Dark Energy and Massive Neutrino Universe Covariances
(DEMNUniCov)

Carmelita Carbone

Physics Dept, Milan University & INAF-Brera

Collaborators: M. Calabrese, M. Zennaro, G. Fabbian, J. Bel

November 30th 2017, ICT workshop, Bologna

"Dark Energy and Massive Neutrino Universe"

(PI Carmelita Carbone)

DEMNUni-Covariances: Class A approved project, INAF/Cineca HPC Resources, 1st call, 3M cpu-hr assigned on Marconi-A1 (BRD) Tier-0 @CINECA (**ONGOING**)

→ 300 cosmological N-body simulations with V=1 $(\mathrm{Gpc}/h)^3$ and $N_{part}=2 \times 1024^3$ (CDM+ν)

Related previously awarded projects:

1. DEMNUni-I: ISCRA class-A, 5M cpu-hr on Fermi Tier-0 @CINECA (**COMPLETED**)
2. DEMNUni-II: ISCRA class-B, 8M cpu-hr on Fermi Tier-0 @CINECA (**COMPLETED**)

14 cosmological simulations with volume: $(2\ \mathrm{Gpc}/h)^3$ and N_{part}: 2×2048^3 (CDM+ν)
baseline Planck cosmology + neutrino masses + dynamical dark energy

$$\mathbf{M_\nu=0,\ 0.17,\ 0.3,\ 0.53\ eV\ \&\ (M_\nu,w_0,w_a)=(0\div0.16,-0.9,\pm0.3),(0\div0.16,-1.1,\pm0.3)}$$

200 TB stored data on DRES @CINECA

Scientific goals

Production of covariance matrices in the presence of massive neutrinos and dynamical dark energy for different probes:
GC, WL, CMB secondary anisotropies, cross-correlations

Six different cosmological models varying ($\mathbf{M_\nu}$, $\mathbf{w_0}$, $\mathbf{w_a}$), 50 sims/model

Final products: covariances from

- mock galaxy catalogs via box-stacking and HOD/SHAM techniques (Zennaro)
- mock shear maps via box-stacking and ML ray-tracing (Calabrese)
- mock CMB weak-lensing and ISW/Rees-Sciama maps (+ runs Carbone)
- Comparison against approximated methods (Bel)
- Possibly use the maximum-likelihood approach from Sellentin&Heavens (in prep) for high-precision covariances-fitting, using only 50 sims per cosmology.

Stored data

- 5 full particle (CDM+ν) snapshots (62.5 GB/snap) between $0<z<2$ (95 TB in total)
- full FoF/Sub halos catalogs from $z<2$ (40 TB in total)
- matter/neutrino power spectra from $z<99$ (negligible)
- Intermediate weak/CMB-lensing & ISW/Rees-Sciama maps (7 TB in total)
- Temporary data: snapshots+potential-grids for 25 simulations/month (180 TB in total)
- Used area: marconi_scratch (temporary data) & marconi_work/INA17_C1A06

Gadget-3: proprietary OpenMP/MPI hybrid version

(Springel 2005 & later versions)

- Hybrid tree-PM method: long-range force calculated at low resolution with a particle-mesh scheme, assigning the particles to a regular cubic mesh, using Fourier methods to obtain the corresponding potential, and numerically differencing the result

- High resolution short-range correction calculated using a tree algorithm. The short-range correction is assembled in real space by collecting contributions from all neighbouring particles.

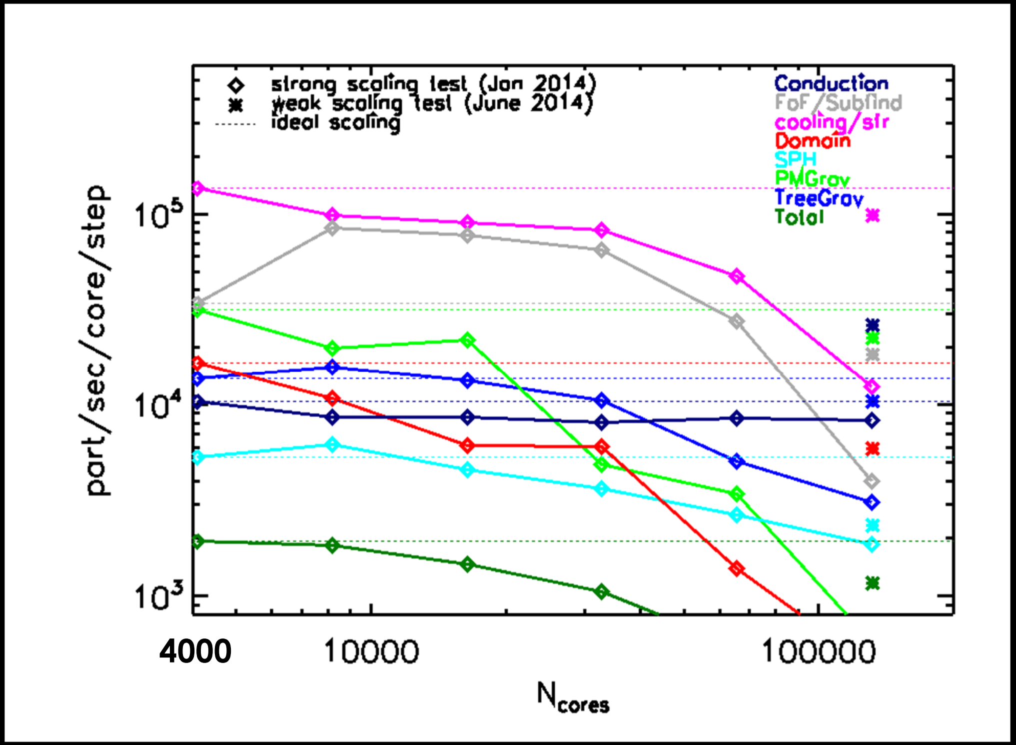
- Space-filling fractal, the Peano-Hilbert curve, to control the domain decomposition associated with parallelisation. The tree decomposition used by the code is independent of the platform, in particular of the number of processors.

- Neutrinos in Gadget-3 (Viel, Haehnelt & Springel 2010): Massive neutrinos are incorporated as a separate low-mass collisionless particle species

Needed Libraries

- Intel-MPI compiler in C (big-bugs big-problems !!)
- GSL - GNU scientific library
- Hierarchical Data Format (HDF5) library
- Szip compression in HDF5
- ZLIB data compression library
- Fastest Fourier Transform in the West (fftw 2.1.5 for the used Gadget-3 version)

OpenMP/MPI Gadget-3 Scalability

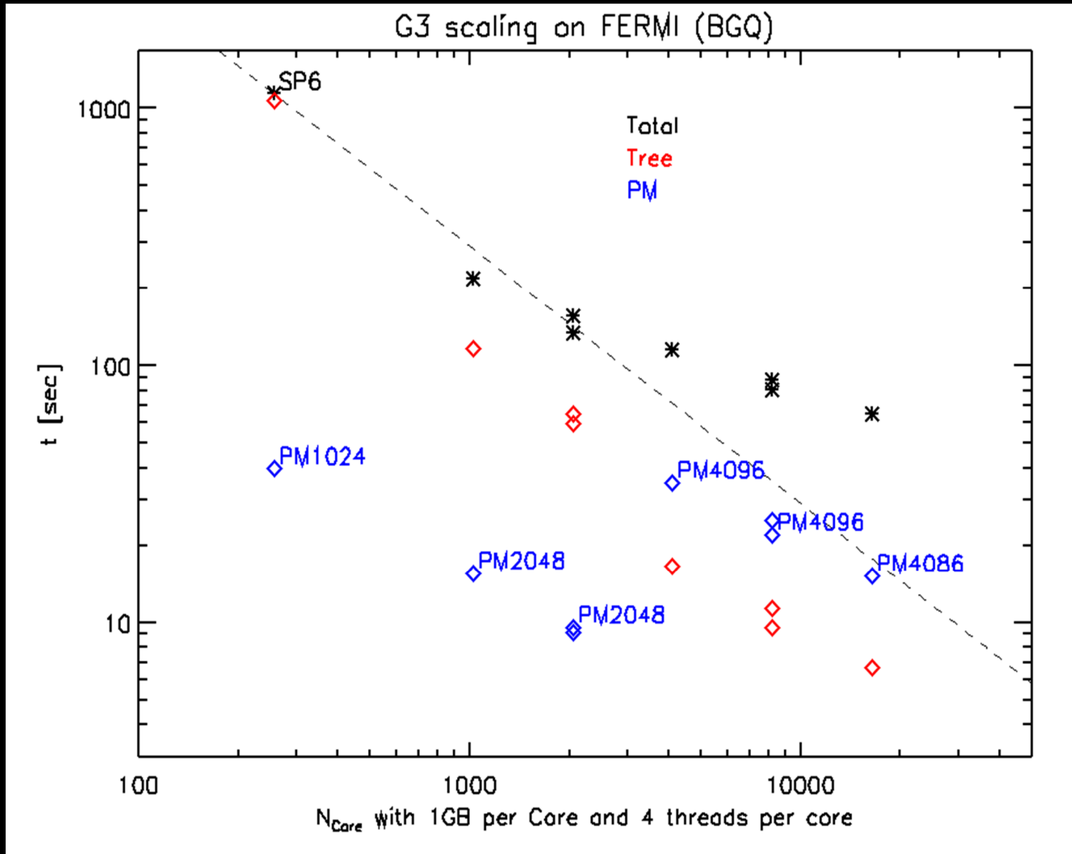


- Scaling tests for the used Gadget-3 version: processed particles within one time-step per second and per core, split into different parts of the simulation. Ideal scaling would be horizontal lines. The timings cover up to the usage of 16 islands of SuperMUC.

- Different parts of the simulation are differently affected due to their different algorithmic properties, and the last data point strongly suffer from the fact that, in the shown case, only 3% of the available memory for each core is used.

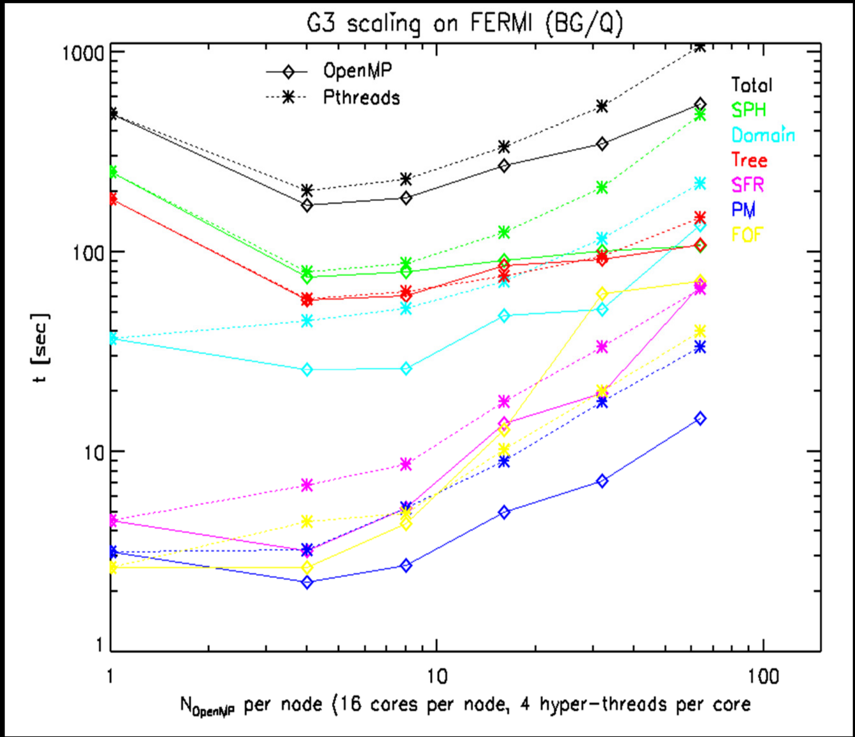
- Perfect scaling for 4096 cores as used on Marconi-A1

OpenMP/MPI Gadget-3 Scalability



Time to solution of Gadget3 on FERMI
for different PM grid-sizes as a function of the core number
(for a DM only simulation)

OpenMP/MPI Gadget-3 Scalability



Shown is the OpenMP performance on FERMI of Gadget3 (version 09/2012), switching between MPI-tasks and OpenMP threads.

(FERMI) # cores	absolute timing (s)	speedup
4096	114,44	1
8192	80,49	1,42
16384	64,68	1,77

Marconi-BRD is 10 times faster than Fermi (when it works ;)

Used resources

```
[ccarbone@r000u08l03 P-Gadget3_5_8_2013]$ saldo -b -a INA17_C1A06

---------------------------------------------------------------------------------------------------------------------------------------------
account                 start        end          total        localCluster   totConsumed      totConsumed      monthTotal     monthConsumed
                                               (local h)   Consumed(local h)    (local h)                %       (local h)        (local h)
---------------------------------------------------------------------------------------------------------------------------------------------
INA17_C1A06          20170515     20180519      2500000             1527734       1527734             61.1          203252           289992
[ccarbone@r000u08l03 P-Gadget3_5_8_2013]$
```

Problems in job execution (about 150k CPUh lost)

- Bug in the Intel-MPI compiler :

writing snapshot file #15...
writing block 0 (Coordinates)...
[MPID_nem_tmi_pending_ssend_dequeue]: ERROR: can not find matching ssend. context=0, rank=28, tag=0
[MPID_nem_tmi_pending_ssend_dequeue]: pending ssend list: (context=0, rank=28, tag=c, sreq=0x2b0cfa0)
MPID_nem_tmi_handle_ssend_ack: &pad1=0x2b2e7ab93688, &ssend_ack_recv_buffer=0x2b2e7ab93690, &pad2=0x2b2e7ab936a0
MPID_nem_tmi_handle_ssend_ack: pad1=0, pad2=0 (both should be 0)

- Instability of Marconi-A1: random errors, sometimes unreproducible

- Jobs running but no outputs

- Jobs running slower than usual

Open problems

- BIG PROJECTS → BIG DATA: What about storage facilities?

Cineca is very helpful and tries to meet user requests, but they cannot provide huge long-time storage for free, obviously…

What about an INAF-CINECA convention on storage facilities?

What about a common storage repository for big simulations in Italy?

- INAF-CINECA convention represents a big step forward HPC usage from the INAF community.

- However the amount of resources should be increased, especially in view of the INAF-PTA.

Calcolo Numerico

Tutti i filoni di ricerca della Cosmologia e astrofisica extra-galattica richiedono anche un'intensa attività di simulazioni numeriche e quindi robuste capacità di calcolo, utilizzando super-computer massivamente paralleli. Lo scopo di tali simulazioni è di descrivere, con tecniche numeriche ad N corpi ed idrodinamiche, la formazione ed evoluzione di galassie, ammassi di galassie e strutture cosmiche su grande scala a partire dalle condizioni iniziali date dal CMB. Le simulazioni cosmologiche forniscono un

quadro interpretativo unico grazie alla loro capacità di catturare la complessità dei fenomeni astrofisici e dinamici che guidano la formazione delle strutture cosmiche. Inoltre esse giocano un ruolo essenziale per la definizione dei casi scientifici e per il disegno di strumenti da terra e missioni spaziali dedicati ad osservazioni cosmologiche ed extra-galattiche. L'attività di ricerca in astrofisica numerica necessita, come l'astrofisica osservativa, di avere a disposizioni infrastrutture hardware ed una strategia di sviluppo di lungo termine, non solo per la costruzione e mantenimento di centri di calcolo, ma anche per lo sviluppo più generale di una "cultura del calcolo".

- Big surveys imply big simulated data:
  why INFN is investing much more than INAF?