Le futur des simulations

E. Athanassoula

Obs. de Marseille

Outline

Massively parallel super-computers :

TOP100 / MareNostrum / BlueGene a l'IDRIS

Hardware advances

GRAPE / FPGAs / GPUs /GRAPE – DR

Software / Collaborations :

NEMO / Starlab / MUSE

Horizon

Rank	Site	Computer	Processors	Year	R_{max}	R_{peak}
1	DOE/NNSA/LLNL United States	BlueGene/L - eServer Blue Gene Solution IBM	212992	2007	478200	596378
2	Forschungszentrum Juelich (FZJ) Germany	JUGENE - Blue Gene/P Solution IBM	65536	2007	167300	222822
3	SGI/New Mexico Computing Applications Center (NMCAC) United States	SGI Altix ICE 8200, Xeon quad core 3.0 GHz SGI	14336	2007	126900	172032
4	Computational Research Laboratories, TATA SONS India	EKA - Cluster Platform 3000 BL460c, Xeon 53xx 3GHz, Infiniband Hewlett-Packard	14240	2007	117900	170880
5	Government Agency Sweden	Cluster Platform 3000 BL460c, Xeon 53xx 2.66GHz, Infiniband Hewlett-Packard	13728	2007	102800	146430
6	NNSA/Sandia National Laboratories United States	Red Storm - Sandia/ Cray Red Storm, Opteron 2.4 GHz dual core Cray Inc.	26569	2007	102200	127531
7	Oak Ridge National Laboratory United States	Jaguar - Cray XT4/XT3 Cray Inc.	23016	2006	101700	119350
8	IBM Thomas J. Watson Research Center United States	BGW - eServer Blue Gene Solution IBM	40960	2005	91290	114688
9	NERSC/LBNL United States	Franklin - Cray XT4, 2.6 GHz Cray Inc.	19320	2007	85368	100464
10	Stony Brook/BNL, New York Center for Computational Sciences United States	New York Blue - eServer Blue Gene Solution IBM	36864	2007	82161	103219
11	DOE/NNSA/LLNL United States	ASC Purple - eServer pSeries p5 575 1.9 GHz IBM	12208	2006	75760	92781
12	Rensselaer Polytechnic Institute, Computational Center for Nanotechnology Innovations United States	eServer Blue Gene Solution IBM	32768	2007	73032	91750
13	Barcelona Supercomputing Center Spain	MareNostrum - BladeCenter JS21 Cluster, PPC 970, 2.3 GHz, Myrinet IBM	10240	2006	63830	94208

IDRIS - IBM

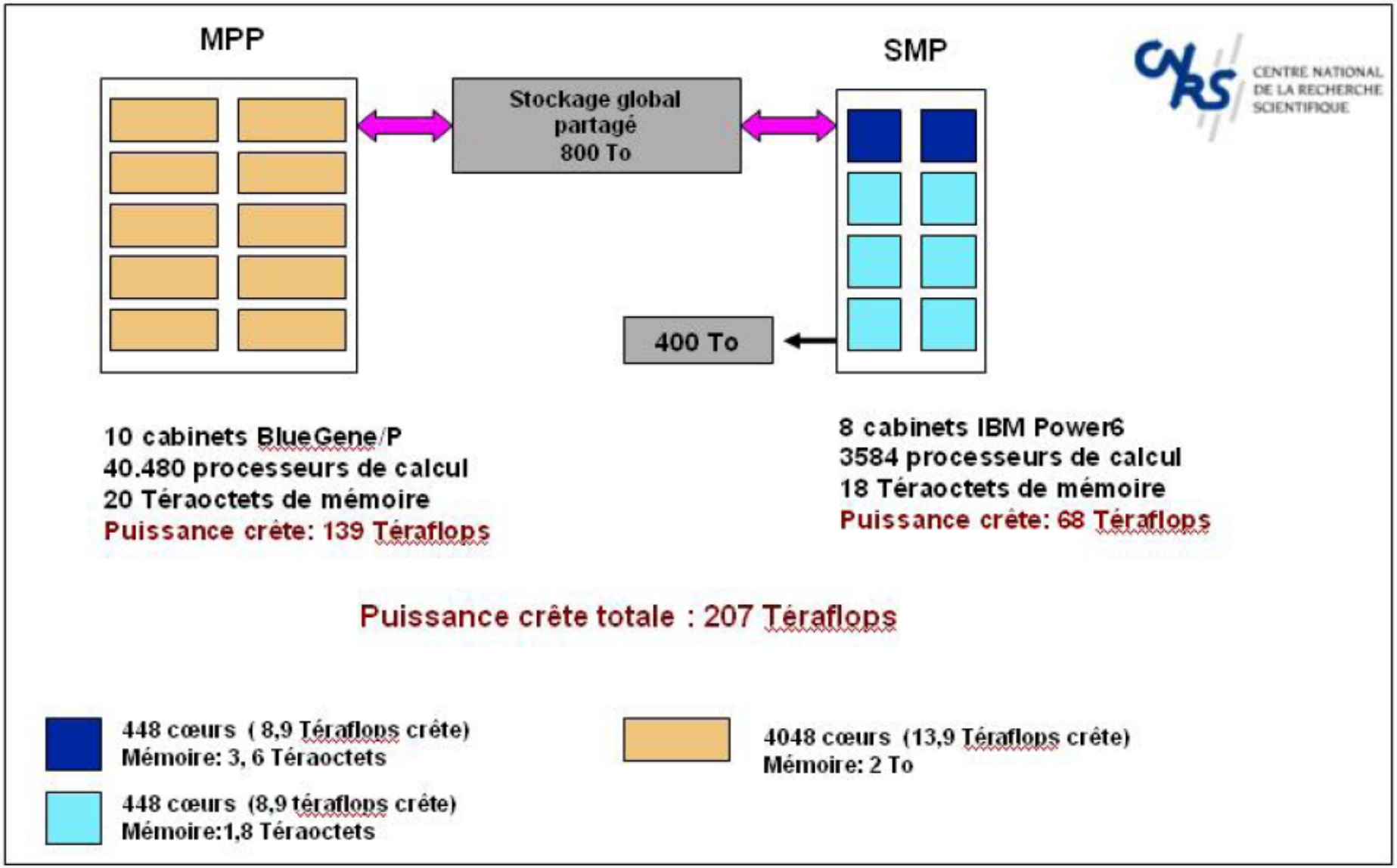
MARENOSTRUM

Barcelona Supercomputing Centre

- Peak Performance of 94,21 Teraflops
- 10240 IBM Power PC 970MP processors at 2.3 GHz (25
- 20 TB of main memory
- 280 + 90 TB of disk storage
- Interconnection networks:
  - Myrinet and Gigabit Ethernet
- Linux: SuSe Distribution

MIMD = Multiple Instructions Multiple Data

La figure ci-dessous schématise la nouvelle plate-forme de calcul intensif du CNRS



MIMD

Bad news : Gadget2 will perform badly

Outline

Massively parallel super-computers :

TOP100 / MareNostrum / BlueGene a l'IDRIS

Hardware advances

GRAPE / FPGAs / GPUs /GRAPE – DR

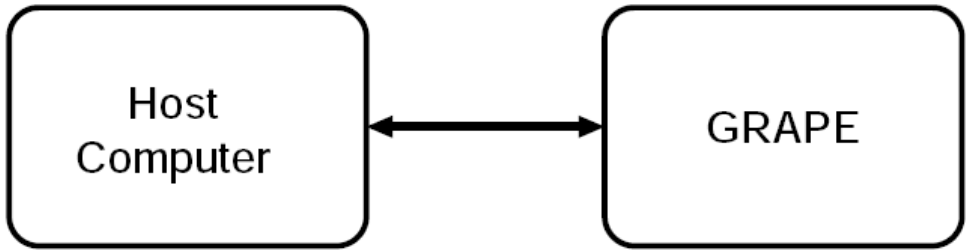
Software / Collaborations :

NEMO / Starlab / MUSE

Horizon

Basic concept

- With N-body simulation, almost all calculation goes to the calculation of particle-particle interaction.
- This is true even for schemes like Barnes-Hut treecode or FMM.
- A simple hardware which just calculates the particle-particle interaction can greatly accelerate overall calculation.



GRAPE6

GPU

Graphics Processing Unit

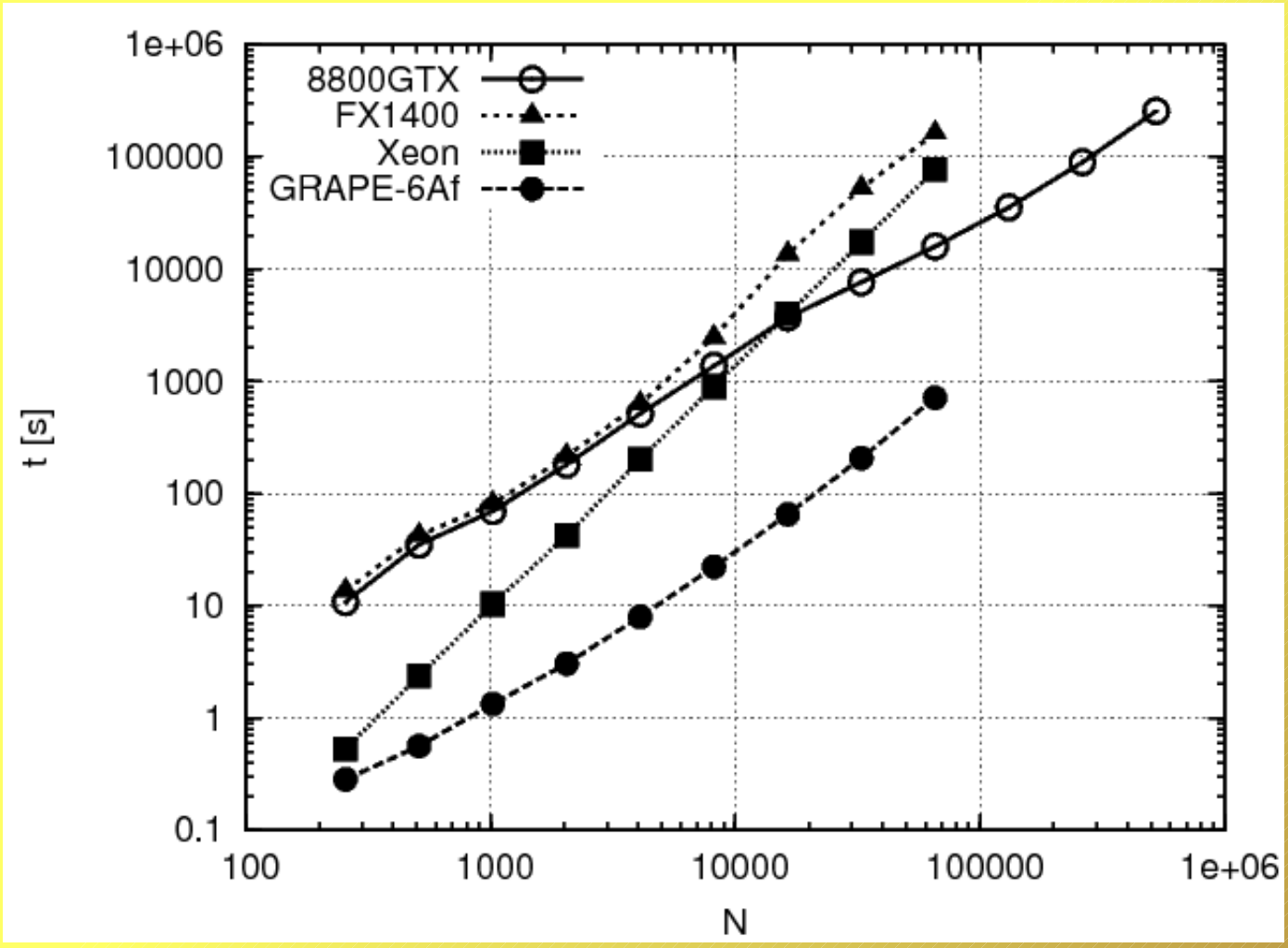
High power parallel vector processors developed by the gaming industry

SIMD = Single Instruction Multiple Data

Advantages :
Very fast, more than the standard CPUs
Moore's law : Doubling in performance over 9 months (instead of 18 for CPUs)
Cheap

Double precision version should be out soon

Ndirect - GPU avec Cg

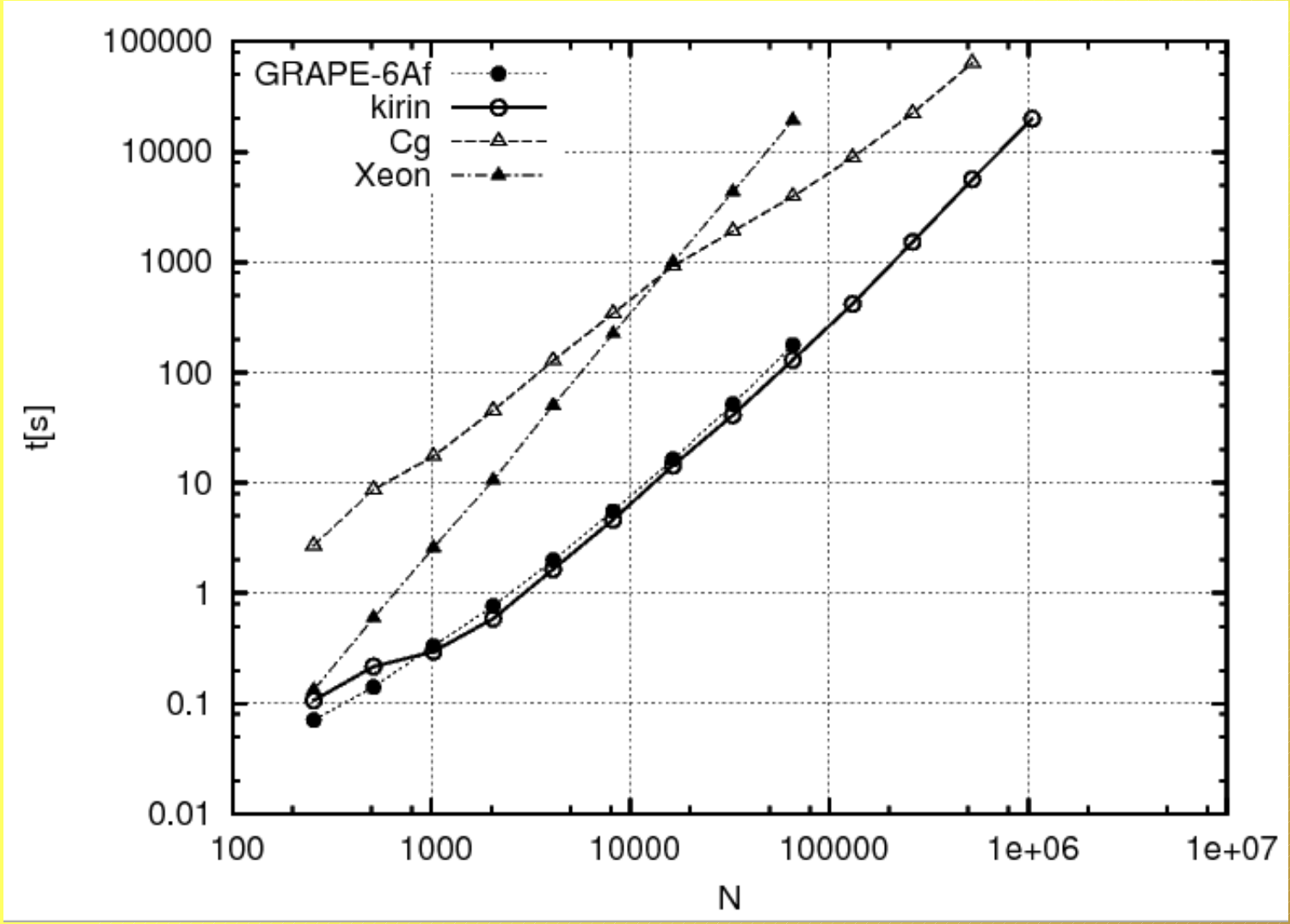


Cg programming language

GPUs also less accurate

Portegies-Zwart et al 2007

Ndirect – GPU avec CUDA

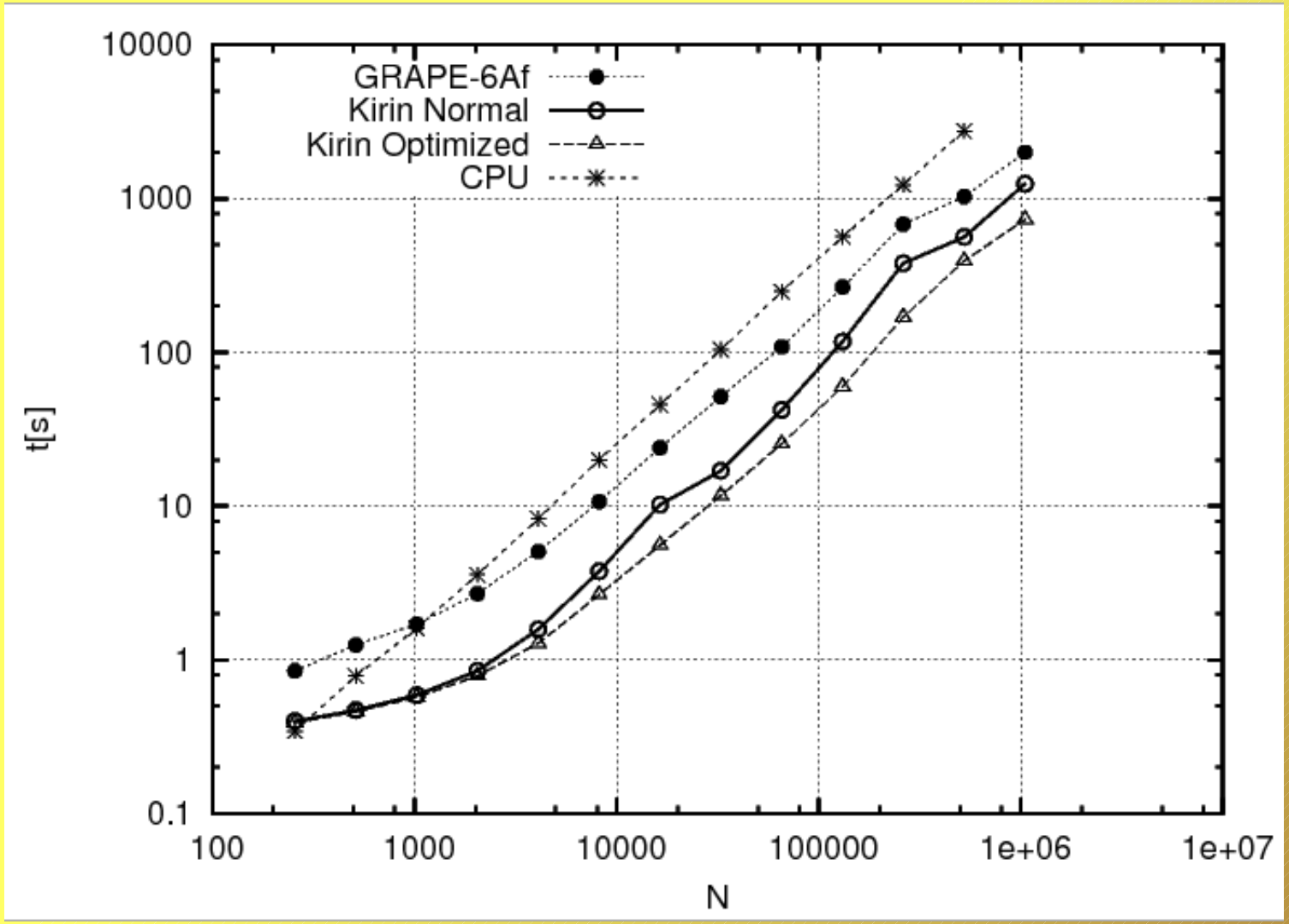


CUDA programming languange

Belleman et al 2008

Comparison not very fair because GRAPE6 dates from 2002 (6 years old!)

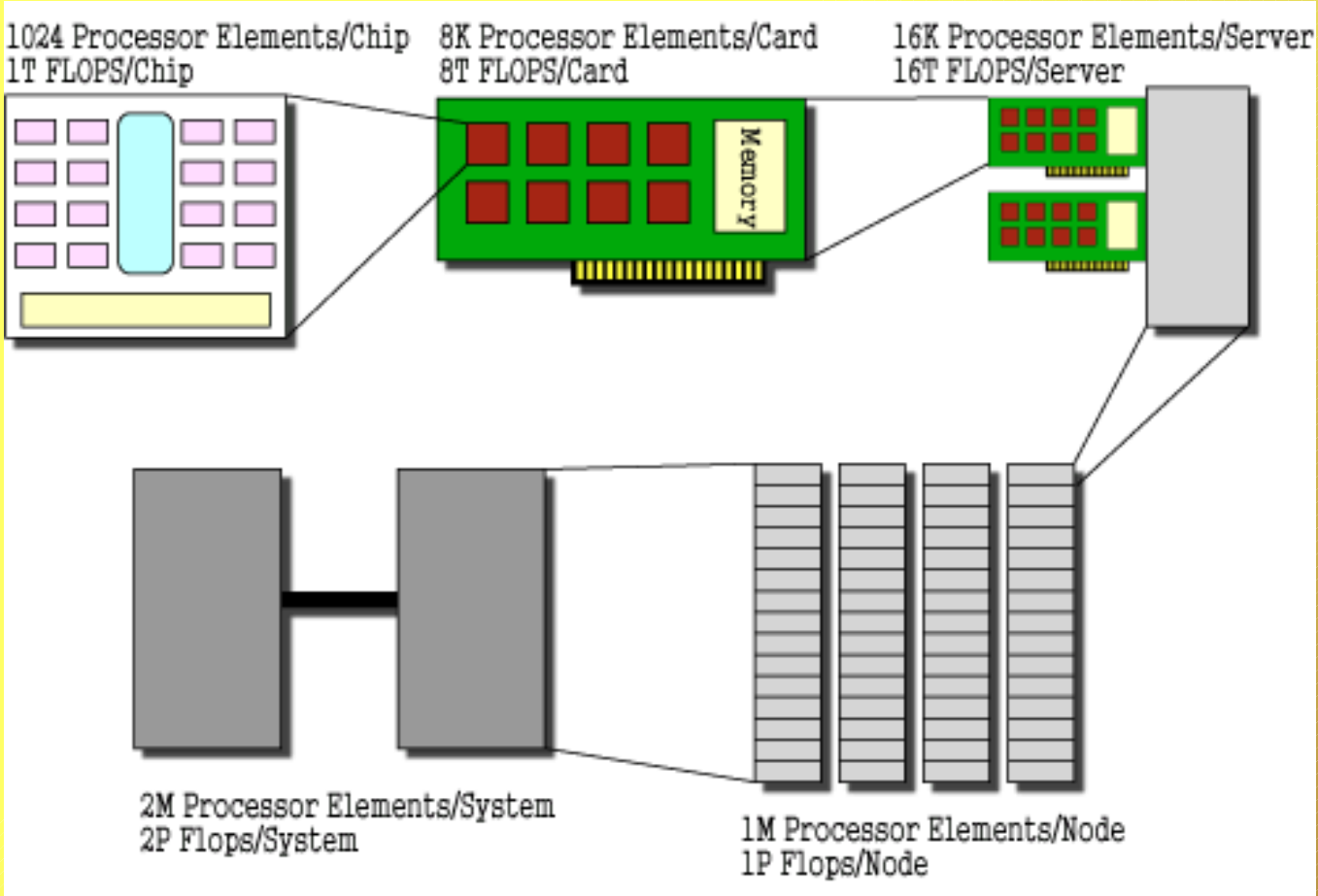
Treecode – GPU avec CUDA



Belleman et al 2008

GRAPE - DR

SIMD



Network used in the experiment

Outline

Massively parallel super-computers :

TOP100 / MareNostrum / BlueGene a l'IDRIS

Hardware advances

GRAPE / FPGAs / GPUs /GRAPE – DR

Software / Collaborations :

NEMO / Starlab / MUSE

Horizon

NEMO

NEMO is an extendible Stellar dynamics toolbox, following an Open-Source software model. It has various programs to create, integrate, analyse and visualise N-body and SPH systems.

4 000 files
1 000 000 lines of code and documentation
C mainly (some C++ and Fortran)

http://bima.astro.umd.edu/nemo

MUSE

- Multiscale/multiphysics/multipurpose/multiuser Scientific Environment (MUSE) is the next step for (astro)physics modeling
  - interchangeable modules with standardised interfaces
  - allows experimentation and comparison
  - individual modules: clean implementations of specific pieces of physics
  - easy to add new modules

muse.li

Work in progress

S. Harfst et al 2008

MUSE motivation

- many codes exists for detailed :
-
    - many- and few-body dynamics
    - stellar evolution
    - stellar collisions
    - hydrodynamics
    - radiative transfer

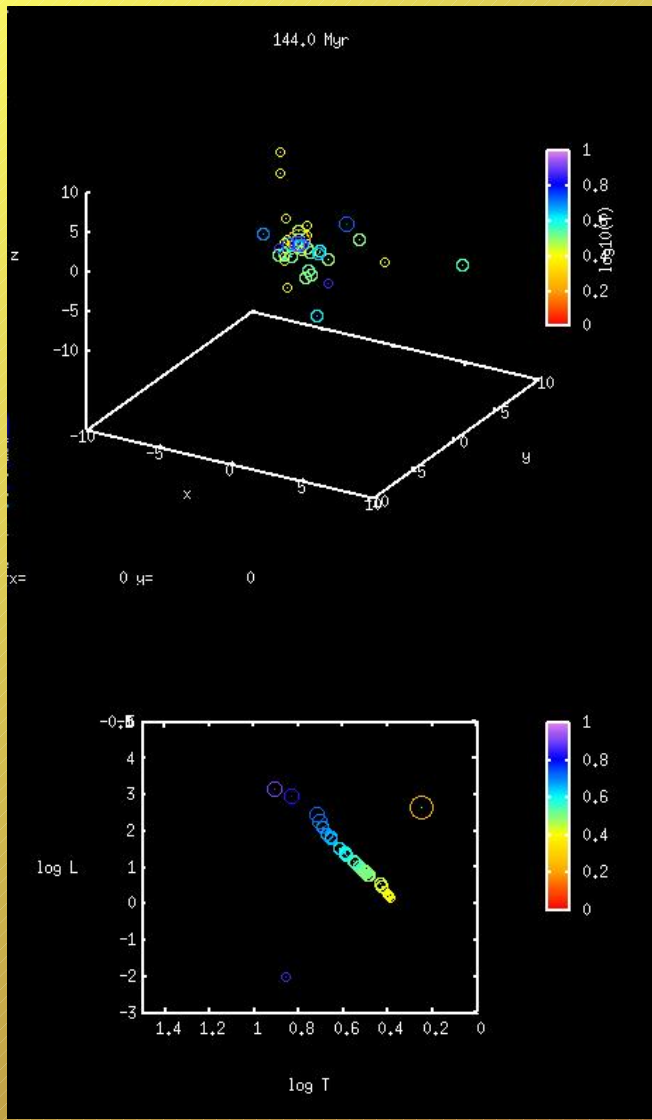
We need a framework for adding them together

The basic N-body scheduler

```
import BHTree as Gravity
grav = Gravity()


LoadNbodySystem()
grav.initialize(0)

t_end   = 10.
dt_step = 0.125

t_step = dt_step
while t_step <= t_end:
  print ''
  print 'integrating to time', t_step
  grav.evolve(t_step)
  output()
  tstep += dt_step

print ‘done :-)’
```

Simulating a Star Cluster in *MUSE*

```
import hermite0 as SD
import EFT89 as SE
import sticky_spheres as SC

.
. (initialisation)
.

while time <= t_end:

  time += dtime
  while SD.get_time()<time:
    id1 = SD.evolve(time)

    if id1 > 0:    # collision detected
      id2 = SD.find_colliding_secondary(id1)
      SE.evolve(SD.get_time())
      SC.collide_pair(id1,id2)

  SE.evolve(time)

  output()

print “end at t = ”,time,“ Nstars = ”,SE.get_number()
```



G3 : Global Grape Grid

Drexel: Steve McMillan

Amsterdam: Derek Groen & SPZ

Tokyo: Jun Makino

Groen et al. 2008

GRID computing in *MUSE*

worker nodes

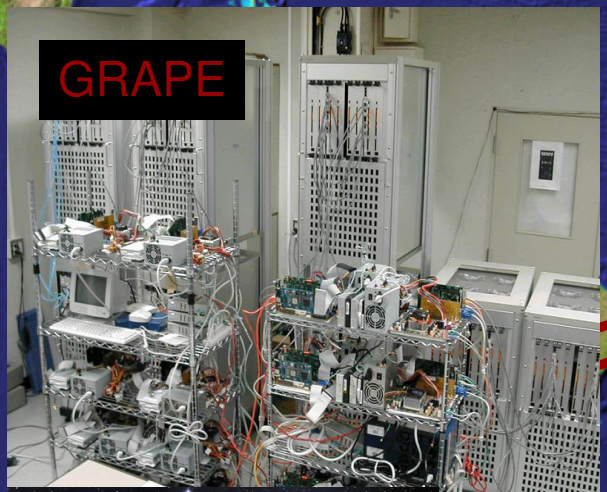
user interface

worker node

Stellar Evolution

Beowulf

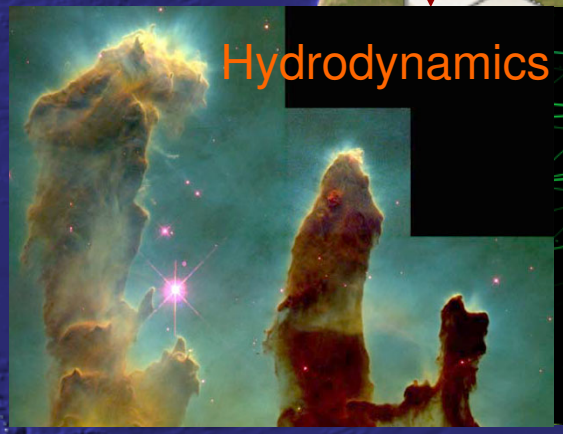
GRID server



GRAPE

nodes

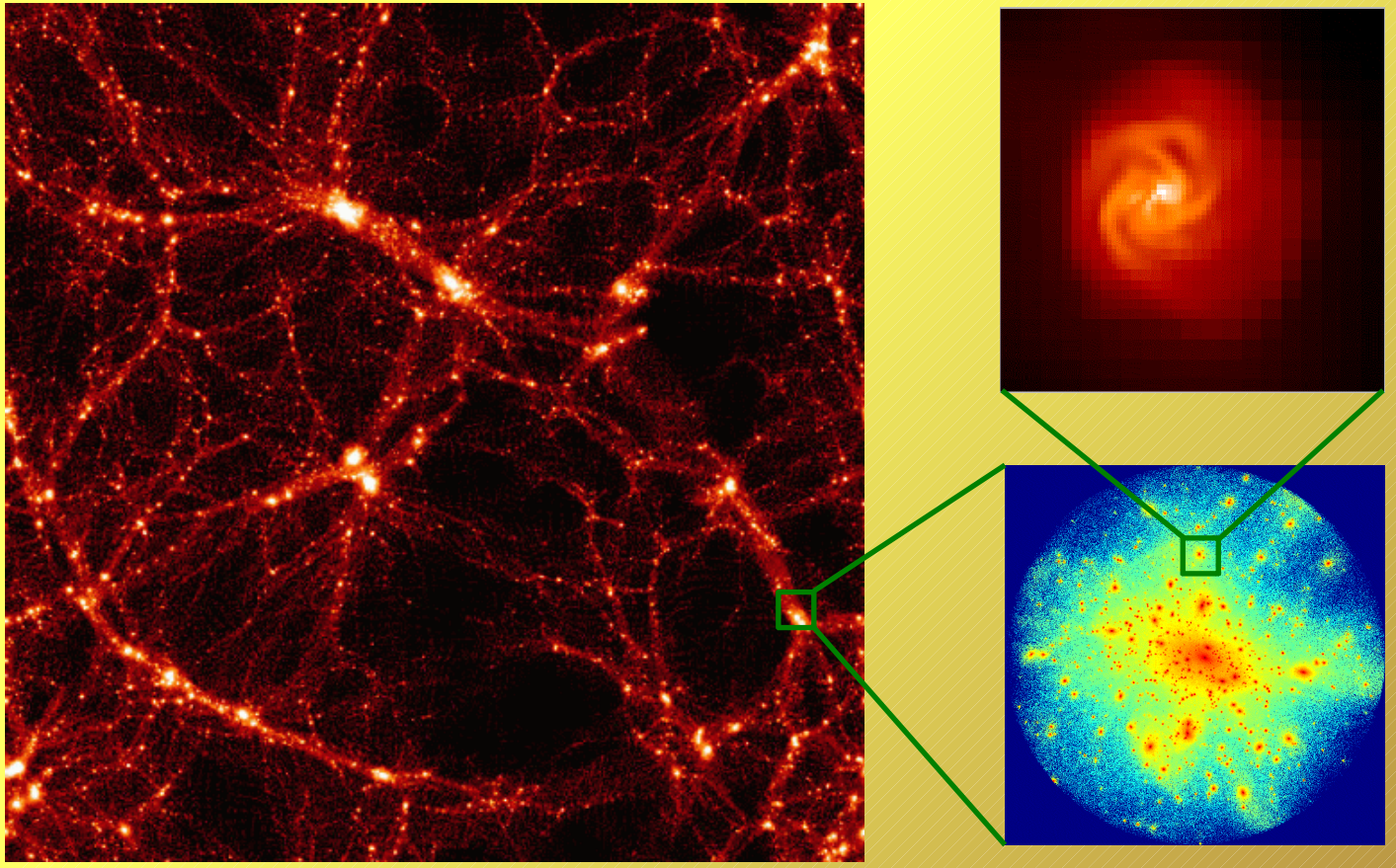
Stellar Dynamics



Hydrodynamics

GPU

Projet Horizon: computational astrophysics on massively parallel systems to understand galaxy formation in the universe.



Computer infrastructure

Supercomputer centers

HPC1

Post-processing on large data sets

Jobs submission

Visualization, Post-processing on Small data sets

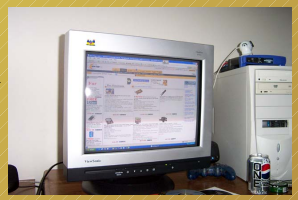
Mini-grid Horizon

Paris

Meudon

Saclay

Lyon

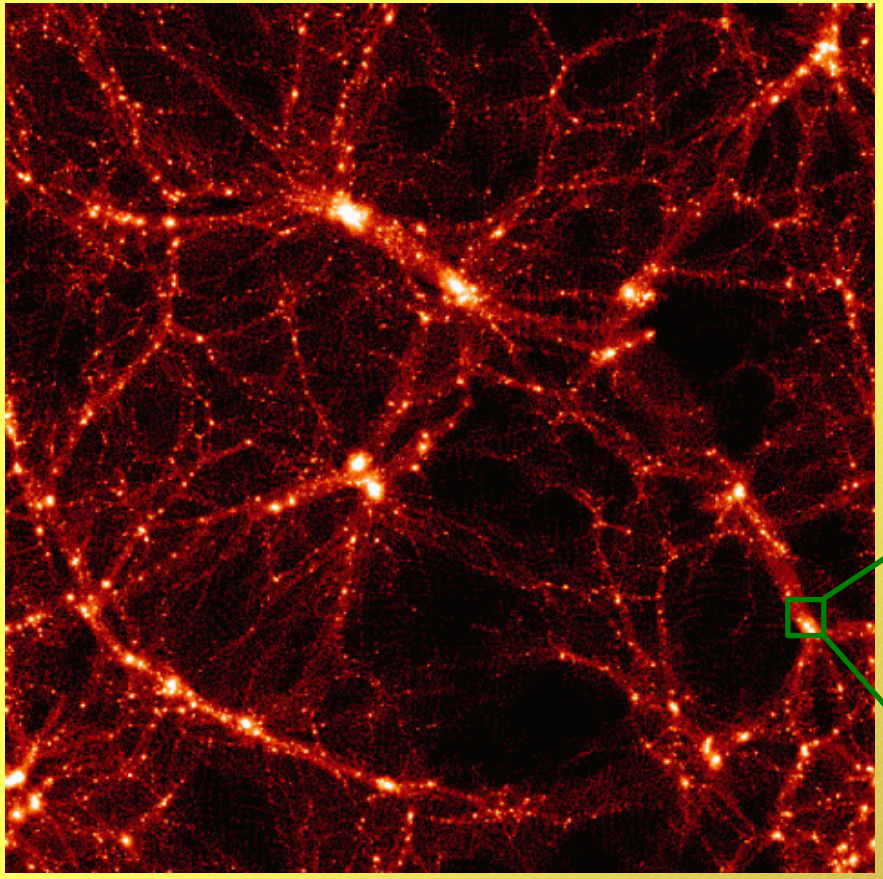


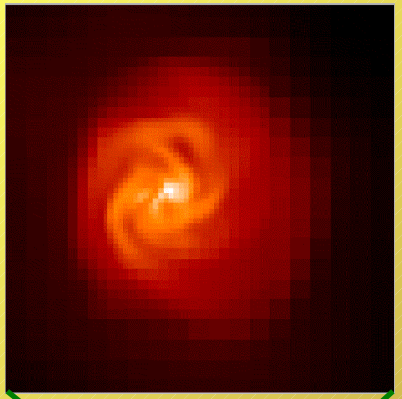
Marseille

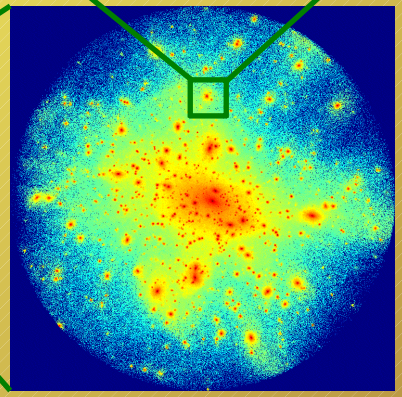
Software resources shared and posted to an internal web site

- Gas dynamics and N body codes
- Data conversion utilities
- Common post-processing tools
- Common visualization tools
- Parallel computing tutorials
- Common initial conditions for code benchmarking
- CVS repository

Projet Horizon: computational astrophysics on massively parallel systems to understand galaxy formation in the universe.







Summary

Many hardware advances on large scale
- Massively parallel MIMD machines (IDRIS , MareNostrum etc)
- Massively parallel SIMD machines (GRAPE-DR etc)

Many hardware advances on a more local scale
- GRAPE7
- GPUs

All these necessitate also a very considerable software investment.

Software libraries and collaborations (NEMO, Starlab, MUSE ....)

More astronomy oriented collaborations (MODEST, HORIZON, VIRGO etc)