



## UvA-DARE (Digital Academic Repository)

### Towards distributed petascale computing

Hoekstra, A.G.; Portegies Zwart, S.; Bubak, M.; Sloot, P.M.A.

**Publication date**  
2007

**Published in**  
Petascale computing: Algorithms and applications

[Link to publication](#)

#### **Citation for published version (APA):**

Hoekstra, A. G., Portegies Zwart, S., Bubak, M., & Sloot, P. M. A. (2007). Towards distributed petascale computing. In D. A. Bader (Ed.), *Petascale computing: Algorithms and applications* (pp. 147-164). (Chapman & Hall/CRC Computational Science; No. 1). Chapman & Hall/CRC. [http://arxiv.org/PS\\_cache/astro-ph/pdf/0703/0703485v1.pdf](http://arxiv.org/PS_cache/astro-ph/pdf/0703/0703485v1.pdf)

#### **General rights**

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

#### **Disclaimer/Complaints regulations**

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

---

## Contents

<b>1 Towards Distributed Petascale Computing</b>	<b>3</b>
<i>Alfons G. Hoekstra, Simon Portegies Zwart, Marian Bubak, and Peter M.A. Sloot</i>	
1.1 Introduction . . . . .	3
1.2 Grid Computing . . . . .	5
1.3 Petascale Computing on the Grid . . . . .	6
1.4 The Virtual Galaxy . . . . .	8
1.4.1 A Multi-Physics model of the Galaxy . . . . .	8
1.4.2 A performance model for simulating the Galaxy . . . . .	11
1.4.3 Petascale simulation of a Virtual Galaxy . . . . .	13
1.5 Discussion and Conclusions . . . . .	15
<b>References</b>	<b>17</b>



# Chapter 1

---

## *Towards Distributed Petascale Computing*

**Alfons G. Hoekstra**

*Section Computational Science, University of Amsterdam, Kruislaan 403, Amsterdam, The Netherlands, alfons@science.uva.nl*

**Simon Portegies Zwart**

*Section Computational Science and Astronomical Institute “Anton Pannekoek”, University of Amsterdam, Kruislaan 403, Amsterdam, The Netherlands, spz@science.uva.nl*

**Marian Bubak**

*Section Computational Science, University of Amsterdam, Kruislaan 403, Amsterdam, The Netherlands, and AGH University of Science and Technology, Kraków, Poland, bubak@science.uva.nl*

**Peter M.A. Sloot**

*Section Computational Science, University of Amsterdam, Kruislaan 403, Amsterdam, The Netherlands, sloot@science.uva.nl*

1.1 Introduction .....	3
1.2 Grid Computing .....	5
1.3 Petascale Computing on the Grid .....	6
1.4 The Virtual Galaxy .....	8
1.5 Discussion and Conclusions .....	15

---

### 1.1 Introduction

Recent advances in experimental techniques have opened up new windows into physical and biological processes on many levels of detail. The resulting data explosion requires sophisticated techniques, such as grid computing and collaborative virtual laboratories, to register, transport, store, manipulate, and share the data. The complete cascade from the individual components to the fully integrated multi-science systems crosses many orders of magnitude in temporal and spatial scales. The challenge is to study not only the fundamental processes on all these separate scales, but also their mutual coupling

through the scales in the overall multi-scale system, and the resulting emergent properties. These complex systems display endless signatures of order, disorder, self-organization and self-annihilation. Understanding, quantifying and handling this complexity is one of the biggest scientific challenges of our time [1].

In this chapter we will argue that studying such multi-scale multi-science systems gives rise to inherently hybrid models containing many different algorithms best serviced by different types of computing environments (ranging from massively parallel computers, via large-scale special purpose machines to clusters of PC's) whose total integrated computing capacity can easily reach the PFlop/s scale. Such hybrid models, in combination with the by now inherently distributed nature of the data on which the models 'feed' suggest a distributed computing model, where parts of the multi-scale multi-science model are executed on the most suitable computing environment, and/or where the computations are carried out close to the required data (i.e. bring the computations to the data instead of the other way around).

Prototypical examples of multi-scale multi-science systems come from biomedicine, where we have data from virtually all levels between 'molecule and man' and yet we have no models where we can study these processes as a whole. The complete cascade from the genome, proteome, metabolome, physiome to health constitutes multi-scale, multi-science systems, and crosses many orders of magnitude in temporal and spatial scales [2, 3]. Studying biological modules, their design principles, and their mutual interactions, through an interplay between experiments and modeling and simulations, should lead to an understanding of biological function and to a prediction of the effects of perturbations (e.g. genetic mutations or presence of drugs). [4]

A good example of the power of this approach, in combination with state-of-the-art computing environments, is provided by the study of the heart physiology, where a true multi-scale simulation, going from genes, to cardiac cells, to the biomechanics of the whole organ, is now feasible. [5] This 'from genes to health' is also the vision of the Physiome project [6, 7], and the Viro-Lab [8, 9], where a multi-scale modeling and simulation of human physiology is the ultimate goal. The wealth of data now available from many years of clinical, epidemiological research and (medical) informatics, advances in high-throughput genomics and bioinformatics, coupled with recent developments in computational modeling and simulation, provides an excellent position to take the next steps towards understanding the physiology of the human body across the relevant  $10^9$  range of spatial scales (nm to m) and  $10^{15}$  range of temporal scales, ( $\mu$ s to human lifetime) and to apply this understanding to the clinic. [6, 10] Examples of multi-scale modeling are increasingly emerging (see for example, [11, 12, 13, 14]).

In Section 1.2 we will consider the Grid as the obvious choice for a distributed computing framework, and we will then explore the potential of computational grids for Petascale computing in Section 1.3. Section 1.4 presents the *Virtual Galaxy* as a typical example of a multi-scale multi-physics appli-

cation, requiring distributed Petaflop/s computational power.

---

## 1.2 Grid Computing

The radical increase in the amount of IT-generated data from physical, living and social systems brings about new challenges related to the sheer size of data. It was this data ‘deluge’ that originally triggered the research into grid computing [15, 16]. Grid computing is an emerging computing model that provides the ability to share data and instruments and to perform high throughput computing by taking advantage of many networked computers able to divide process execution across a distributed infrastructure.

As the Grid is ever more frequently used for collaborative problem solving in research and science, the real challenge is in the development of new applications for a new kind of users through virtual organizations. Existing grid programming models are discussed in [17, 18].

Workflow is a convenient way of distribution of computations across a grid. A large group of composition languages have been studied for formal description of workflows [19] and they are used for orchestration, instantiation, and execution of workflows [20]. Collaborative applications are also supported by problem solving environments which enable users to handle application complexity with web-accessible portals for sharing software, data, and other resources [21]. Systematic ways to building grid applications are provided through object-oriented and component technology, for instance the Common Component Architecture which combines the IDL-based distributed framework concept with requirements of scientific applications [22]. Some recent experiments with computing across grid boundaries, workflow composition of Grid services with semantic description, and development of collaborative problem solving environments are reported in [23, 24, 25]. These new computational approaches should transparently exploit the dynamic nature of Grid and virtualization of grid infrastructure. The challenges are efficient usage of knowledge for automatic composition of applications [26].

Allen et al. in [27] distinguish four main types of grid applications: (1) Community-centric; (2) Data-centric; (3) Computation-centric; and (4) Interaction-centric. Data-centric applications are, and will continue to be the main driving force behind the Grid. Community-centric applications are about bringing people or communities together, as e.g. in the Access Grid, or in distributed collaborative engineering. Interaction-centric applications are those that require ‘a man in the loop’, for instance in real-time computational steering of simulations or visualizations (as e.g. demonstrated by the CrossGrid project [25]).

In this chapter we focus on Computation-centric applications. These are

the traditional High Performance Computing (HPC) and High Throughput Computing (HTC) applications which, according to Allen et al. [27] “turned to parallel computing to overcome the limitations of a single processor, and many of them will turn to Grid computing to overcome the limitations of a parallel computer.” In the case of parameter sweep (i.e. HTC) applications this has already happened. Several groups have demonstrated successful parameter sweeps on a computational Grid (see e.g. [28]). For tightly coupled HPC applications this is not so clear, as common wisdom is that running a tightly coupled parallel application in a computational grid (in other words, a parallel job actually running on several parallel machines that communicate with each other in a Grid) is of no general use because of the large overheads that will be induced by communications between computing elements (see e.g. [17]). However, in our opinion this certainly is a viable option, provided the granularity of the computation is large enough to overcome the admittedly large communication latencies that exist between compute elements in a Grid. [29] For PFlop/s scale computing we can assume that such required large granularity will be reached. Recently a Computation-centric application running in parallel on compute elements located in Poland, Cyprus, Portugal, and the Netherlands was successfully demonstrated [30, 31].

---

### 1.3 Petascale Computing on the Grid

Execution of multi-scale multi-science models on computational grids will in general involve a diversity of computing paradigms. On the highest level functional decompositions may be performed, splitting the model in sub-models that may involve different types of physics. For instance, in a fluid-structure interaction application the functional decomposition leads to one part modeling the structural mechanics, and another part modeling the fluid flow. In this example the models are tightly coupled and exchange detailed information (typically, boundary conditions at each time step). On a lower level one may again find a functional decomposition, but at some point one encounters single-scale, single-physics sub-models, that can be considered as the basic units of the multi-scale multi-science model. For instance, in a multi-scale model for crack propagation, the basic units are continuum mechanics at the macroscale, modeled with finite elements, and molecular dynamics at the microscale [32]. Another example is provided by Plasma Enhanced Vapor Deposition where mutually coupled chemical, plasma physical and mechanical models can be distinguished [33]. In principle all basic modeling units can be executed on a single (parallel) computer, but they can also be distributed to several machines in a computational grid.

These basic model units will be large scale simulations by themselves. With

an overall performance on the PFlop/s scale, it is clear that the basic units will also be running at impressive speeds. It is difficult to estimate the number of such basic model units. In the example of the fluid-structure interaction, there are two, running concurrently. However, in case of for instance a multi-scale system modeled with the Heterogeneous Multiscale Method [34] there could be millions of instances of a microscopic model that in principle can execute concurrently (one on each macroscopic grid point). So, for the basic model units we will find anything between single processor execution and massively parallel computations.

A computational grid offers many options of mapping the computations to computational resources. First, the basic model units can be mapped to the most suitable resources. So, a parallel solver may be mapped to massively parallel computers, whereas for other solvers special purpose hardware may be available, or just single PC's in a cluster. Next, a distributed simulation system is required to orchestrate the execution of the multi-scale multi-science models.

A computational grid is an appropriate environment for running functionally decomposed distributed applications. A good example of research and development in this area is the CrossGrid Project which aimed at elaboration of an unified approach to development and running large scale interactive distributed, compute- and data-intensive applications, like biomedical simulation and visualization for vascular surgical procedures, a flooding crisis team decision support system, distributed data analysis in high energy physics, and air pollution combined with weather forecasting [25]. The following issues were of key importance in this research and will also play a pivotal role on the road towards distributed PFlop/s scale computing on the Grid: porting applications to the grid environment; development of user interaction services for interactive startup of applications, online output control, parameter study in the cascade, and runtime steering, and on-line, interactive performance analysis based on-line monitoring of grid applications. The elaborated CrossGrid architecture consists of a set of self-contained subsystems divided into layers of applications, software development tools and Grid services [35].

Large scale grid applications require on-line performance analysis. The application monitoring system, OCM-G, is a unique online monitoring system in which requests and response events are generated dynamically and can be toggled at runtime. This imposes much less overhead on the application and therefore can provide more accurate measurements for the performance analysis tool like G-PM, which can display (in form of various metrics) the behavior of Grid applications [36].

The High Level Architecture (HLA) fulfills many requirements of distributed interactive applications. HLA and the Grid may complement each other to support distributed interactive simulations. The G-HLAM system supports for execution of legacy HLA federates on the Grid without imposing major modifications of applications. To achieve efficient execution of HLA-based simulations on the Grid, we introduced migration and monitoring mechanisms



for such applications. This system has been applied to run two complex distributed interactive applications: N-body simulation and virtual bypass surgery [37].

In the next section we explore in some detail a prototypical application where all the aforementioned aspects need to be addressed to obtain distributed Petascale computing.

---

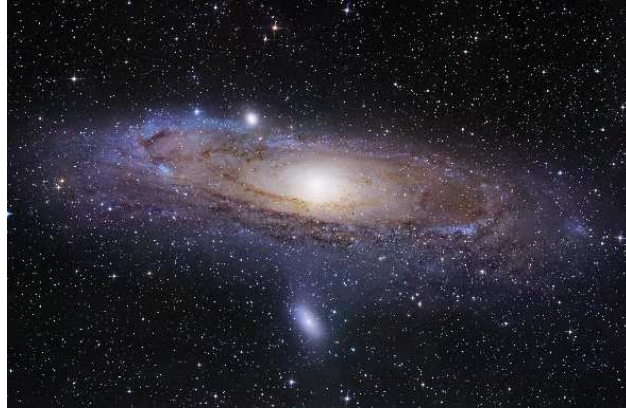
## 1.4 The Virtual Galaxy

A grand challenge in computational astrophysics, requiring *at least* the PFlop/s scale, is the simulation of the physics of formation and evolution of large spiral galaxies like the Milky-way. This requires the development of a hybrid simulation environment to cope with the multiple time scales, the broad range of physics and the sheer number of simulation operations [38, 39]. The nearby grand design spiral galaxy M31 in the constellation andromeda, as displayed in Fig. 1.1, provides an excellent birdseye view of how the Milky-way probably looks.

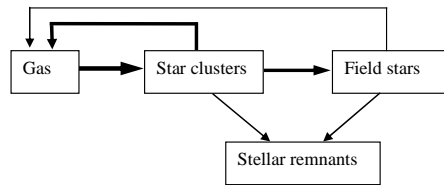
This section presents the Virtual Galaxy as a typical example of a multi-physics application that requires PFlop/s computational speeds, and has all the right properties to be mapped to distributed computing resources. We will introduce in some detail the relevant physics and the expected amount of computations (i.e. Flop) needed to simulate a Virtual Galaxy. Solving Newton's equations of motion for any number of stars is a challenge by itself, but to perform this in an environment with the number of stars as in the Galaxy, and over the enormous range of density contrasts and with the inclusion of additional chemical and nuclear physics, doesn't make the task easier. No single computer will be able to perform the resulting multitude of computations, and therefore it provides an excellent example for a hybrid simulation environment containing a wide variety of distributed hardware. We end this section with a discussion on how a Virtual Galaxy simulation could be mapped to a PFlop/s scale grid computing environment. We believe that the scenarios that we outline are prototypical and also apply to a multitude of other multi-science multi-scale systems, like the ones that were discussed section 1.1 and 1.3.

### 1.4.1 A Multi-Physics model of the Galaxy

The Galaxy today contains a few times  $10^{11}$  the solar mass ( $M_{\odot}$ ) in gas and stars. The life cycle of the gas in the Galaxy is illustrated in Fig. 1.2, where we show how gas transforms to star clusters, which again dissolve to individual stars. The ingredients for a self consistent model of the Milky-way Galaxy



**FIGURE 1.1:** The Andromeda Nebula, M31. A mosaic of hundreds of Earth based telescope pointings were needed to make this image.



**FIGURE 1.2:** Schematic representation of the evolution of the gas content of the Galaxy.

is based on these same three ingredients: the gas, the star clusters and the field stellar population. The computational cost and physical complexity for simulating each of these ingredients can be estimated based on the adopted algorithms.

#### 1.4.1.1 How gas turns into star clusters

Stars and star clusters form from giant molecular clouds which collapse when they become dynamically unstable. The formation of stars and star clusters is coupled with the galaxy formation process. The formation of star clusters themselves has been addressed by many research teams and most of the calculations in this regard are a technical endeavor which is mainly limited by the lack of resources.

Simulations of the evolution of a molecular cloud up to the moment it forms stars are generally performed with adaptive mesh refinement and smoothed particles hydrodynamics algorithms. These simulations are complex, and some calculations include turbulent motion of the gas [40], solve the full mag-

netic hydrodynamic equations [41, 42], or include radiative transport [43]. All the currently performed dynamical cloud collapse simulations are computed with a relatively limited accuracy in the gravitational dynamics. We adopt the smoothed particle hydrodynamics methodology to calculate the gravitational collapse of a molecular cloud, as it is relatively simple to implement and has scalable numerical complexity. These simulation environments are generally based on the Barnes-Hut tree code [44] for resolving the self gravity between the gas or dust volume or mass elements, and have a  $\mathcal{O}(n_{\text{SPH}} \log n_{\text{SPH}})$  time complexity [45].

Simulating the collapse of a molecular cloud requires at least  $\sim 10^3$  SPH particles per star, a star cluster that eventually (after the simulation) consisting of  $\mathcal{O}(10^4)$  stars then requires about  $n_{\text{SPH}} \sim 10^7$  SPH particles.

The collapse of a molecular cloud lasts for about  $\tau_J \simeq 1/\sqrt{G\rho}$ , which for a  $10^4 M_\odot$  molecular cloud with a size of 10 pc is about a million years. Within this time span the molecular cloud will have experienced roughly  $10^4$  dynamical time scales totaling the CPU requirements to about  $\mathcal{O}(10^{11})$  Flop for calculating the gravitational collapse of one molecular cloud.

#### 1.4.1.2 The evolution of the individual stars

Once most of the gas is cleared from the cluster environment, an epoch of rather clean dynamical evolution mixed with the evolution of single stars and binaries starts. In general, star cluster evolution in this phase may be characterized by a competition between stellar dynamics and stellar evolution. Here we focus mainly on the nuclear evolution of the stars.

With the development of shell based Henyé codes [46] the nuclear evolution of a single star for its entire lifetime requires about  $10^9$  Flop [47]. Due to efficient step size refinement the performance of the algorithm is independent of the lifetime of the star; a  $100 M_\odot$  star is as expensive in terms of compute time as a  $1 M_\odot$  star. Adopting the mass distribution with which stars are born [48] about one in 6 stars require a complete evolutionary calculation. The total compute time for evolving all the stars in the Galaxy over its full life time then turns out to be about  $10^{20}$  Flop.

Most ( $\gtrsim 99\%$ ) of all the stars in the Galaxy will not do much apart from burning their internal fuel. To reduce the cost of stellar evolution we can therefore parameterize the evolution of such stars. Excellent stellar evolution prescriptions at a fraction of the cost ( $\lesssim 10^4$  Flop) are available [49, 50], and could be used for the majority of stars (which is also what we adopted in § 1.4.2).

#### 1.4.1.3 Dynamical evolution

When a giant molecular cloud collapses one is left with a conglomeration of bound stars and some residual gas. The latter is blown away from the cluster by the stellar winds and supernovae of the young stars. The remaining gas

depleted cluster may subsequently dissolve in the background on a time scale of about  $10^8$  years.

The majority (50-90%) of star clusters which are formed in the Galaxy dissolve due to the expulsion of the residual gas [51, 52]. Recent reanalysis of the cluster population of the Large Magellanic cloud indicates that this process of *infant mortality* is independent of the mass of the cluster [53]. Star clusters that survive their infancy engage in a complicated dynamical evolution which is quite intricately coupled with the nuclear evolution of the stars [54].

The dynamical evolution of a star cluster is best simulated using direct  $N$ -body integration techniques, like NBODY4 [55, 56] or the `starlab` software environment [54].

For dense star clusters the compute time is completely dominated by the force evaluation. Since each star has a gravitational pull at all other stars this operation scales with  $\mathcal{O}(N^2)$  for one dynamical time step. The good news is that the large density contrast between the cluster central regions and its outskirts can cover 9 orders of magnitude, and stars far from the cluster center are regularly moving whereas central stars have less regular orbits [57]. By applying smart time stepping algorithms one can reduce the  $\mathcal{O}(N^2)$  to  $\mathcal{O}(N^{4/3})$  without loss of accuracy [58]. In fact one actually gains accuracy since taking many unnecessary small steps for a regularly integrable star suffers from numerical round-off.

The GRAPE-6, a special purpose computer for gravitational  $N$ -body simulations, performs dynamical evolution simulations at a peak speed of about 64 Tflop/s [59], and is extremely suitable for large scale  $N$ -body simulations.

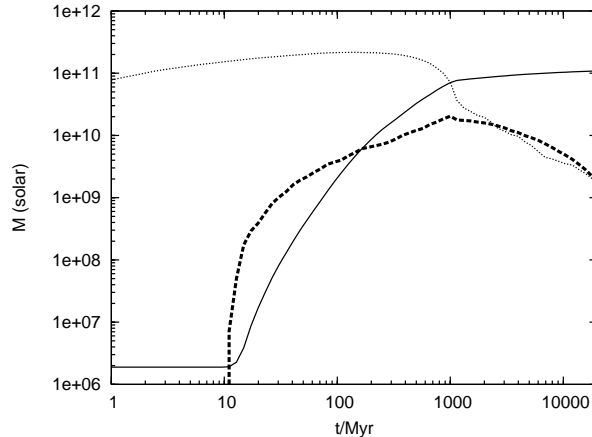
#### 1.4.1.4 The galactic field stars

Stars that are liberated by star clusters become part of the Galactic tidal field. These stars, like the Sun, orbit the Galactic center in regular orbits. The average time scale for one orbital revolution for a field star is about 250 Myr. These regularly orbiting stars can be resolved dynamically using a relatively unprecise  $N$ -body technique, we adopt here the  $\mathcal{O}(N)$  integration algorithm which we introduced in § 1.4.1.1.

In order to resolve a stellar orbit in the Galactic potential about 100 integration time steps are needed. Per Galactic crossing time (250 Myr) this code then requires about  $10^6$  operations per star, resulting in a few times  $10^7 N$  Flop for simulating the field population. Note that simulating the galactic field population is a trivially parallel operation, as the stars hover around in their self generated potential

### 1.4.2 A performance model for simulating the Galaxy

Next we describe the required computer resources as a function of life time of a Virtual Galaxy. The model is relatively simple and the embedded physics is only approximate, but it will give an indication on what type of calculation



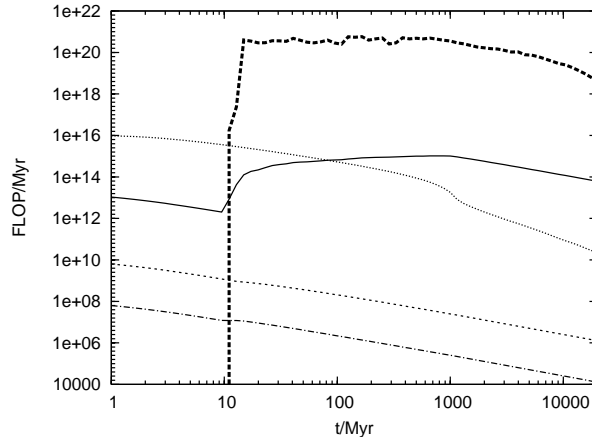
**FIGURE 1.3:** The evolution of the mass content in the Galaxy via the simple model described in §1.4.2. The dotted curve give the total mass in giant molecular clouds, the thick dashed curve in star clusters and the solid curve in field stars, which come from dissolved star clusters.

is most relevant in what state of the evolution of the Galaxy.

According to the model we start the evolution of the Galaxy with amorphous gas. We subsequently assume that molecular clouds are formed with power-law mass function with an index of -2 between  $10^3 M_\odot$  and  $10^7 M_\odot$ , with distribution in time which is flat in  $\log t$ . We assume that the molecular cloud lives for between 10 Myr and 1 Gyr (with an equal probability between these moments). The star formation efficiency is 50%, and the cluster has an 80% chance to dissolve within 100 Myr (irrespective of the cluster mass). The other 20% clusters dissolve on a time scale of about  $t_{\text{diss}} \sim 10\sqrt{R^3 M}$  Myr. During this period they lose mass at a constant rate. The field population is enriched with the same amount of mass.

The resulting total mass in molecular clouds, star clusters and field stars is presented in fig. 1.3. At early age, the galaxy completely consists of molecular clouds. After about 10 Myr some of these clouds collapse to form star clusters and single stars, indicated by the rapidly rising solid (field stars) and dashed (star clusters) curves. The maximum number of star clusters is reached when the Galaxy is about a Gyr old. The field population continues to rise to reach a value of a few times  $10^{11} M_\odot$  at today's age of about 10 Gyr. By that time the total mass in star clusters has dropped to several  $10^9 M_\odot$  quite comparable with the observed masses of the field population and the star cluster content.

In Fig. 1.4 we show the evolution of the amount of Flop required to simulate the entire galaxy, as a function of its life time. The Flop count along the vertical axis are given in units of number of floating points operations per million



**FIGURE 1.4:** The number of floating points operations expenditure per million years for the various ingredients in the performance model. The solid, thick short dash and dotted curve are as in Fig. 1.3. New in this figure are the two dotted and dash-dotted lines near the bottom, which represent the CPU time needed for evolving the field star population (lower dotted curve) and dark matter (bottom curve).

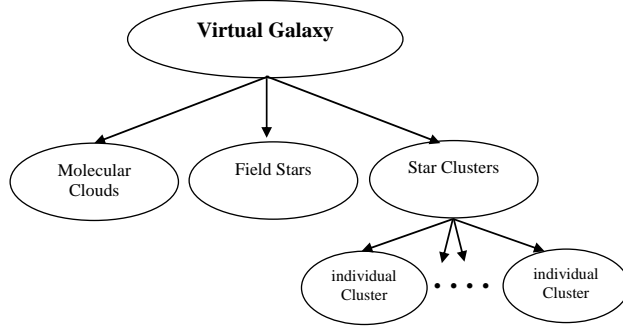
years in Galactic evolution. For example, to evolve the Galaxy's population of molecular clouds from 1000 Myr to 1001 Myr requires about  $10^{16}$  Flop.

### 1.4.3 Petascale simulation of a Virtual Galaxy

From Fig. 1.4 we see that the most expensive submodels in a Virtual Galaxy are the star cluster simulations, the molecular cloud simulations, and the field star simulations. In the following discussion we neglect the other components. A Virtual Galaxy model, viewed as a multi-scale multi-physics model, can then be decomposed as in Fig. 1.5.

The by far most expensive operation is the star cluster computations. We have  $O(10^4)$  star clusters, each cluster can be simulated independent of the others. This means that a further decomposition is possible, down to the individual cluster level. A single star cluster simulation, containing  $O(10^4)$  stars, still requires computational speeds at the TFlop/s scale (see also below). The clusters simulations require  $10^{21}$  Flop per simulated Myr of lifetime of the Galaxy. The molecular clouds plus the field stars need, on average over the full life time of the Galaxy,  $10^{15}$  Flop per simulated Myr of lifetime, and can be executed on general purpose parallel machines.

A distributed Petascale computing infrastructure for the Virtual Galaxy could consist of a single or two general purpose parallel machines to execute the molecular clouds and fields stars at a sustained performance of 1 TFlop/s,

**FIGURE 1.5:** Functional decomposition of the Virtual Galaxy**TABLE 1.1:** Estimated run times of the Virtual Galaxy simulation on a distributed Petascale architecture as described in the main text.

Age	Milky Way Galaxy	Factor 10 reduction	Dwarf Galaxy (factor 100 reduction)
10 Myr	3 hour	17 min.	2 min.
100 Myr	3 year	104 days	10 days
1 Gyr	31 year	3 year	115 days
10 Gyr	320 year	32 year	3 year

and a distributed grid of special purpose Grapes to simulate the star clusters. We envision for instance 100 next generation GrapeDR systems<sup>1</sup>, each delivering 10 Tflop/s, providing a sustained 1 PFlop/s for the star cluster computations. We can now estimate the expected runtime on a Virtual Galaxy simulation on this infrastructure. In Table 1.1 we present the estimated wall-clock time needed for simulating the Milky-way Galaxy, a smaller subset and a dwarf galaxy using the distributed Petascale resource described above. Note that in the reduced Galaxies the execution time goes linearly down with the reduction factor, which should be understood as a reduction of mass in the molecular clouds and a reduction of the total number of star clusters (but with the same amount of stars per star cluster).

With such a performance it will be possible to simulate the entire Milky-way Galaxy for about 10 Myr which is an interesting time scale on which stars form, massive stars evolve and infant mortality of young newly born star clusters operates. By simulating the entire Milky-way Galaxy on this important time scale will enable us to study these phenomena with unprecedented detail.

At the same performance it will be possible to simulate part (1/10th) of the Galaxy on a time scale of 100 Myr. This time scale is important for the

<sup>1</sup>Currently some 100 Grape6 systems, delivering an average performance of 100 GFlop/s are deployed all over the world.

evolution of young and dense star clusters, the major star formation mode in the Galaxy.

Simulating a dwarf galaxy, like the Large Magellanic Cloud for its entire lifetime will become possible with a PFlop/s scale distributed computer. The entire physiology of this galaxy is largely not understood, as well as the intricate coupling between stellar dynamics, gas dynamics, stellar evolution and dark matter.

---

## 1.5 Discussion and Conclusions

Multi-scale multi-science modeling is the next (grand) challenge in Computational Science. Not only in terms of formulating the required couplings across the scales or between multi-science models, but also in terms of the sheer computational complexity of such models. The later can easily result in requirements on the PFlop/s scale.

We have argued that simulating these models involves high level functional decompositions, finally resulting in some collection of single-scale single-science sub-models, that by themselves could be quite large, requiring simulations on e.g. massively parallel computers. In other words, the single-scale single-science sub-models would typically involve some form of High Performance - or High Throughput Computing. Moreover, they may have quite different demands for compute infrastructure, ranging from Supercomputers, via special purpose machines, to the single workstation. We have illustrated this by pointing to a few models from biomedicine and in more detail in the discussion on the Virtual Galaxy.

We believe that the Grid provides the natural distributed computing environment for such functionally decomposed models. The Grid has reached a stage of maturity that in essence all the necessary ingredients needed to develop a PFlop/s scale computational grid for multi-scale multi-science simulations are available. Moreover, in a number of projects grid enabled functionally decomposed distributed computing has been successfully demonstrated, using many of the tools that were discussed in Section 1.2.

Despite these successes the experience with computational grids is still relatively small. Therefore, a real challenge lies ahead in actually demonstrating the feasibility of Grids for distributed Petascale computing, and realizing Grid-enabled Problem Solving Environments for multi-scale multi-science applications.





---

## References

- [1] A. Barabasi, Taming Complexity, *Nature Physics* **1**, 68–70, 2005.
- [2] A. Finkelstein, J. Hetherington, L. Li, O. Margoninski, P. Saffrey, R. Seymour and A. Warner, Computational Challenges of System Biology, *Computer* **37**, 26–33, 2004.
- [3] P.M.A. Sloot, D. Frenkel, H.A. Van der Vorst, A. van Kampen, H. Bal, P. Klint, R.M.M. Mattheij, J. van Wijk, J. Schaye, H.-J. Langevelde, R.H. Bisseling, B. Smit, E. Valenteyn, H. Sips, J.B.T.M. Roerdink, and K.G. Langedoen, White paper on Computational e-Science: Studying complex systems in silico, A National Research Initiative, December 2006.
- [4] B. Di Ventura, C. Lemerle, K. Michalodimitrakis, and L. Serrano, From in vivo to in silico biology and back, *Nature* **443**, 527–533, 2006.
- [5] D. Noble, Modeling the Heart – from Genes to Cells to the Whole Organ, *Science* **295**, 1678–1682, 2002.
- [6] P.J. Hunter and T.K. Borg, Integration from Proteins to Organs: The Physiome Project, *Nature Reviews Molecular and Cell Biology* **4**, 237–243, 2003.
- [7] P.J. Hunter, W.W. Li, A.D. McCulloch, and D. Noble, Multiscale Modeling: Physiome Project Standards, Tools, and Databases, *Computer* **39**, 48–54, 2006.
- [8] Virolab project – <http://www.virolab.org>
- [9] P.M.A. Sloot, A. Tirado-Ramos, I. Altintas, M.T. Bubak, and C.A. Boucher, From Molecule to Man: Decision Support in Individualized E-Health, *IEEE Computer* **39**, 40–46, 2006.
- [10] N. Ayache, J.-P. Boissel, S. Brunak, G. Clapworthy, J. Fingberg, G. Lonsdale, A. Frangi, G. Deco, P. Hunter, P. Nielsen, M. Halstead, R. Hose, I. Magnin, F. Martin-Sanchez, P. Sloot, J. Kaandorp, A. Hoekstra, S. van Sint Jan, M. Viceconti M, Towards Virtual Physiological Human: Multi-level modelling and simulation of the human anatomy and physiology. White paper, edited by S. Norager, I. Lakovidis, M. Carbrera, and M. Ozcivelek.

- [http://ec.europa.eu/information\\_society/activities/health/docs/events/barcelona2005/ec-vph-white-paper2005nov.pdf](http://ec.europa.eu/information_society/activities/health/docs/events/barcelona2005/ec-vph-white-paper2005nov.pdf), 2005
- [11] P.F. Davies, J.A. Spaan, R. Krams, Shear stress biology of the endothelium, *Ann Biomed Eng.* **33**, 1714–1718, 2005.
  - [12] G. Iribe, P. Kohl, D. Noble, odulatory effect of calmodulin-dependent kinase II (CaMKII) on sarcoplasmic reticulum Ca<sup>2+</sup> handling and interval-force relations: a modelling study, *Philos Transact A Math Phys Eng Sci.* **364**, 1107–1133, 2006.
  - [13] D. Kelly, L. Mackenzie, P. Hunter, B. Smaill, D.A. Saint, Gene expression of stretch-activated channels and mechanoelectric feedback in the heart, *Clin Exp Pharmacol Physiol.* **33**, 642–648, 2006.
  - [14] P.M.A. Sloot, A.V. Boukhanovsky, W. Keulen, A. Tirado-Ramos, C.A. Boucher A Grid-based HIV expert system, *J Clin Monit Comput.* **19**, 263–278, 2005.
  - [15] I. Foster, C. Kesselman, and S. Tuecke, The anatomy of the grid: Enabling scalable virtual organizations, *International Journal of High Performance Computing Applications* **15**, 200–222, 2001.
  - [16] A.J.G. Hey and A.E. Trefethen, The Data Deluge: An e-Science Perspective, in F. Berman, G. Fox, and A.J.G. Hey (Eds.), *Grid Computing, Making the Global Infrastructure a Reality*, chapter 36, 2003.
  - [17] C. Lee, D. Talia, Grid Programming Models: Current Tools, Issues, and Directions, in F. Berman, G.C. Fox, and A.J.G. Hey (Eds.), *Grid Computing - Making a Global Infrastructure a Reality*, chapter 21, pp. 555–578, 2003.
  - [18] H. Bal, H. Casanova, J. Dongarra, and S. Matsuoka, Application Tools, in I. Foster and C. Kesselman (Eds.), *The Grid2 - Blueprint for a New Computing Infrastructure*, Chapter 24, pp. 463–489, 2004.
  - [19] W.M.P. van der Aalst and A.H.M. ter Hofstede, YAWL: Yet Another Workflow Language, *Information Systems* **30**, 245–275, 2005
  - [20] B. Ludäscher, I. Altintas, C. Berkley, D. Higgins, E. Jaeger, M. Jones, E.A. Lee, J. Tao, Y. Zhao, Scientific Workflow Management and the Kepler System, *Concurrency and Computation: Practice and Experience*, **18**, 1039–1065, 2006.
  - [21] D.W. Walker, E. Houstis (guest editors), *Complex Problem-Solving Environments for Grid Computing*, *Fut. Gen. Comp. Sys. Grid Computing* **21**, 841–968, 2005.
  - [22] R. Armstrong, G. Kumfert, L. Curfman McInnes, S. Parker, B. Allan, M. Sottile, T. Epperly, and T. Dahlgren, The CCA component model

- for high-performance scientific computing, *Concurr. Comput. : Pract. Exper.* **18**, 215–229, 2006.
- [23] M. Malawski, M. Bubak, M. Placzek, D. Kurzyniec, and V. Sunderam, Experiments with Distributed Component Computing Across Grid Boundaries, in *IEEE HPC-GECO/CompFrame 2006*, Paris, June 19–20, pp. 109–116, 2006.
- [24] M. Bubak, T. Gubala, M. Kapalka, M. Malawski, and K. Rycerz, Workflow Composer and Service Registry for Grid Applications, *Fut. Gen. Comp. Sys. Grid Computing* **21**, 77–86, 2005.
- [25] CrossGrid project – <http://www.crossgrid.org>
- [26] K-WfGrid Project – <http://www.kwfgrid.eu>
- [27] G. Allen, T. Goodale, M. Russell, E. Seidel, J. Shalf, Classifying and Enabling Grid Applications, in F. Berman, G. Fox, and A.J.G. Hey (Eds.), *Grid Computing, Making the Global Infrastructure a Reality*, chapter 23, 2003.
- [28] W. Sudholt, K. Baldridge, D. Abramson, C. Enticott, and S. Garic, Parameter Scan of an Effective Group Difference Pseudopotential Using Grid Computing, *New Generation Computing* **22**, 125–135, 2004.
- [29] A.G. Hoekstra and P.M.A. Sloot, Introducing Grid Speedup Gamma: A Scalability Metric for Parallel Applications on the Grid, in P.M.A. Sloot, A.G. Hoekstra, T. Priol, A. Reinefeld, and M.T. Bubak (Eds.), *Advances in Grid Computing - EGC 2005*, Lecture Notes in Computer Science vol. **3470** (Springer, Berlin, Heidelberg), 245–249, 2005.
- [30] A. Tirado-Ramos, A. Gualandris, and S.F. Portegies Zwart, Performance of a Parallel Astrophysical N-body Solver on pan-European Computational Grids, in P.M.A. Sloot, A.G. Hoekstra, T. Priol, A. Reinefeld, and M.T. Bubak (Eds.), *Advances in Grid Computing - EGC 2005*, Lecture Notes in Computer Science vol. **3470** (Springer, Berlin, Heidelberg) 2005.
- [31] A. Gualandris, S. Portegies Zwart, A. Tirado-Ramos, Performance analysis of direct N-body algorithms for astrophysical simulations on distributed systems, *Parallel Computing*, in press, 2007 (also: ArXiv Astrophysics e-prints astro-ph/0412206).
- [32] J.Q. Broughton, F.F. Abraham, N. Bernstein, and E. Kaxiras, Concurrent coupling of length scales: Methodology and application, *Phys. Rev. B* **60**, 2391–2403, 1999.
- [33] V.V. Krzhizhanovskaya, P.M.A. Sloot, and Y.E. Gorbachev, Grid-based Simulation of Industrial Thin-Film Production, *Simulation: Transactions of the Society for Modeling and Simulation International*, (Special

- Issue on Applications of Parallel and Distributed Simulation) **81**, 77-85, (ISSN 0037-5497) 2005.
- [34] W. E. B. Engquist, X. Li, W. Ren, and E. van den Eijnden, Heterogeneous Multiscale Methods: A Review, *Commun. Comput. Phys.* **2**, 367–450, 2007.
- [35] M. Bubak, M. Malawski, and K. Zajac, , Architecture of the Grid for Interactive Applications, in P.M.A. Sloot et al. (Eds.), *Proceedings of Computational Science - ICCS 2003*, International Conference Melbourne, Australia and St. Petersburg, Russia, June 2003, Lecture Notes in Computer Science **2657**, Part I, 207–213 Springer, 2003.
- [36] B. Balis, M. Bubak, W. Funika, R. Wismueller, M. Radecki, T. Szepieniec, T. Arodz, and M. Kurdziel, Grid Environment for On-line Application Monitoring and Performance Analysis, *Scientific Programming* **12**, pp. 239–251 2004.
- [37] K. Rycerz, M. Bubak, M. Malawski, and P. M. A. Sloot, A Framework for HLA-Based Interactive Simulations on the Grid, *SIMULATION* **81**, 67–76, 2005.
- [38] J. Makino, GRAPE and Project Milkyway, *Journal of Korean Astronomical Society* **38**, 165–168, 2005.
- [39] P. Hut, Dense Stellar Systems as Laboratories for Fundamental Physics, *ArXiv Astrophysics e-prints*, 2006.
- [40] M. R. Bate, I. A. Bonnell, The origin of the initial mass function and its dependence on the mean Jeans mass in molecular clouds *MNRAS* , **356**, 1201–1221, 2005.
- [41] D. Zengin, E. R. Pekünlü, E. Tigrak, Collapse of Interstellar Molecular Clouds, in B. Uyaniker, W. Reich, R. Wielebinski (eds.), *The Magnetized Interstellar Medium*, p. 133–136, 2004.
- [42] S. C. Whitehouse, M. R. Bate, The thermodynamics of collapsing molecular cloud cores using smoothed particle hydrodynamics with radiative transfer, *MNRAS* **367**, 32–38, 2006.
- [43] P. Padoan, and Nordlund, The Stellar Initial Mass Function from Turbulent Fragmentation, *ApJ* **576**, 870–879, 2002.
- [44] J. Barnes, P. Hut, A Hierarchical  $O(N \log N)$  Force-Calculation Algorithm, *Nature* , 324, 446–449, 1986.
- [45] A. Kawai, J. Makino, T. Ebisuzaki, Performance Analysis of High-Accuracy Tree Code Based on the Pseudoparticle Multipole Method, *ApJS* , 151, 13–33, 2004.
- [46] P. Eggleton, *Evolutionary Processes in Binary and Multiple Stars*, Cambridge University Press (Cambridge, UK), ISBN 0521855578, 2006.

- [47] J. Makino, P. Hut, Bottlenecks in simulations of dense stellar systems, *ApJ* **365**, 208–218, 1990.
- [48] P. Kroupa, C. A. Tout, G. Gilmore, The low-luminosity stellar mass function, *MNRAS* **244**, 76–85, 1990.
- [49] P. P. Eggleton, M. J. Fitchett, C. A. Tout, The distribution of visual binaries with two bright components, *ApJ* **347**, 998–1011, 1989.
- [50] J. R. Hurley, O. R. Pols, C. A. Tout, Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity, *MNRAS* **315**, 543–569, 2000.
- [51] S. P. Goodwin, The initial conditions of young globular clusters in the Large Magellanic Cloud, *MNRAS* **286**, 669–680, 1997.
- [52] C. M. Boily, P. Kroupa, The impact of mass loss on star cluster formation - II. Numerical N-body integration and further applications, *MNRAS* **338**, 673–686, 2003.
- [53] H. J. G. L. M. Lamers, M. Gieles, , S. F. Portegies Zwart, Disruption time scales of star clusters in different galaxies, *A&A* **429**, 173–179, 2005.
- [54] S. F. Portegies Zwart, S. L. W. McMillan, , P. Hut, J. Makino, Star cluster ecology - IV. Dissection of an open star cluster: photometry, *MNRAS* **321**, 199–226, 2001.
- [55] S. J. Aarseth, M. Lecar, Computer simulations of stellar systems, *ARA&A* **13**, 1–88, 1975.
- [56] S. J. Aarseth, *PASP From NBODY1 to NBODY6: The Growth of an Industry*, **111**, 1333–1346, 1999.
- [57] J. Gemmeke, S. Portegies Zwart, C. Kruip, Detecting Irregular Orbits in Gravitational N-body Simulations, *ArXiv Astrophysics e-prints*, 2006.
- [58] J. Makino, P. Hut, Performance analysis of direct N-body calculations, *ApJS* **68**, 833–856, 1988.
- [59] J. Makino, Direct Simulation of Dense Stellar Systems with GRAPE-6, in S. Deiters, B. Fuchs, A. Just, R. Spurzem, and R. Wielen, *ASP Conf. Ser. 228: Dynamics of Star Clusters and the Milky Way*, 87, 2001.