One decade of GPUs for cosmological simulations (in Strasbourg) : fortunes & misfortunes

Dominique Aubert

with P. Ocvirk, J. Chardin, J. Lewis, N. Deparis (Strasbourg, F) & N. Gillet (SNS Pisa, It)
**The epoch of Reionization**

Reionization

- end by $z \sim 6$; $t \sim 1$ Gyr
- a great and rapid cosmological transition (few $100s$ Myrs)
- Driven by IONIZING LIGHT and the first sources (and *dominated by stars*)
- network of HII regions (*ionized bubbles*)
- ionization ($0.9999$) & heating ($\sim 10,000$ K) of IGM (by stars and quasars)
- UV Background
- the initial stages of galaxy formation
CODA I-AMR
(Aubert+ 18)
91 Mpc/2048³
16 billions resolution elements with AMR @ z=6
32768 cores+4096 GPUs on Titan(DOE/ORNL) using EMMA simulation code (Aubert+ 15)
20+ millions cpu hours
Jan.-Mar 2017

INCITE CODA
(PI : Shapiro)

WMAP5-CLUES ICs
spatial res.=500 pc
mass res~2e6 Msol
stellar res ~70 000 Msol

Aubert+ 2018
A few words about GPUs

- Graphics Processing Units
- Large number of Multiprocessors (~100) + Scheduler

- Fit to Large number of Independent, intensive and memory friendly tasks

- x10 to x100 compared to CPU
- High-level interface with CUDA (Nvidia), OpenCL (Kronos)
- High-end GPUs: up to 16 GB VRAM
- Volta GPUs (e.g. Summit/ORNL):
  - unified CPU/GPU memory,
  - GPU/GPU comm,
  - GPU shared among CPUs
  - hardwired tensor operations

- Explicit conservative solver of Radiative Transfer equations + local thermo-chemistry: ensures high load and local computations for cell update
- also true for grid-based Hydrodynamics with non trivial Riemann solvers, reconstructions, etc..
2008 Direct N-Body Calculation
(With Christian Boily & Romaric David, Strasbourg)

100k bodies

similar throughput as GRAPE boards
(with hardwired $1/r^2$ calculators)
with
GPUs GeForce 8800 GTX (with a cost 10x smaller !)
A Particle-Mesh Integrator for Galactic Dynamics Powered by GPGPUs

Dominique Aubert\textsuperscript{1}, Mehdi Amini\textsuperscript{2}, and Romaric David\textsuperscript{2}

\textsuperscript{1} Observatoire Astronomique, Université de Strasbourg, France
\textsuperscript{2} Direction Informatique, Université de Strasbourg, France

Abstract. We present a particle-mesh N-body integrator running on GPU using CUDA. Relying on a grid-based description of the gravitational potential, it can simulate the evolution of self-interacting ‘stars’ in order to model e.g. galaxies. All the steps of the application have been ported on the GPU, namely 1/ an histogramming algorithm with CUDPP, 2/ of the resolution of the Poisson equation by means of FFT with CUFFT and multi-grid relaxation, 3/ of an optimized finite-difference scheme to compute the accelerations of stars and 4/ of an update procedure for positions and velocities. We present several tests at different resolution, and reach a speedup from 2 to 50 depending on the resolution and on the test case.
Multi-GPU PM

1.2 billion particles
(1024^3 real particles +2 \times 10^8 ghosts)
8 sec/tstep on 64 Teslas
with 25 % spent in communications
with sort optimisation
we may expect 6 sec/tstep
communication~40%
asynchronous coms?

presented
@ GPU technology conference 2010,
unpublished
Simulations of the Cosmic Reionization
Explicit treatment of a fluid-like radiation

RT Moment **benefits** from grid based hydrodynamics methods

\[
\frac{\partial N_\nu}{\partial t} + \frac{\partial F_\nu}{\partial r} = S_\nu - \kappa_N N_\nu, \\
\frac{\partial F_\nu}{\partial t} + c^2 \frac{\partial P_\nu}{\partial r} = -\kappa_F F_\nu.
\]

RT Moment must also **satisfy the same constraints** as grid based hydrodynamics methods.

**We use an explicit solver:**
Because of the speed of light, RT can be quite expensive compared to collisionless and hydro-dynamics

\[ \Delta t < \frac{\Delta x}{v} \]
Courant Condition

Typically **200-600 rad. steps per dyn. step**

One solution (among others) : **Hardware acceleration**
Radiative Post-Processing with ATON (Aubert & Teyssier 2010)

128 GPUs (Titane/CEA/CCRT) 100 Mpc - $1024^3$

**Close-circuit GPU code** without transfers between the CPU host and the devices

**Static cartesian mesh**: ensures predictability of computation and memory access patterns

x80 acceleration factor: $c=1$
<table>
<thead>
<tr>
<th>Purpose</th>
<th>ATON (or CUDATON) (Aubert &amp; Teyssier 2008,2010)</th>
<th>RAMSES-CUDATON 2013 Ocvirk, Gillet, Shapiro, Aubert et al. 2016</th>
<th>EMMA Aubert et al. 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rad. Post-Processing</td>
<td>Rad. Post-Processing of a pre-existing hydro simulation</td>
<td>On the fly interaction of RAMSES (CPU) &amp; CUDATON (GPU)</td>
<td>Multi-purpose cosmological simulation code with Radiative transfer</td>
</tr>
<tr>
<td>Radiative hydrodynamics</td>
<td>[Red]</td>
<td>[Green]</td>
<td>[Green]</td>
</tr>
<tr>
<td>Adaptive Mesh refinement</td>
<td>[Red]</td>
<td>[Red]</td>
<td>[Red]</td>
</tr>
<tr>
<td>GPU</td>
<td>x 80 (RT only)</td>
<td>(x80) RT almost free</td>
<td>x5 Overall (Poisson+Hydro+RT)</td>
</tr>
</tbody>
</table>
We can match $x(z)$ deduced from the observations without matching the observation themselves…

(Aubert & Teyssier 2010)
Calibration of Reionization Simulations using LyA Forest

Ly-A data put constraints on reionization models: overlap by z~7, accelerated evolution of opacity at z>6, large scatter of IGM opacities at late times!

Probable significant contribution of rare and bright sources (quasars?) to the UV background at z~6
Indications of a late reionization?

Radiative **Post-Processing** with cudATON (Aubert & Teyssier, 2010).
160 Mpc/h - $1024^3$ - SPH-Gadget Sherwood Simulation

Kulkarni+ 2018
Indications of a late reionization?

A much better fit to observed distributions of Intergalactic medium opacities, from Ly-Alpha Forest, is recovered with a z~5.3 reionization (without quasars). (« observed » xion(z) is likely to be a poor constraint of the reionization history)

Kulkarni+ 2018
<table>
<thead>
<tr>
<th></th>
<th>ATON (or CUDATON) (Aubert &amp; Teyssier 2008,2010)</th>
<th>RAMSES-CUDATON 2013 Ocvirk, Gillet, Shapiro, Aubert et al. 2016</th>
<th>EMMA Aubert et al. 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Rad. Post-Processing of a pre-existing hydro simulation</td>
<td>On the fly interaction of RAMSES (CPU) &amp; CUDATON (GPU)</td>
<td>Multi-purpose cosmological simulation code with RT</td>
</tr>
<tr>
<td>Radiative hydrodynamics</td>
<td>Red</td>
<td>Green</td>
<td>Green</td>
</tr>
<tr>
<td>Adaptive Mesh refinement</td>
<td>Red</td>
<td>Red</td>
<td>Red</td>
</tr>
<tr>
<td>GPU</td>
<td>x 80 (RT only)</td>
<td>(x80) RT almost free</td>
<td>x5 Overall (Poisson+Hydro +RT)</td>
</tr>
</tbody>
</table>
Setup: CoDall specs

- 16384 GPUs, 65536 CPUs
- 64 h^{-1} Mpc side, 4096^3 grid
- M_{halo_{min}} \sim 1 \times 10^8 \, M_\odot
- \Delta x \sim 22 \, \text{kpc comoving} \quad (< 3.2 \, \text{kpc physical})
- z_{end}=5.8
- \sim 6 \, \text{days runtime, 2 PB data}
- Planck 2013 cosmology
- New ICs: M_{Virgo} = 2.14 \times 10^9 \, M_\odot

(taken from illustris website)
Coda I
(Ocvirk+ 16,
Dawoodbhoy+18)
& Coda II
(Ocvirk+ in prep.,
Lewis+ in prep.)

Galaxy population properties
during the reionization with
radiative hydrodynamics
Ramses-CUDATON

Star Formation &
suppression, environmental
dependance
Photon budget & Escape
creation fraction
Luminosity functions

(see Ocvirk’s talk tomorrow)

Next Stages :
on Piz Daint (CSSC, CH) &
Summit (ORNL, US)
<table>
<thead>
<tr>
<th></th>
<th>ATON (or CUDATON) (Aubert &amp; Teyssier 2008,2010)</th>
<th>RAMSES-CUDATON 2013 Ocvirk, Gillet, Shapiro, Aubert et al. 2016</th>
<th>EMMA Aubert et al. 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Purpose</strong></td>
<td>Rad. Post-Processing of a pre-existing hydro simulation</td>
<td>On the fly interaction of RAMSES (CPU) &amp; CUDATON (GPU)</td>
<td>Multi-purpose cosmological simulation code with RT</td>
</tr>
<tr>
<td><strong>Radiative</strong></td>
<td>Red</td>
<td>Green</td>
<td>Green</td>
</tr>
<tr>
<td><strong>hydrodynamics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adaptive</strong></td>
<td>Red</td>
<td>Green</td>
<td>Green</td>
</tr>
<tr>
<td><strong>Mesh refinement</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GPU</strong></td>
<td>x 80 (RT only)</td>
<td>(x80) RT almost free</td>
<td>x5 Overall (Poisson+Hydro +RT)</td>
</tr>
</tbody>
</table>
AMR Cosmological RT with **EMMA**

- Electromagnétisme et Mécanique sur Maille Adaptative
- Full **standalone** cosmological code
- Collisionless Dynamics (PM) + Hydro (Godunov/MUSCL/HLL) + Radiative transfer (M1)
- Full **AMR** radiative transport (like e.g. Ramses-RT (Rosdahl et al. 2013)) or restricted to the Coarse grid with thermo-chemistry on refined levels
- Star Formation + SN Feedback
- C+MPI Parallelisation (scales up to 2048 cores and 1024^3 coarse cells)
- **Optional GPU** (CUDA) acceleration for the Poisson, Hydro and RT solver

4 Mpc - 128^3 + 5 AMR levels

Aubert, Deparis & Ocvirk, MNRAS 2015
Global performances

x4-x15 acceleration (including non-GPU procedures and general overheads)

Acceleration can be obtained at the x15 level when strict floating point operations are required (gcc O2 or icc O2 -strict-fp operation)

CPU can be quite competitive if this restriction is lifted: x4 acceleration only (icc O2)

Figure 22. Comparison of the cumulative time spent to reach a given expansion factor for a 4 Mpc/h-128^3 cosmological simulation of the reionization. Times are given for a single computing device (i.e. 1 GPU or 1 CPU core). The thick black dashed line stands for the GPU run performed on a M2090 Nvidia GPU whereas the thin dashed blue line stands for the contribution of radiative transfer to this cost. The black solid (resp. dotted) line stands for a single CPU core (2.7 GHz Sandybridge Westermere) using the gcc-O2 (resp. icc-O2) binary. The symbols stand for a 4-core CPU calculation using icc-O2 on a Curie node.

Aubert, Deparis & Ocvirk, MNRAS 2015
Decoupling Physics from Data Logistics - Stream vectors to accelerators

Physics Kernels

Single Core, Multi-Core, GPU, Xeon-Phi,…

Currently:
flat array of data to update +
flat array of neighbors =
redundant data
Data preparation and transfer currently limit the performance
Note that GS/Transfer x2 for GPUs
CODA I-AMR
91 Mpc/2048³
16 billions resolution elements with AMR
@ z=6

32768 cores+4096 GPUs
on Titan(DOE/ORNL)
using EMMA simulation code (Aubert+ 15)
(1.4 GPU acceleration rate)

20+ millions
cpu hours
Jan.-Mar 2017

INCITE CODA
(PI : Shapiro)

WMAP5-CLUES ICs
spatial res.=500 pc
mass res~2e6 Msol
stellar res ~70 000 Msol
reduced speed of light
(10%)
Reionization times of $z=0$ haloes

- Massive objects reionize early on
- First sources in most massive $z=0$ halos
- Light Objects reionize like the full volume
- The scatter is significant (~200 Myrs)
- Light Objects have median reionization times at later times than the full box (late reionization of faint objects?)
- At low-mass, progenitor-based predictions are biased
- Particle-based predictions tend to predict later reionization times (sensitive to the full reionization history of the object, even diffuse matter)

Most massive progenitor Reion. redshift

$<\text{Reion. Redshift}> \quad \text{of all DM particles}$

Median & 5-95% interval

Aubert+18
The case of the Local Group (I)

CLUES Initial Conditions produces a MW and M31 pair in proper cluster environment (Virgo & Fornax)
Ionization front speeds during reionization

reduced speed-of-light impact front speeds during the overlap stage
Cosmic variance of ionization front speeds and median reionization times

Deparis+ 18

512 x 12 Mpc boxes in a 91 Mpc large volume choice of c results in greater relative discrepancies in underdense regions
EMMA v2.0
codename ‘Kereon’

we trash the fully threaded AMR tree
and replace it with Morton multi-level ordering

will we be able to take advantage of:
• shared GPU/CPU memory?
• GPU/GPU comm?
• Zerocopy/Shared GPU/CPU memory?
• tensor operations?

For sure: a much much simpler code
What’s next for GPU-driven simulations?

The elephant in the room: **Machine Learning**

Machine learning may push hardware: we should be able to use the same hardware (even intrinsics for e.g. tensor operations or hard-wired upscaling/downscaling) = GPUs during simulation

Cognitive networks might actually replace physics modules = GPUs before simulations for greater efficiency (and greater modularity ?)

2015

2009

2020+