Mem. S.A.It. Vol. 80, 357

Memorie della

The HORIZON project

H. Wozniak (on behalf of the HORIZON consortium)

Centre de Recherche Astrophysique de Lyon, F-69561 Saint-Genis-Laval cedex, France
e-mail: `herve.wozniak@obs.univ-lyon1.fr`

Abstract. I briefly summarize the main results obtained by the HORIZON computational astrophysics project and described two on-line services (GalMer and GalICS/MoMaF) that intent to give access to the simulations. I discuss some issues that were raised when developing these services.

Key words. Computational astrophysics - On-line services - Virtual Observatory

1. Introduction

Within the past two decades, computational astrophysics has seen a phase transition from being dominated by individual researchers writing and running their own "private" simulations on their personal workstation, to large projects that gather experts from different fields and use the huge computational power available through parallel machines or GRID technologies. Thanks to recent progresses in numerical techniques and available hardware, numerical cosmology has become one of the most important fields among the scientific topics that require high-performance computing.

A series of large simulations have been recently produced by, among others, (Cen & Ostriker 2000), the Virgo consortium (Frenk et al. 2000), Millenium simulation (Springel et al. 2005), and the Gasoline team (Wadsley et al. 2004).

Along the same path, the objective of the Horizon Project is to federate numerical simulations activities within the French community on topics such as large-scale structure formation in a cosmological framework, the formation of galaxies, and the prediction of its observational signatures.

The HORIZON Project is built on several research teams in different institutes, namely the CEA/SAp in Saclay, the Observatories of Paris (LUTh and LERMA laboratories), Lyons and Marseilles. The scientific objective is specifically oriented towards studying galaxy formation in a cosmological framework. Its transverse and federative nature will however allow to develop in a few years high-level expertise in parallel and distributed (GRID) computing, in database management and virtual observations, in applied mathematics and computer science, and build in the same time a strong theoretical knowledge in astrophysics. The consortium also studies the influence on the predictions of the resolution, the numerical codes, the self-consistent treatment of the baryons and of the physics included.

The main objectives are:

- the numerical study of galaxy formation in a cosmological framework using Grand Challenge applications;
- the development of advanced techniques in parallel computing and in applied math-

ematics to model galaxy formation and predict their observational signatures, as a function of physical parameters
- the gathering of renowned experts in computational astrophysics to share their software and expertise, and to optimize their access to national and international supercomputing facilities
- the delivery of a friendly access to state-of-the-art numerical simulations to the scientific community of both observers and theoreticians.

2. The simulations

2.1. The cosmological simulations

The cosmological simulations were performed on the initial conditions (ICs) described in Prunet et al. (2008). The ICs generation relied on MPGrafic and the associated software suite. The large Horizon ICs involve two boxes of 50 and 2000 Mpc h^{-1} comoving size with, respectively, 1024^3 and 4096^3 initial resolution elements (particles or grid points), following a ΛCDM concordance cosmology. They were generated from an existing set of initial conditions created for the MareNostrum simulation (Gottlöber & Yepes 2007): they share the same phases but with different box sizes and resolutions.

2.1.1. Horizon-Mare Nostrum.

The ICs with a 50 h^{-1} Mpc box and 1024^3 elements were used as inputs to the AMR code RAMSES (Teyssier 2002) in a simulation that included dark matter dynamics, hydrodynamics, star formation, metal enrichment of the gas, and feedback. The best resolution (i.e. the size of the smallest cell) is 1-2 kpc wide with 5 levels of refinement (leading to 4 billion AMR cells). This simulation directly compares to the MareNostrum simulation in terms of cosmology and physics, and we refer to it as the Horizon-MareNostrum simulation (Ocvirk et al. 2008). The simulation ran on the Mare Nostrum Supercomputer in Barcelona (Spain). The run took 4 weeks (from Sept. 2006 to Sept. 2007) for a total CPU time of 1.5 Mhr on 2048 processors (cf. Fig. 1).

2.1.2. 4Π Horizon.

The ICs with a 2000 h^{-1} Mpc box ($\approx$ 50% of the observable Universe) and 4096^3 elements (roughly 70 billions of particles) served as a starting point for the Horizon-4Π simulation (Teyssier et al. 2008). Horizon-4Π is a pure dark-matter simulation and assumes a cosmology constrained by the WMAP three-year results (cf. Fig. 2). We used a new French supercomputer BULL Novascale 3045 recently commissioned at CCRT (Centre de Calcul Recherche et Technologie, CEA). We ran RAMSES in pure N-body mode with 6144 processors for 2 months. Starting with a base grid of 4096^3 cells, we used 6 additional level of refinements for a formal resolution of 262144^3.

It is currently used to investigate the full-sky gravitational lensing signal that could be observed by the Dark Universe Explorer (hence the 4Π).

2.2. The smaller scale simulations

2.2.1. The GALaxy MERgers simulations (GalMer)

With the GalMer sub-project, a statistical line of attack has been chosen to the problem of galaxy mergers, by performing and analyzing thousands of simulations of interacting pairs, with initial conditions relevant to all epochs of the universe, at different redshifts.

di Matteo et al. (2007a) have investigated the enhancement of star formation efficiency in galaxy interactions and mergers by numerical simulations of several hundred galaxy collisions. All morphological types along the Hubble sequence have been considered in the ICs of the two colliding galaxies, with varying bulge-to-disk ratios and gas mass fractions. Different types of orbits have been simulated, direct and retrograde, according to the initial relative energy and impact parameter, and the resulting star formation history has been compared to that occurring in the two galaxies

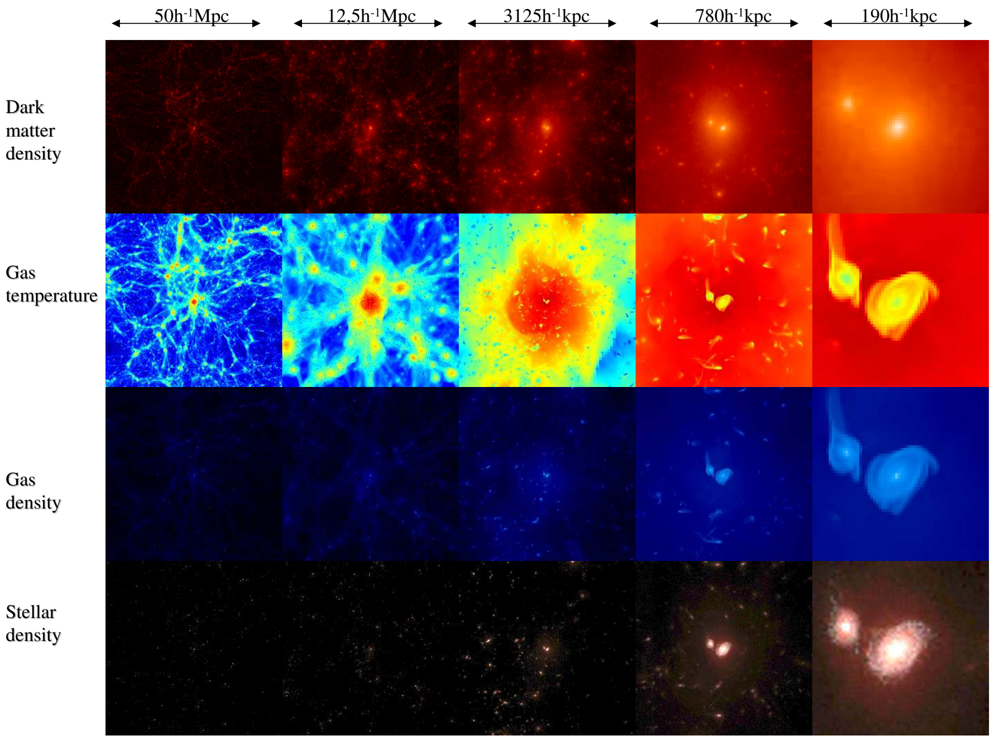

Fig. 1. Example of Horizon-Mare Nostrum simulation output. The field of view decreases from the left to the right.

when they are isolated. Based on their statistical results they derived general laws for the enhancement of star formation in galaxy interactions and mergers, as a function of the main parameters of the encounter.

2.2.2. Other simulations

A large number of small-scale simulations has also been performed. They include 'zoom'-simulations of regions of interest extracted from Dark Matter only simulations, simulations of cluster of galaxies, 'milky-way' like galaxies, 'idealized' and 'chemo-dynamical' simulation of isolated discs (e.g. Michel-Dansac & Wozniak 2006; Wozniak 2007).

2.3. The agenda

The Horizon agenda was the following:

- Jun 03: first discussions
- Sep 04: Kick-off meeting
- Oct 05: Selection by DEISA
- Oct 06: start of Mare Nostrum simulation
- Jul-Aug 07: Horizon 4Pi run
- Sep 07: end of Mare Nostrum simulation
- Dec 07: GALMER database opening
- Spring 09: GalICS database opening

3. On-line access to the simulations

Two main categories of data have been planned to be released for the astrophysical community: 1) the simulations of binary interacting galaxies and 2) the catalogs of galaxy properties in dark matter structures computed with

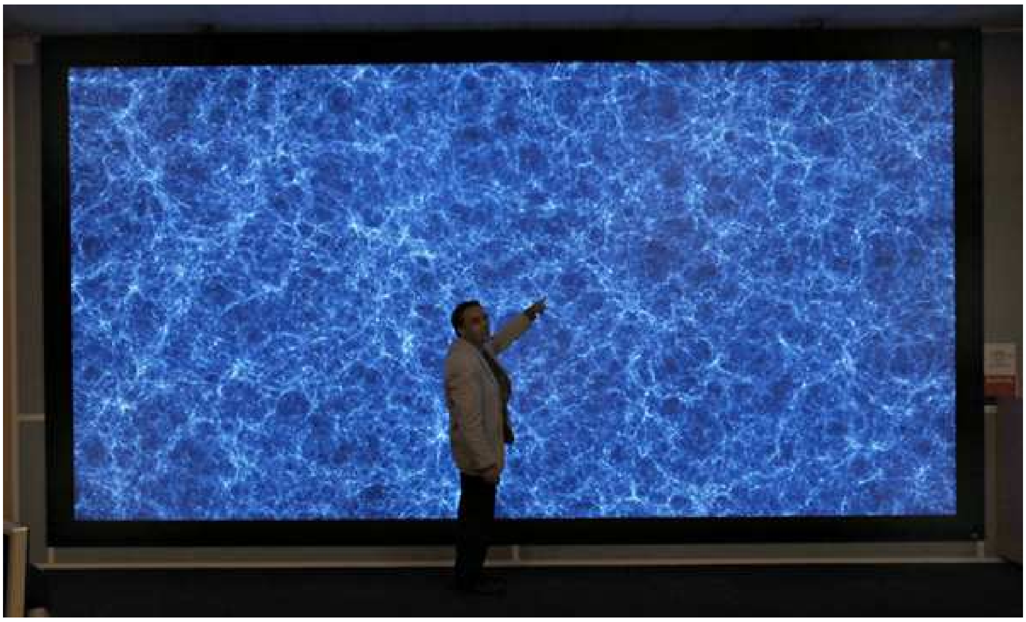

Fig. 2. Full 2000 h^{-1} Mpc projection of the 4Π simulation at redshift 0.

a semi-analytical modeling tool called GalICS (Hatton et al. 2003).

3.1. GalMer

The snapshots of the GalMer simulations are accessible on-line through the GalMer website (di Matteo et al. 2007b, http://galmer.obspm.fr, Fig. 3). All the necessary information (the metadata) has been stored in a database according to the VObs (Virtual Observatory) preliminary standards. GalMer provides the scientific community a free and user-friendly access to all the runs performed, together with the possibility to make on-the-fly analysis of the runs. In particular, it will be possible: to produce maps of gas, stars and dark matter, at different stages of the interactions, as seen from different directions; to trace the star formation histories and metal enrichment during galaxy encounters, and to compare these evolutions to those of the corresponding galaxies evolved in isolation; to compute spectra and colors for any region of the simulated galaxies, as seen from any direction.

GalMer server also provides the community with snapshots and metadata in standard FITS and VOTable format. The outputs can thus be sent to VObs visualization or analysis tools, like Aladin or Topcat, that are able to handle these formats (Fig. 4).

3.2. GalICS/MoMaF

The GalICS software package (for Galaxies in Cosmological Simulations) produces synthetic catalogs of galaxies within dark matter structures (see Devriendt 2007, for a review). These catalogs gather a large set of physical properties, as well as rest-frame absolute magnitudes through many filters, and rest-frame spectra.

The MoMaF software package (for Mock Map Facility, Fig. 5) is a facility that is able to make observing cones (Blaizot et al. 2005). The companion catalogs gather observer-frame apparent magnitudes through many filters, observer-frame apparent spectra, and mock images of wide/deep fields. The mock images can

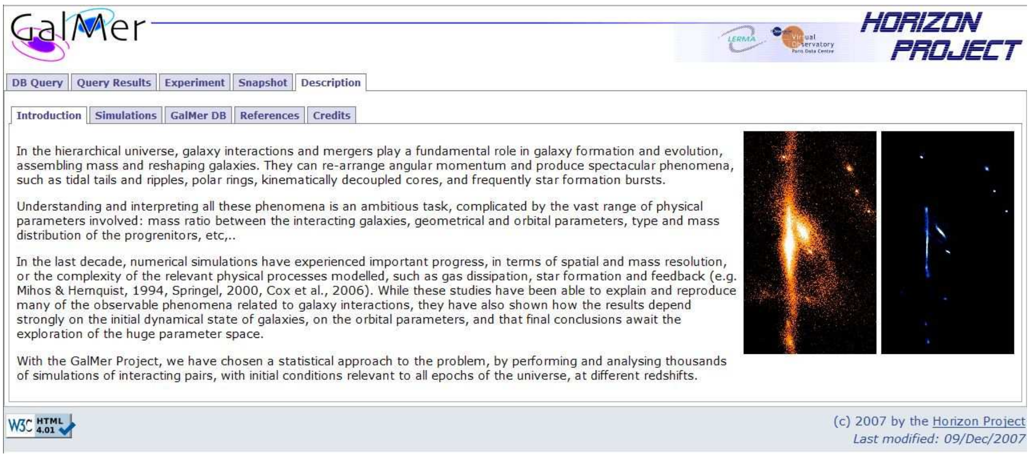

Fig. 3. GalMer web site opened in 2007 (http://galmer.obspm.fr)

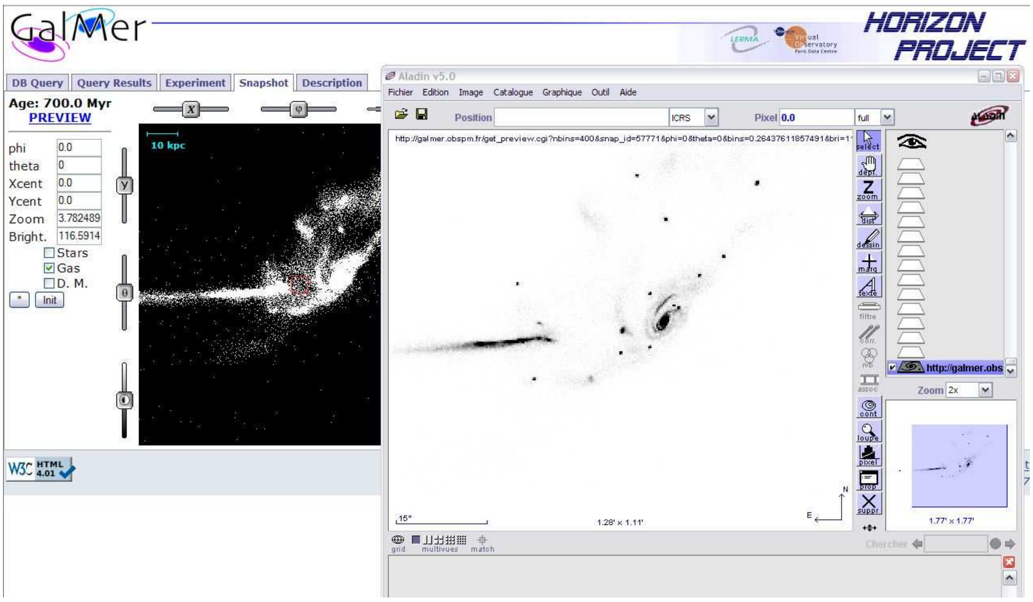

Fig. 4. GalMer provides files in VOTable and FITS formats that can be then used in VObs-compliant visualizer tools like Aladin, developed by the CDS, or TOPCAT, developed by AstroGrid.

then be observed with any specific combination of a telescope and detector.

The products of the GalICS/MoMaF packages can be used for different purposes:

- To get random samples of mock galaxies computed within hierarchical galaxy formation and compare them to data;
- To get mock images produced within hierarchical galaxy formation and process them with the same pipeline as actual data to compare model predictions with observations;
- To test data-processing pipelines with mock images where all properties of galaxies are known;

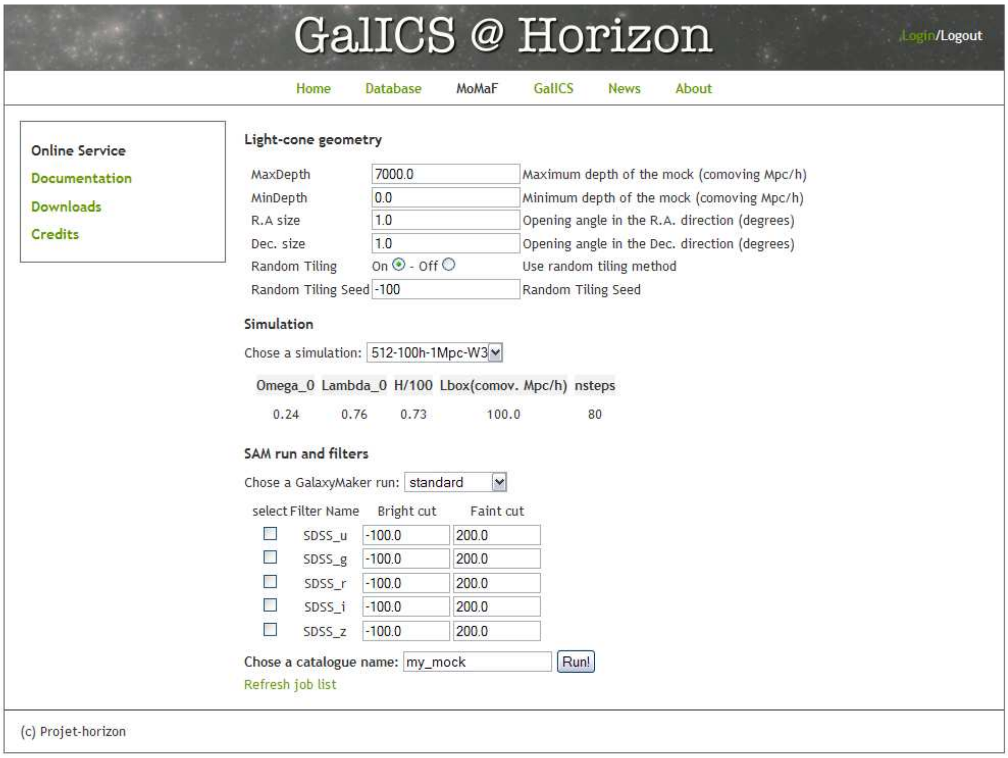

Fig. 5. A page of the Horizon/GalICS web site (under development). The user can select the geometrical properties of a light-cone, the dark-matter simulation and the input parameters of the semi-analytical modeling. The catalog of semi-analytical galaxy includes the photometric properties in a various filters that have been also selected by the user.

- To prepare observational strategies for forthcoming instruments;
- To get samples of synthetic spectra that are computed within hierarchical galaxy formation and compare them to observed spectra.

The future GalICS/MoMaF Database of Galaxies (http://horizon-vo.univ-lyon1.fr/GalICS) includes information on:

- the physical properties of galaxies,
- their location in dark matter structures,
- the merging history trees that link their progenitors and descendants,
- the rest-frame spectral energy distributions in the ultraviolet, optical, infrared and submillimeter wavelength ranges;
- and the absolute and apparent magnitudes in many photometric bands ranging from the ultraviolet to the submillimeter.

3.3. Future directions

The Horizon simulations form the basis for many other kinds of theoretical products: mock images of galaxies, 'deep' fields, light cones, etc. For instance, it can be imagined to implement a server giving access to photometrically calibrated images of galaxies extracted from the Horizon-Mare Nostrum simulation (cf. Fig. 6) or calibrated 'deep' fields. These mock images are very useful for defining the survey strategy and designing future complex instruments.

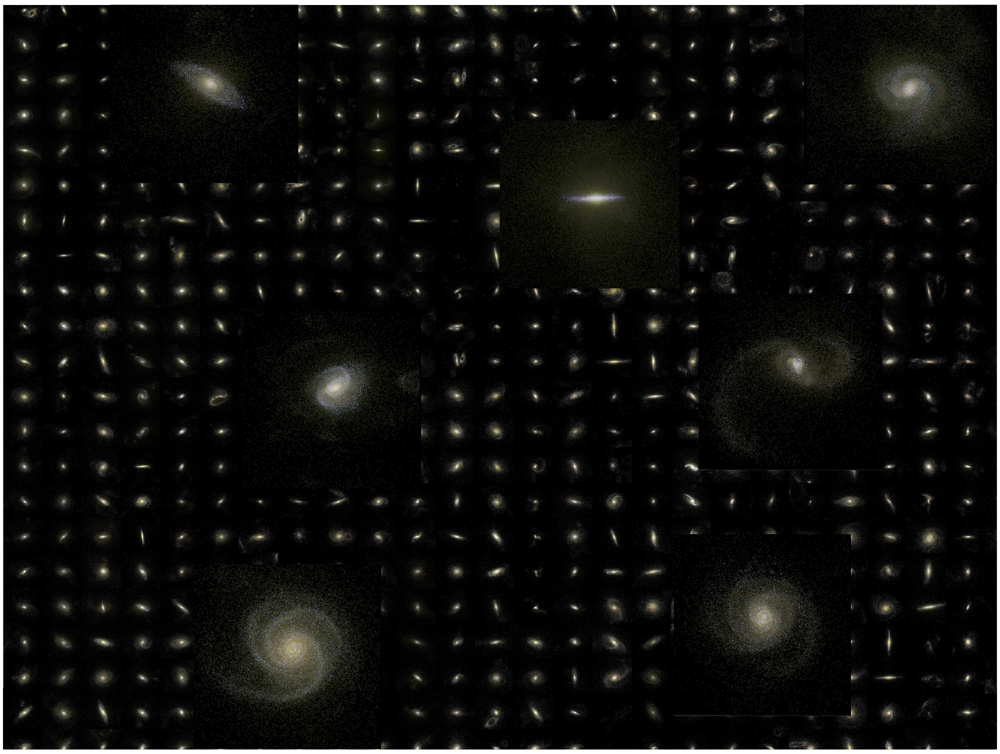

Fig. 6. Examples of galaxies extracted from the Horizon-Mare Nostrum simulation and calibrated in true colors (C. Pichon, http://www.projet-horizon.fr/rubrique38.html).

4. Discussion

There are a number of issues that have to be preliminary solved if we want to give a full access of the results of the Horizon numerical simulation.

4.1. Data access

- The ICs are stored in Grafic file format. The production of initial conditions deserves the development of specific codes. Not only the physical and algorithmic parameters have thus to be provided to the community but also the ICs files. They have to be accessible directly from an on-line service since a user might want to re-compute the simulation with his own code or set of parameters.
- The raw snapshots can be available for small simulations only since a full cosmological simulation is several tens of Tb of data. In the case of large-scale high resolution cosmological simulations, it is thus not reasonable to give a direct access to the snapshots.
- Post-processed data can be accessible in the form of catalogs thru database access. In the case of the cosmological Horizon simulations, only post-processed data will be available (e.g. GalICS/MoMaF). For smaller simulations, snapshots are available (e.g. GalMer).

4.2. Theory-theory comparisons

The same ICs have used by different codes (e.g. AMR vs SPH) allowing for detailed comparisons of various algorithms. But how to compare results? Correlation function, positions, phase-space,... ? When one wants to compare two simulations obtained with different codes, it could be useful to have on-line services that are able to handle the different formats and make the comparison with basic tools.

In the case of zoom-simulations or simulations using different spatial resolutions, it should be valuable to give access to all the results with links between the simulations related to the same part of the Universe.

4.3. Theory-observation interface

The production of observation-like products, or mock catalogs is not always an easy task. In some case, it can be an on-line service (e.g. GalMer on-the-fly spectral calibration) but also a time consuming post-process (e.g. GalICS semi-analytical modeling computations) that can be considered as another kind of simulation.

4.4. A lot of work (beyond the raw simulations...)

The distribution of final products via databases or web services is not a 'natural' path for numerical products. People involved in code development or running the codes of highly-parallel architecture are not specialist in database design or not familiar with web developments (using languages such as Java or PHP). It could quickly become a pain to develop a friendly interface, write the documentation and, overall, respect some standardized way to publish the simulations. This problem deserves not to be overlooked.

Acknowledgements. I am grateful to R. Teyssier, PI of the Horizon project, I. Chilingarian and J. Blaizot for providing me with helpful material and for fruitful discussions.

References

Blaizot, J., Wadadekar, Y., Guiderdoni, B.et al., 2005, MNRAS 360, 159
Cen, R., & Ostriker, J. P. 2000, ApJ, 538, 83
Devriendt, J.E.B., 2007, CRAL-2006. Chemodynamics: From First Stars to Local Galaxies E. Emsellem, H. Wozniak, G. Massacrier, J.-F. Gonzalez, J. Devriendt and N. Champavert (eds) EAS Publications Series, 24, 215-220
di Matteo, P., Combes, F., Melchior, A.-L., Semelin, B., 2007, A&A 468, 61
di Matteo, P., Combes, F., Chilingarian, I., et al., 2007, Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics held in Grenoble, France, July 2-6, 2007, Eds.: J. Bouvier, A. Chalabaev, and C. Charbonnel, p.322
Frenk, C. S., et al. 2000, preprint (astro-ph/0007362)
Gottlöber, S., & Yepes, G. 2007, ApJ, 664, 117
Hatton, S., Devriendt, J.E.G., Ninin, S., et al., 2003, MNRAS, 343, 75
Michel-Dansac, L., Wozniak, H., 2006 A&A 452, 97
Ocvirk, P., Pichon, C., & Teyssier, R. 2008, MNRAS, 390, 1326
Prunet, S., Pichon, C., Aubert, D., et al., 2008, ApJS, 178, 179
Springel, V., et al. 2005, Nature, 435, 629
Teyssier, R. 2002, A&A, 385, 337
Teyssier, R., et al. 2008, A&A, submitted
Wadsley, J. W., Stadel, J., & Quinn, T. 2004, NewA, 9, 137
Wozniak, H. 2007, A&A Letters, 465, L1