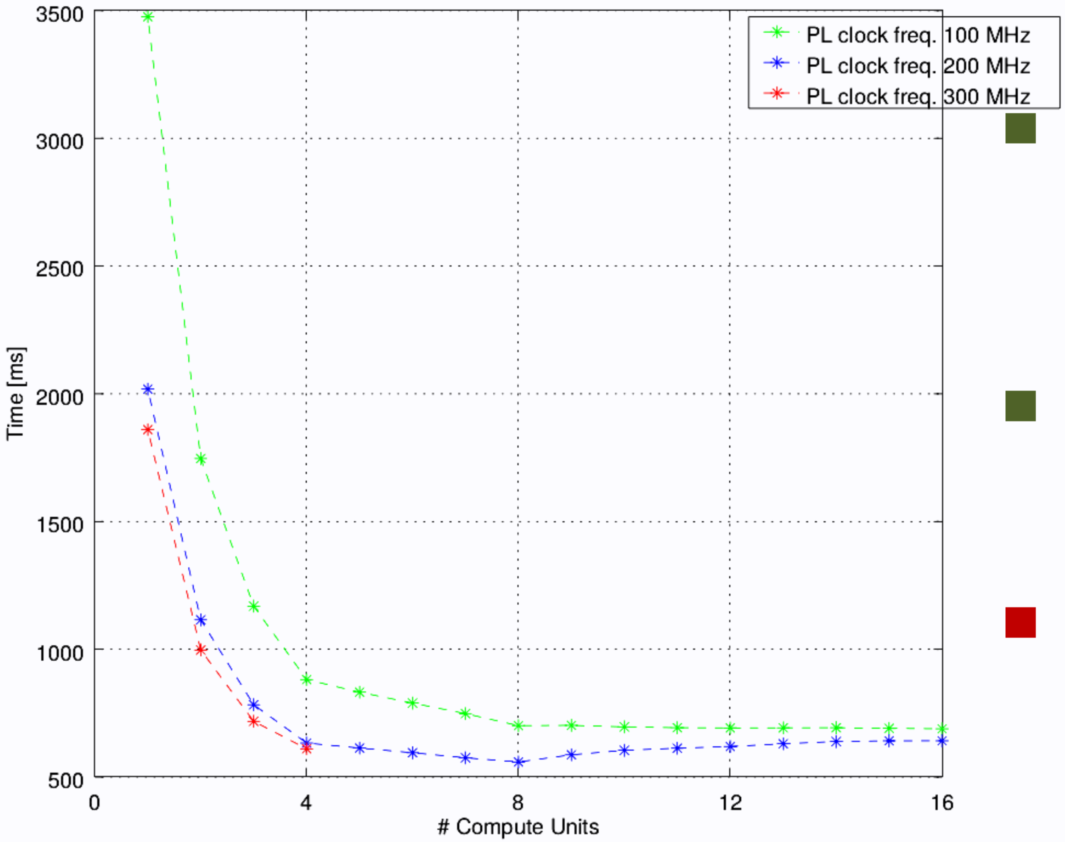
Third code

- Astrophysics: PINOCCHIO
- Astrophysics: HiGPUs
- Material science: **LAMMPS**
  P. Gorlani, G.P. Brandino, S. Cozzini – Exact-LAB

Porting of miniMD on Zynq UltraScale+

- MiniMD is a miniApp for **molecular dynamics**, modelled after **LAMMPS**
- Same algorithmic complexity but reduced features and code base size
- The OpenCL kernels of miniMD have been ported and optimized on the FGPA

The biggest limit: memory bandwidth

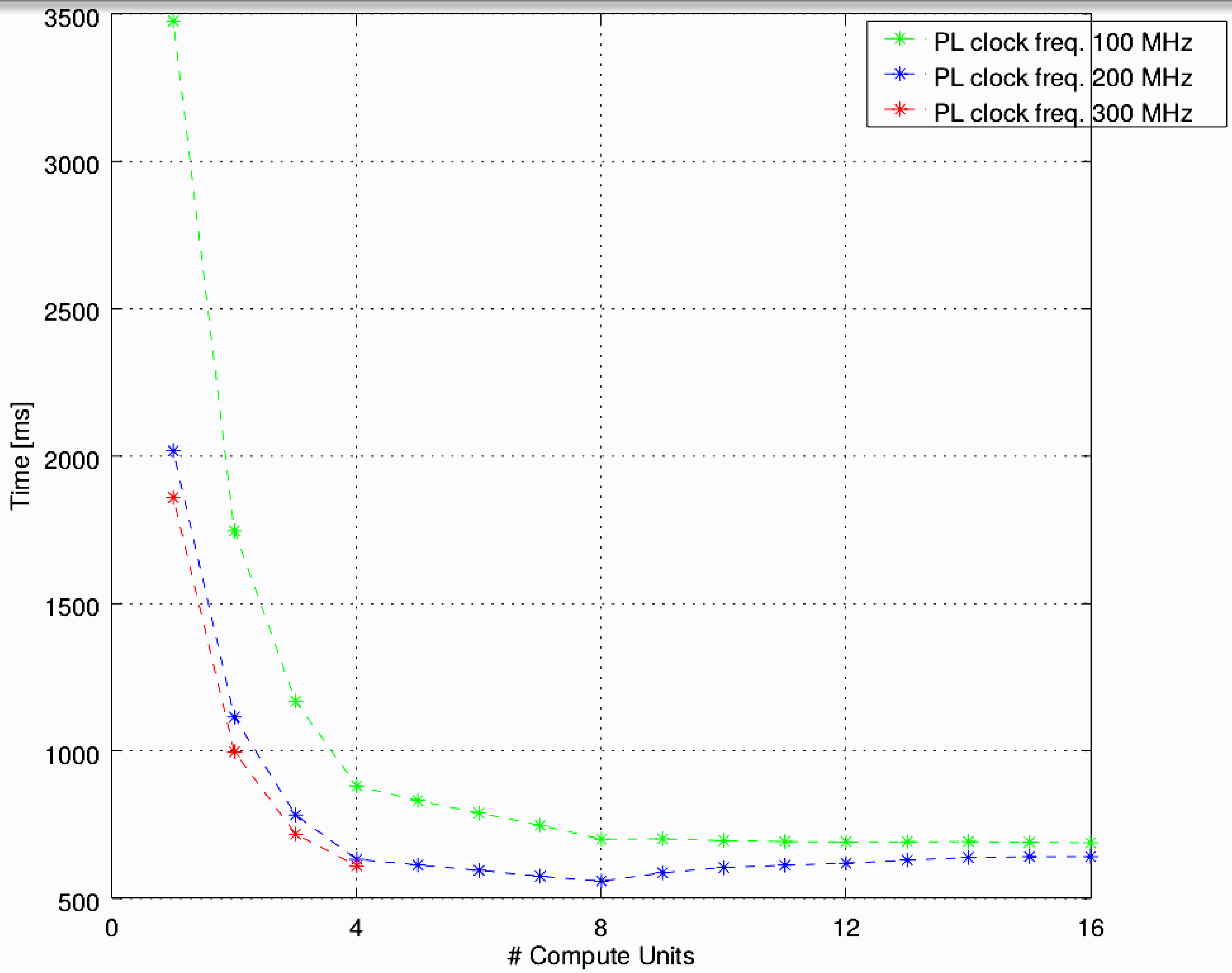


- Good improvement with respect to running on ARM cores, still much slower than GPUs
- No separation of host memory/device memory (and related data transfers)
- Even for compute-bound kernels, the small memory bandwidth is still the biggest limitation.

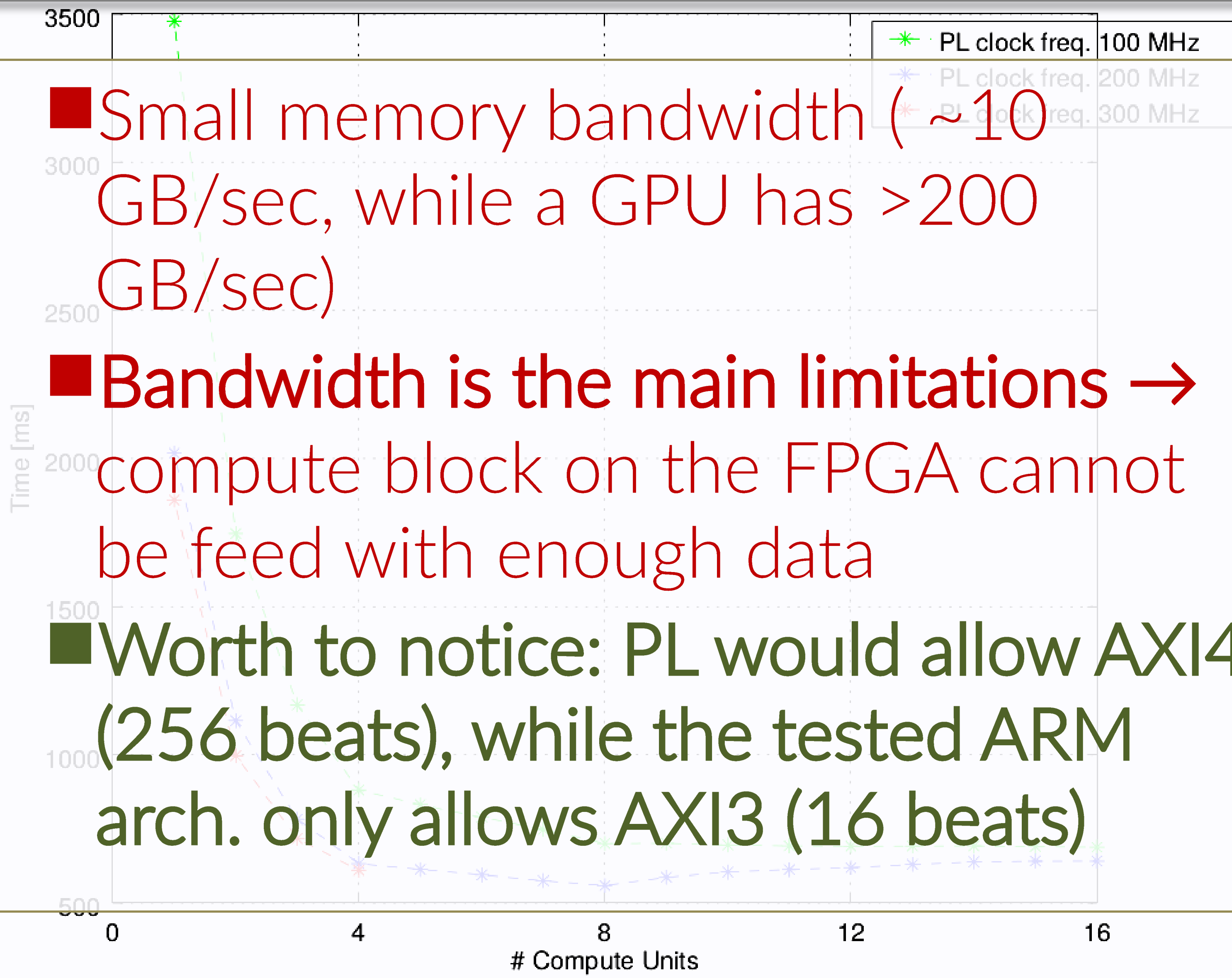
ARM CPU	FPGA	K20
1300 ms	560 ms	10 ms

Execution time of the force Kernel on 1024^2 particles

The biggest limit: memory bandwidth



The biggest limit: memory bandwidth



- Small memory bandwidth (~10 GB/sec, while a GPU has >200 GB/sec)
- **Bandwidth is the main limitations →** compute block on the FPGA cannot be feed with enough data
- Worth to notice: PL would allow AXI4 (256 beats), while the tested ARM arch. only allows AXI3 (16 beats)