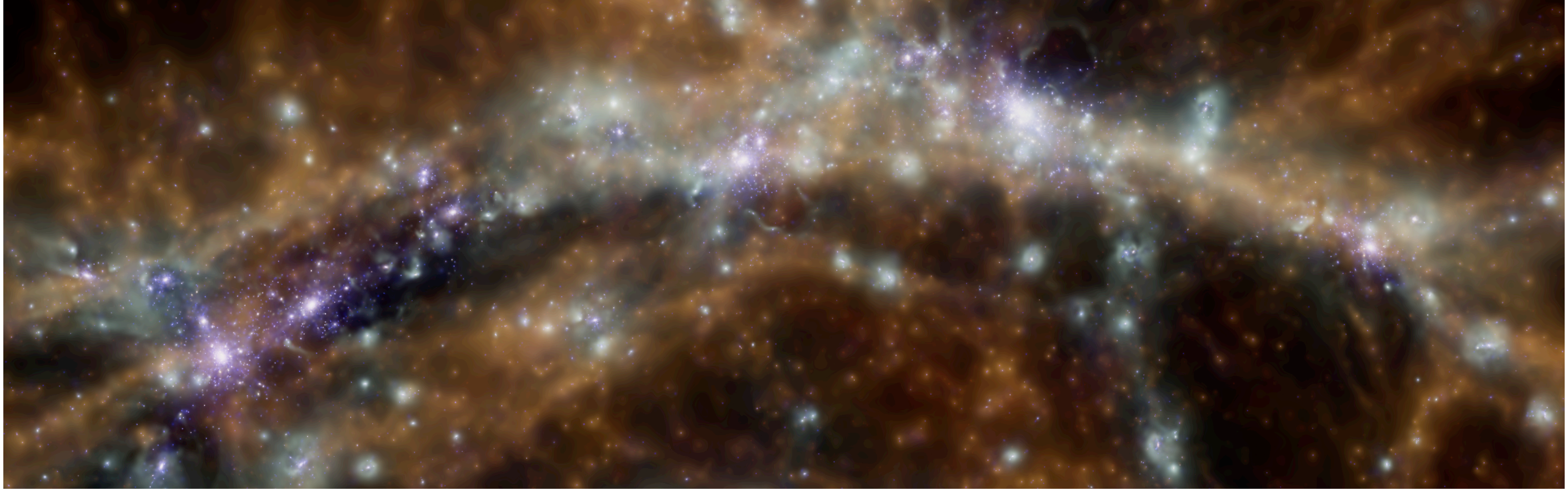


Scientific Visualization on GPU-Enabled, Hybrid HPC Systems

Mel Krokos

University of Portsmouth, UK

mel.krokos@port.ac.uk

GPU Technology Conference, San Jose, March 2014

Splotch Team

K. Dolag (*Munich Astrophysics Observatory*, Germany)

M. Reinecke (*Max Planck Institute for Astrophysics*, Germany)

C. Gheller (*ETH-CSCS*, Switzerland)

M. Rivi (*University of Oxford*, UK)

T. Dykes, M. Krokos (*University of Portsmouth*, UK)

MAX-PLANCK-GESELLSCHAFT

CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

UNIVERSITY OF OXFORD

University of Portsmouth

Overview

- Big Data in Astrophysics
- Splotch on GPUs
- Performance Tests
- Future Developments

BIG DATA IN ASTROPHYSICS

Big Data

Big Data : coming from Observations / Numerical Simulations with dramatically increasing sizes;

- need to be suitably *stored* and *managed*;
- *and be accessible* through suitable tools for exploring, processing and analysis.

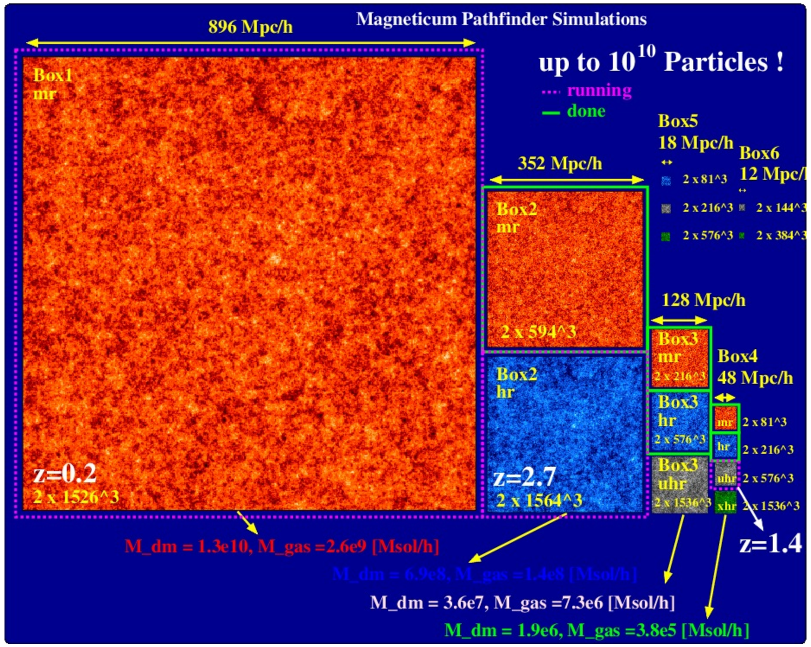
Examples:

- *SQUARE KILOMETER ARRAY* (*SKA*)
  - most ambitious astronomy project yet, comparable to CERN's LHC (*http://www.skatelescope.org*);
  - to generate over an *exabyte* (10^{18}) of data per day.
- Cosmological N-body / SPH simulations
  - GADGET (http://www.mpa-garching.mpg.de/gadget/)
  - magneticum (our target simulation, http://www.mpa-garching.mpg.de/~kdolag/Simulations/MagneticumPathfinder).

Pure gravity (N-body)	**Size**
10^{11} particles (100 billion)	10^{11} elements
A particle is represented at a minimum by position and velocity in 3D, i.e. 6 variables;	6×10^{11} elements
Each variable is typically represented as a float, i.e. 4 bytes;	~2.4 TBs
We normally want to save a number of time steps: e.g. 100.	~240 TBs

Millennium I	(WMAP1-GAD)	500 /h Mpc	10 billion particles
Millennium II	(WMAP1 –GAD)	100/h Mpc	10 billion particle
Millenium XXL	(WMAP1-GAD)	3 /h Gpc	303 billion particles
Bolshoi	(WMAP7-ART)	250/h Mpc	8 billion particles
Multidark	(WMAP7-ART)	1 /h Gpc	8 billion particles
BigMD	(WMAP7-GAD)	2.5/h Gpc	56.6 billion particles
MICE	(WMAP5-GAD)	7 /h Gpc	8 billion particles
Horizon (FR)	(WMAP3-RAMSES)	2 /h Gpc	68 billion particles
Horizon (KR)	(WMAP5-GOTPM)	10.7 /h Gpc	372 billion.
DEUS (FR)	(WMAP7-RAMSES)	21/h Gpc	550 billion particles
JUBILEE	*(WMAP7-CP3M)*	*6/h Gpc*	*216 billion particles*

Magneticum Pathfinder Simulations



- to study formation of cosmological structures
  - large scale / high resolution (N-Body+SPH) simulations
  - to allow comparison with a variety of multi-wavelength observational datasets
- we deploy a 6 billion dataset ~200 GB (~100 snapshots)

Visualization

- **Visual exploration** and **discovery** can be invaluable tools providing scientists with prompt and intuitive insights, enabling them to identify characteristics and manageable regions of interest in large-scale datasets for applying time consuming methods.

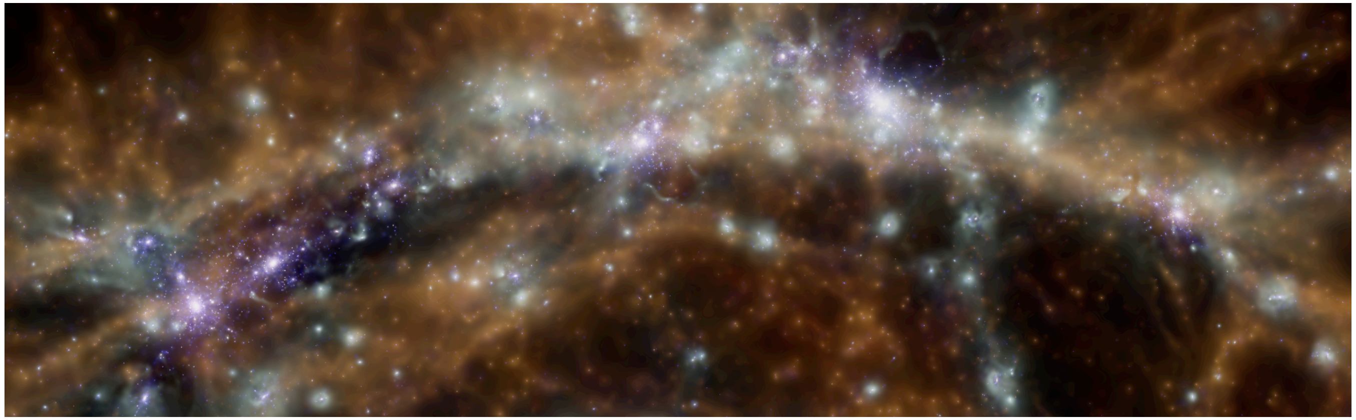
- **Splotch** is a rendering approach for large-scale datasets exploiting High Performance Computing architectures.

- To handle big data effectively visualization solutions need to satisfy:
    - high level of **customisation** to suit specialised **scientific datasets***;*
    - handling of **terabytes** (and more) of datasets;
    - exploitation of **large HPC infrastructures** (to reduce processing time and avoid data movements);
    - production of **high resolution / accurate imagery**.

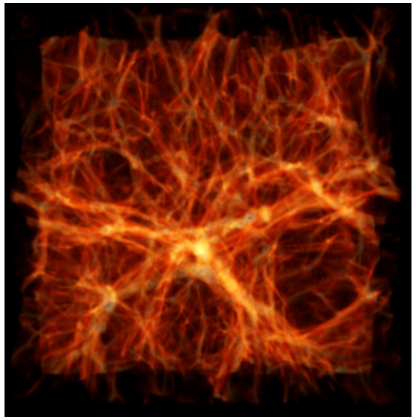
Splotch

[1]

- Customised, Volume Ray-casting algorithm for large datasets
- Main target is astrophysical N-body simulations
- Applicable to any data representable as point-like elements with attributes
- Particle contribution to image determined using radiative transfer equation and a Gaussian distribution function



[2]



[3]

[1] 3D Modelling and Visualization of Galaxies By B.Koribalsky, C.Gheller and K.Dolag (ATNF, Australia, ETH-CSCS, Switzerland, Univ. Observatory Munich, Germany)
[2] Filaments Connecting Galaxy Clusters By K.Dolag and M.Reineke (Univ. Observatory Munich, Max Planck Institute, Germany)
[3] Modified Gravity Models By Gong-Bo Zhao (Univ. of Portsmouth, UK)

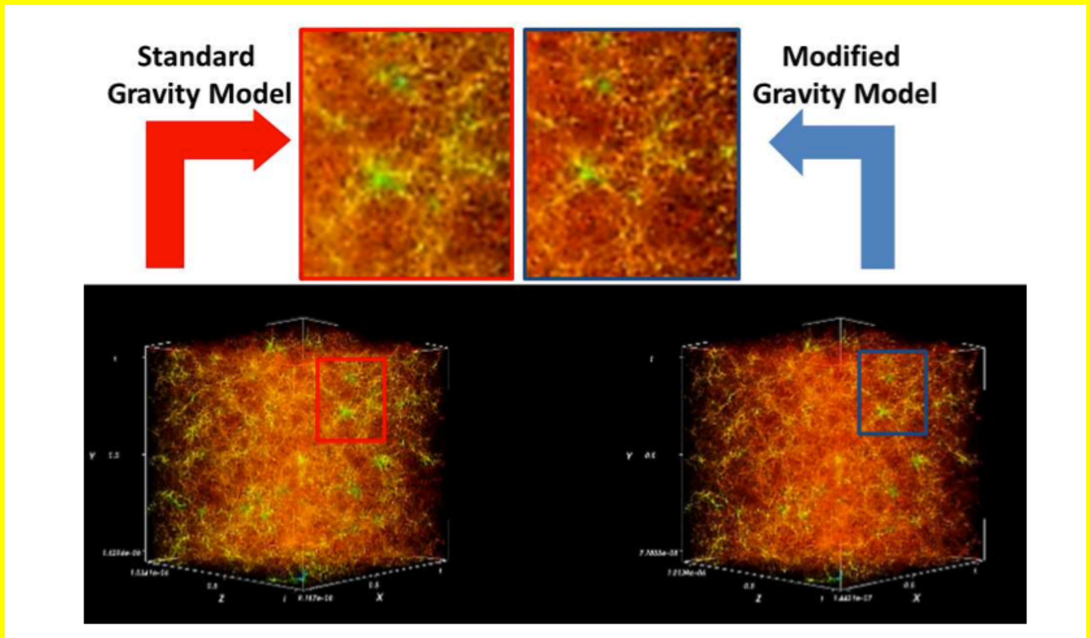
Modified Gravity

By Gong-Bo Zhao (ICG, Univ. of Portsmouth, UK)

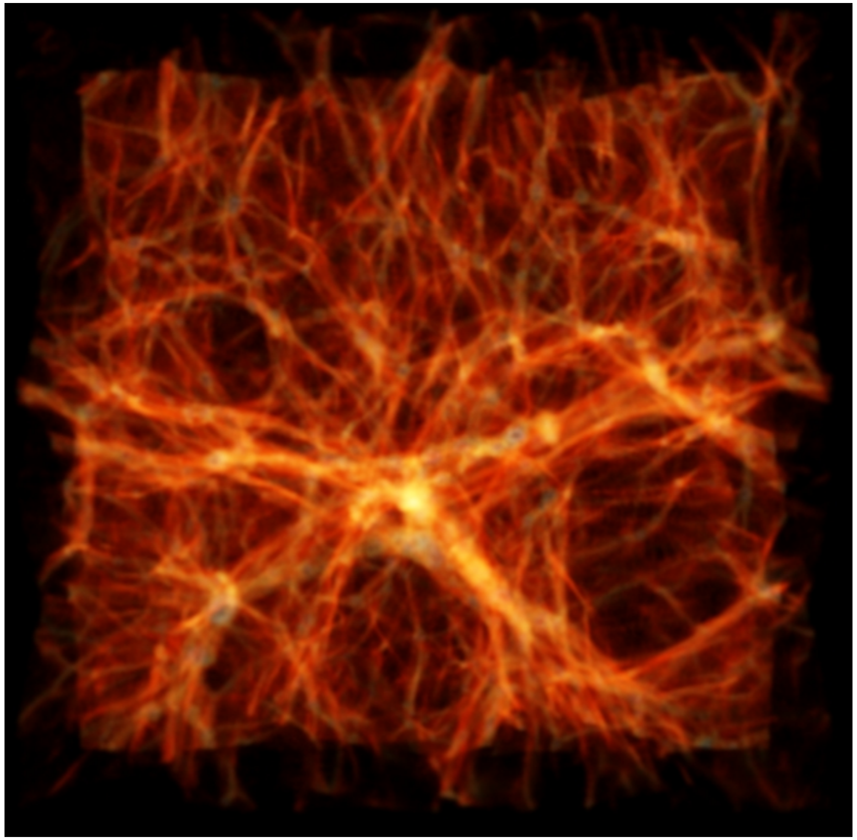
To understand the acceleration of the universe a new component is introduced, called dark energy, in the framework of General Relativity (GR).

Alternatively, *we can modify GR itself on cosmological scales* to realise the acceleration *without introducing dark energy*.

Observing the large scale structure of the universe could in principle *provide a new test of GR on cosmic scales.*



Ramses simulation
Mass distribution

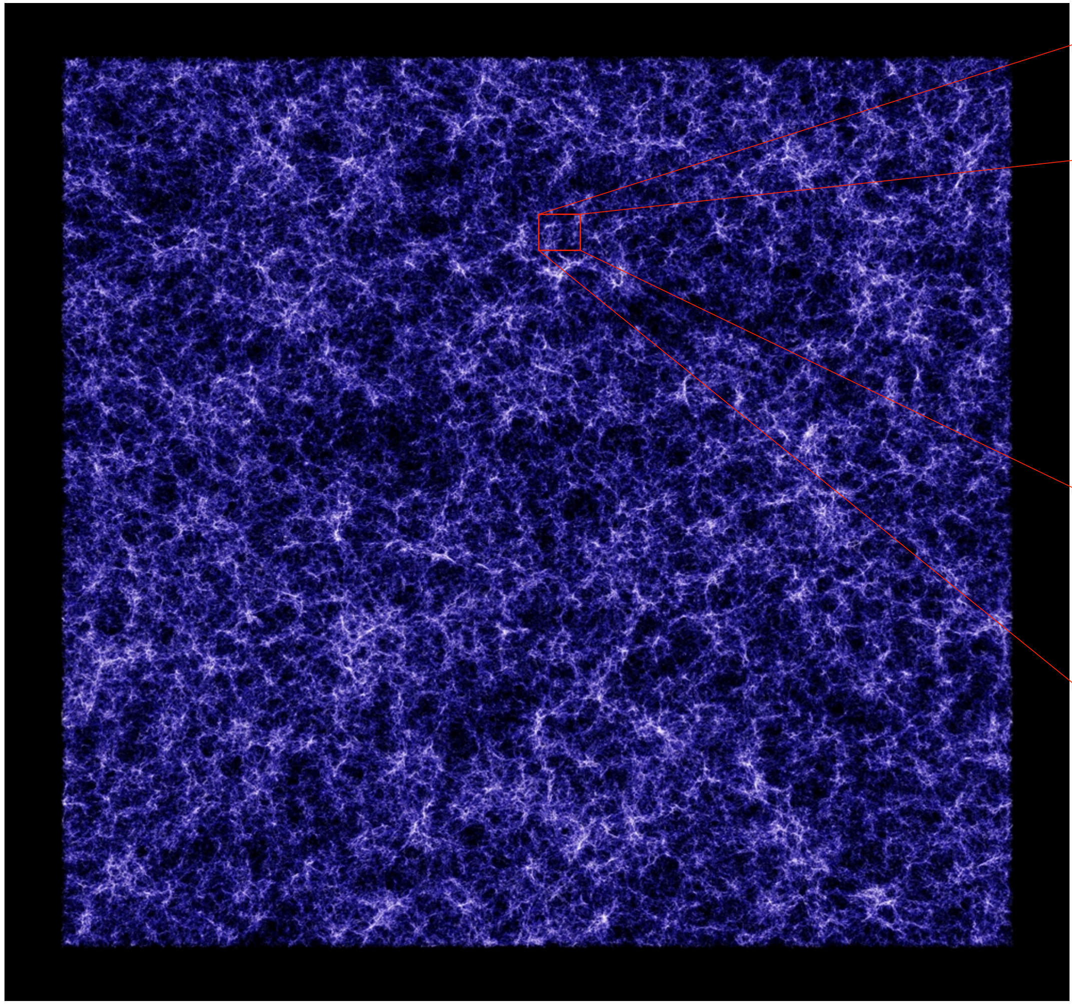


- Comparing different simulations
- Exploring datasets individually or simultaneously

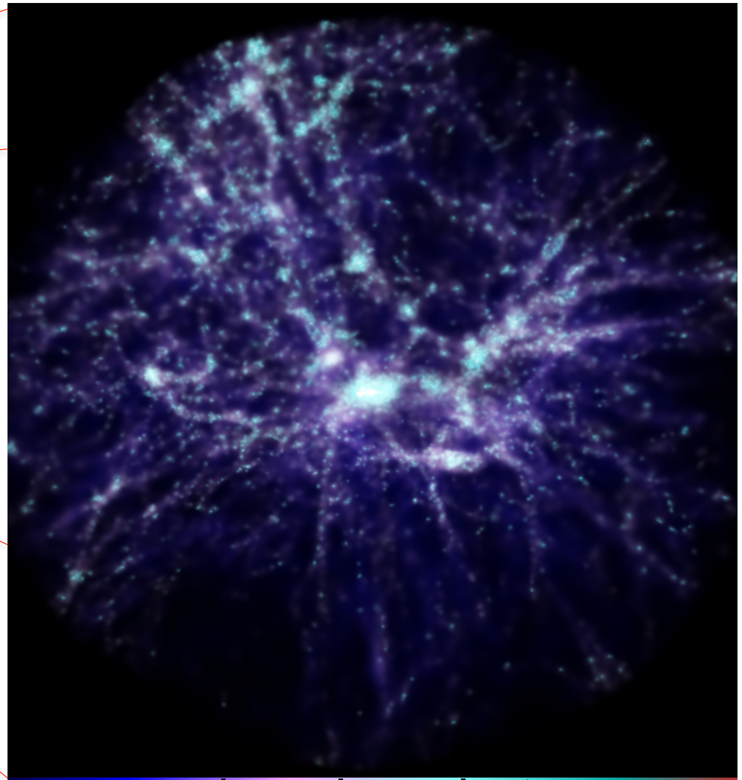
Large Scale Structures

By *A.D'Aloisio, H.Park, I.T.Iliev, P.R.Shapiro (University of Texas, Austin, USA)*

GADGET N-body cosmological simulation box: **5000^3 particles (**1.6TBytes per snapshot)



Dark matter mass density
image resolution: 10,000 × 10,000



zoom in: galaxy clusters

Images up to **32000^2 pixels** have been produced!

Ramses Cosmological Simulations

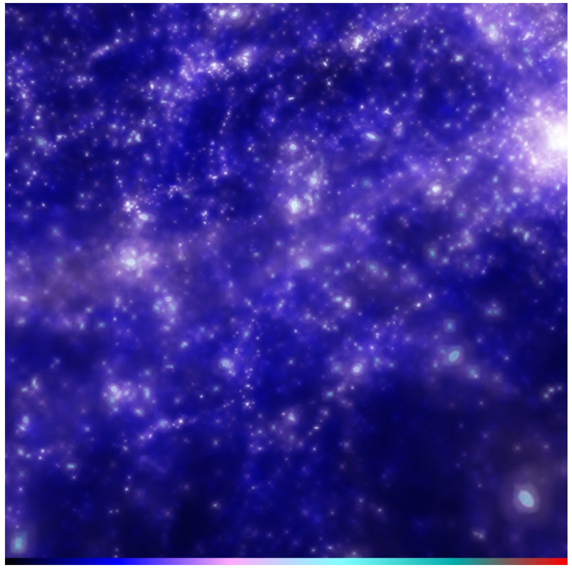
By *R. Teyssier (University of Zurich, Switzerland)*
http://www.itp.uzh.ch/~teyssier/ramses

Splotch can be easily deployed also for rendering of **mesh** or **AMR** based data.

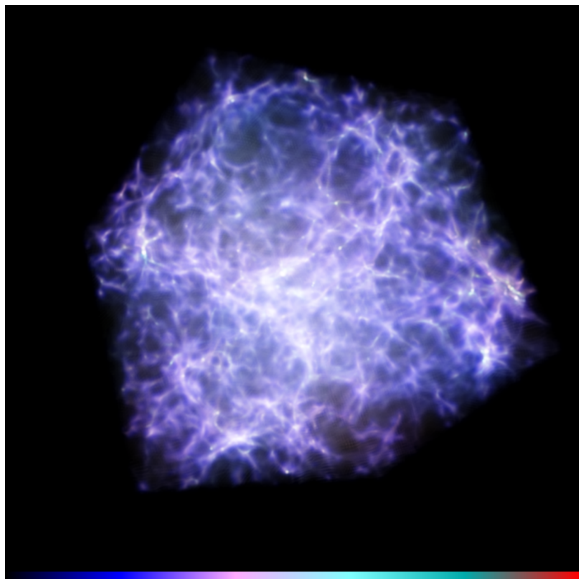
Ramses AMR reader:

- Navigate tree structure in AMR file
- Read grid positions & son indices from AMR hierarchy
- Read corresponding variables per cell from hydro file
- Generate particle for cells
- Randomize location within cell
- Set attributes based on user parameters

Combined method for visualizing both particle and AMR data.



mass density



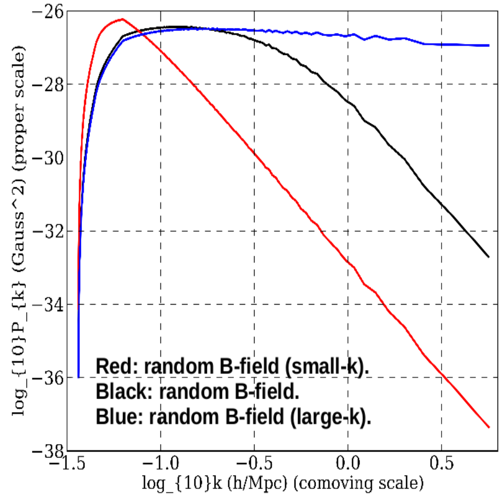
Magnetic Hydro Dynamics

By *Min-Su Shin, Julian Devriendt, Adrianne Slyz* *(University of Oxford, UK)*
http://www.physics.ox.ac.uk/users/msshin/research.htm

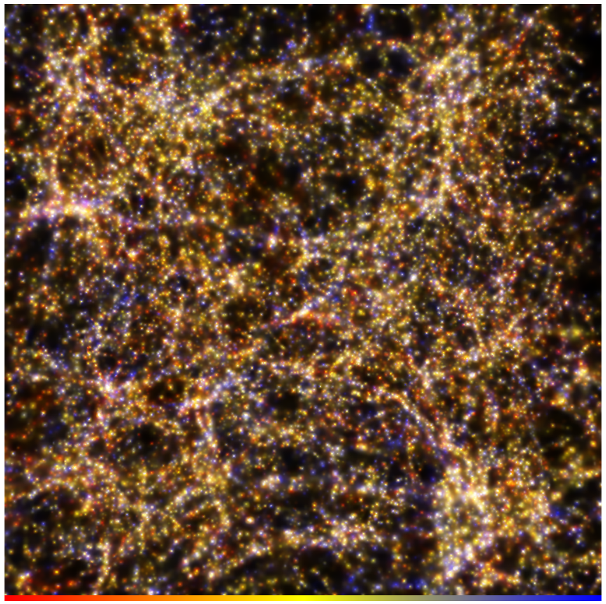
Magnetic field structures coloured according to cos(θ), where θ is the angle between velocity field vector and magnetic field vector in the grid cells.

Ramses MHD cosmological simulations
200 Mpc/h box,
AMR baryonic matter, 256^3 dark matter particles,
evolution from z=50 to z=0. Only gas is visualized.

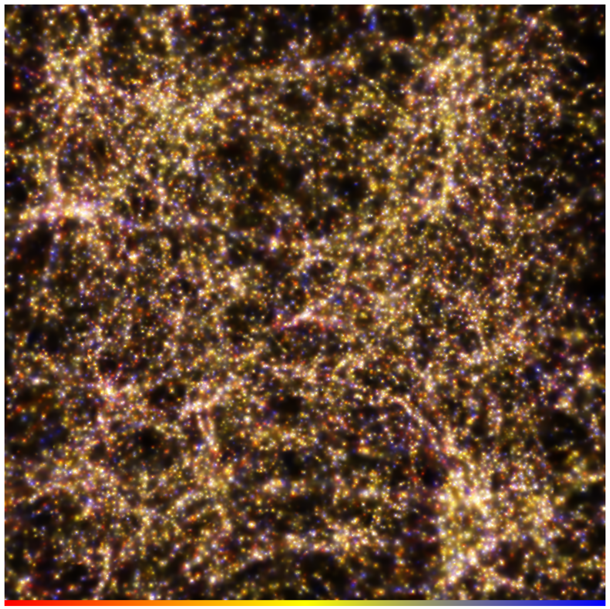
Power spectra of initial B-fields for 256^3 runs.



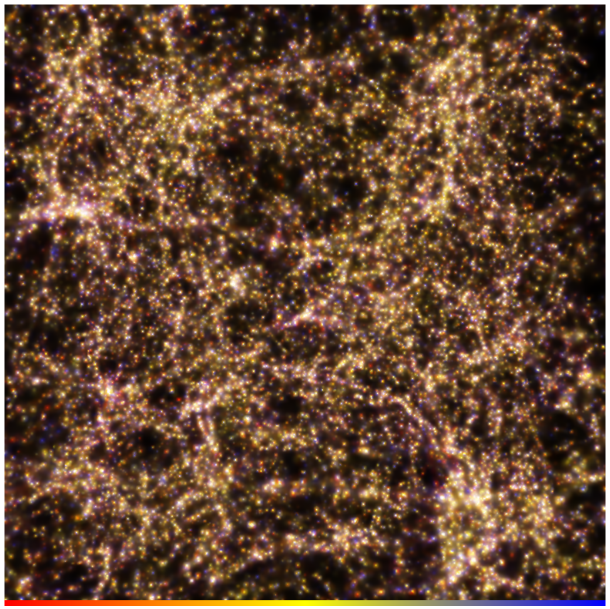
large *k* scales



fiducial run



small *k* scales

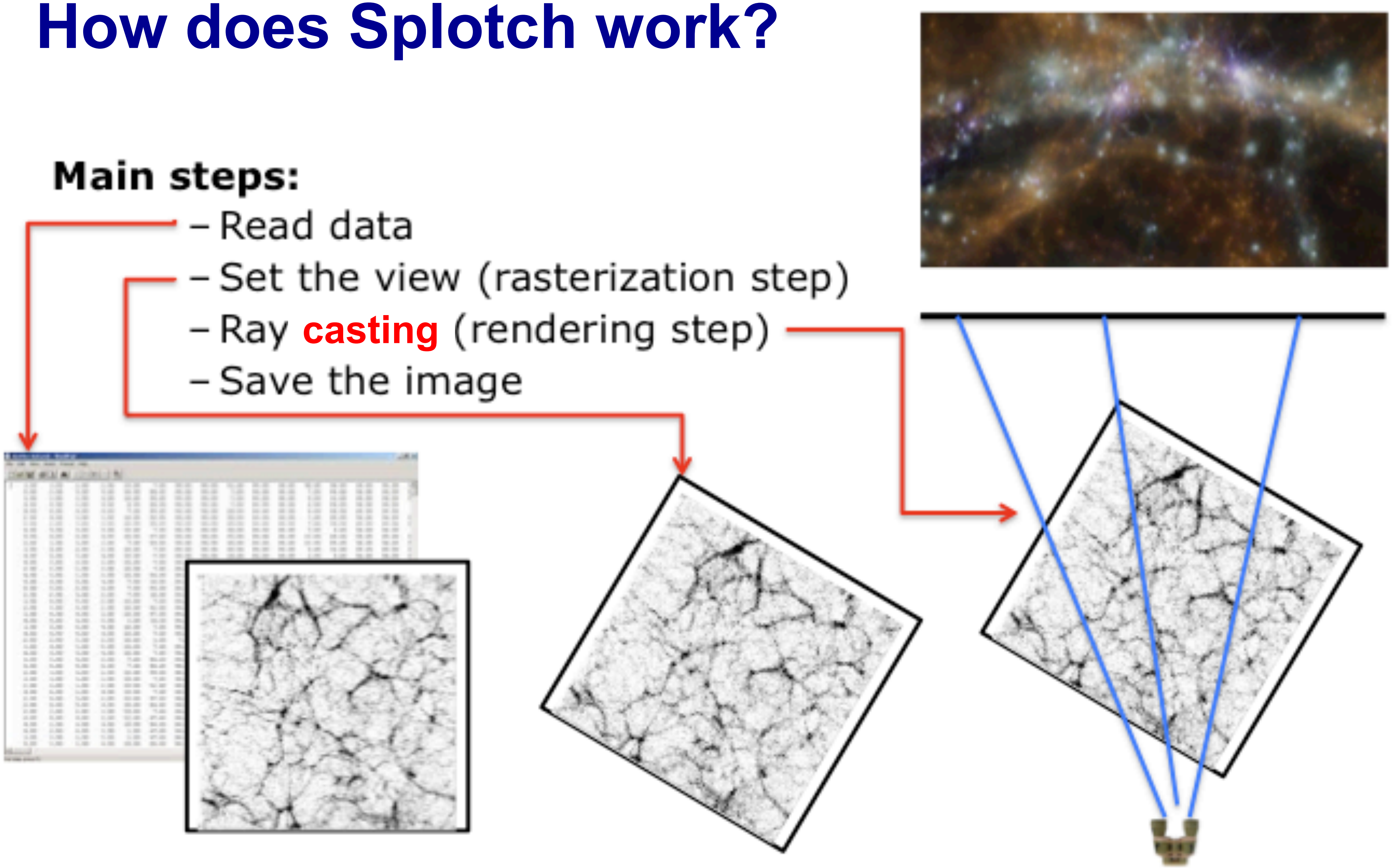


SPLOTCH ON GPUS

How does Splotch work?

Main steps:

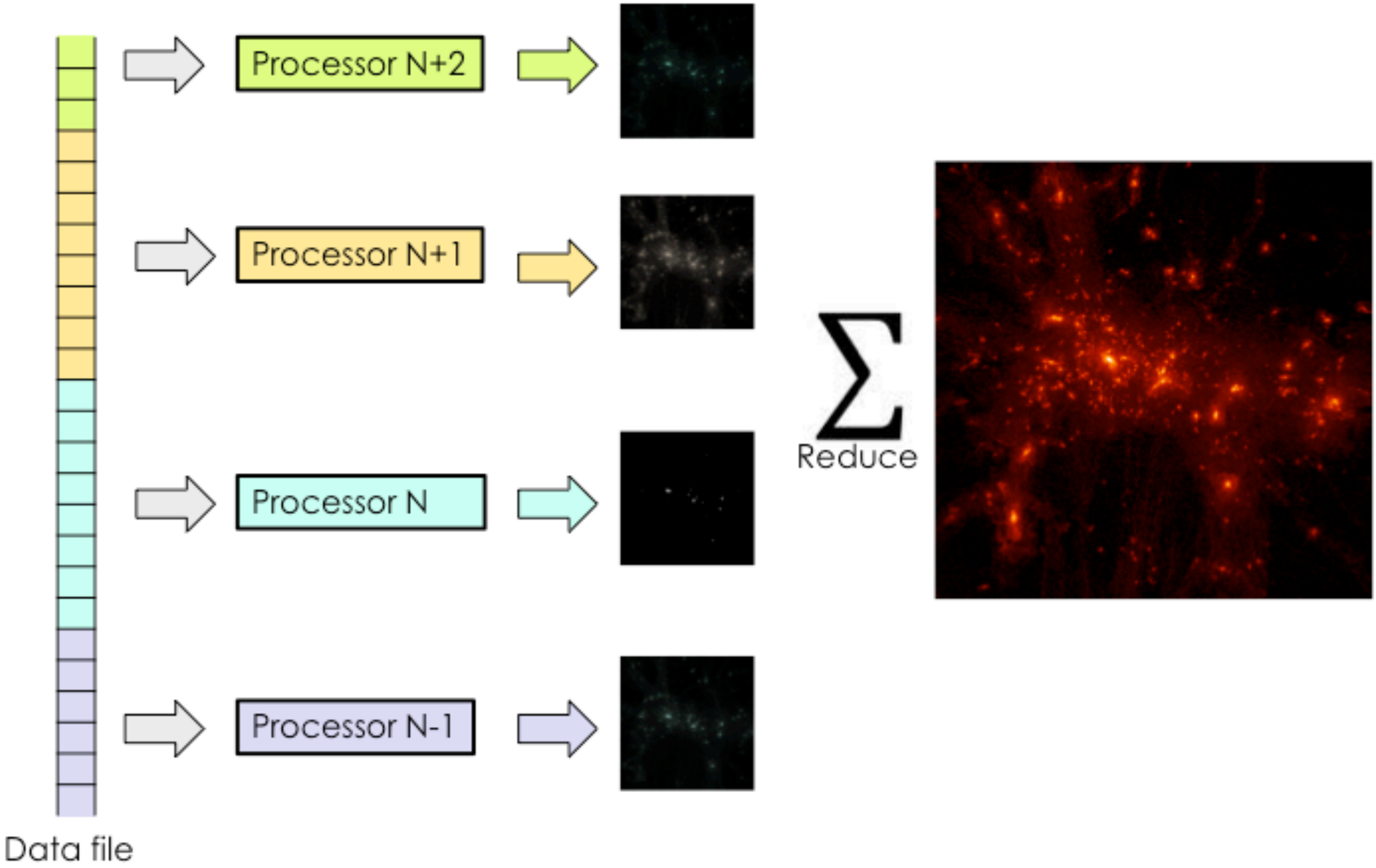
- Read data
- Set the view (rasterization step)
- Ray **casting** (rendering step)
- Save the image



Splotch on HPC Systems

- To exploit all available hardware in potentially hybrid HPC systems, Splotch has various complimentary implementations.
- For shared memory models, the data processing can be handled in parallel using OpenMP.
- For distributed memory models, the entire execution flow can be handled in parallel using MPI.
- In the case of hybrid systems, the data processing can be offloaded to GPU and processed with CUDA.
- This allows us to maximise computational efficiency.

Splotch MPI Implementation



Our Development / Testing / Target System

"Piz Daint" CRAY XC30 system @ CSCS (N.6 in Top500)

Nodes:

5272 CPUs 8-core Intel SandyBridge equipped with:

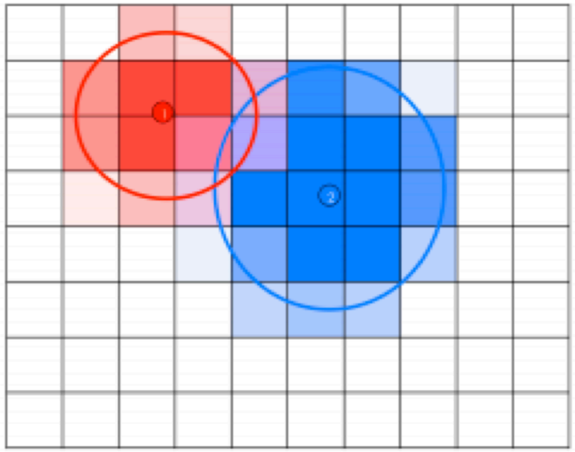
- 32 GB DDR3 memory
- One NVIDIA Tesla K20X GPU with 6 GB of GDDR5 memory

Overall system

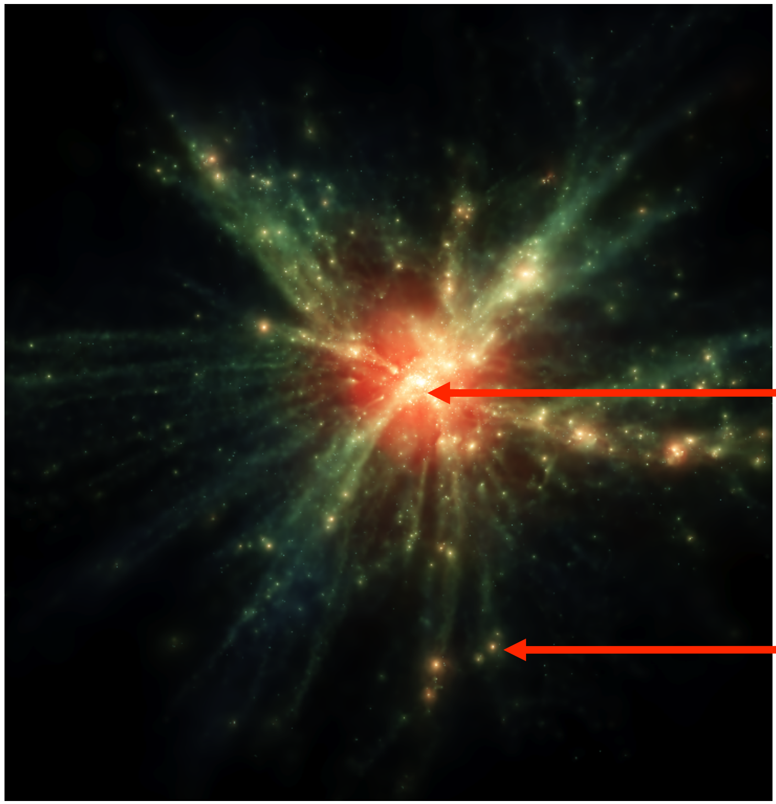
- 42176 cores and 5272 GPUs
- Memory: 170 (CPU)+32 (GPU) TBs
- Interconnect: Aries routing and communications ASIC, and Dragonfly network topology
- Peak performance: 7.787 Petaflops

Challenges for GPU Implementation

Potential race conditions due to single pixels potentially affected by many elements



Load balancing issues for data spread unevenly throughout the image



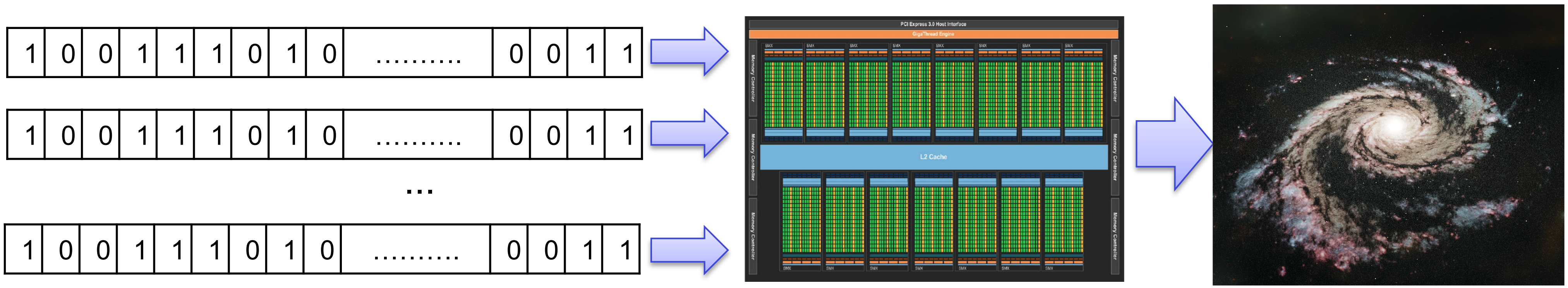
High concentration of particles in small area

Low concentration of particles spread across large area

GPU solution

The CUDA based approach consists of:

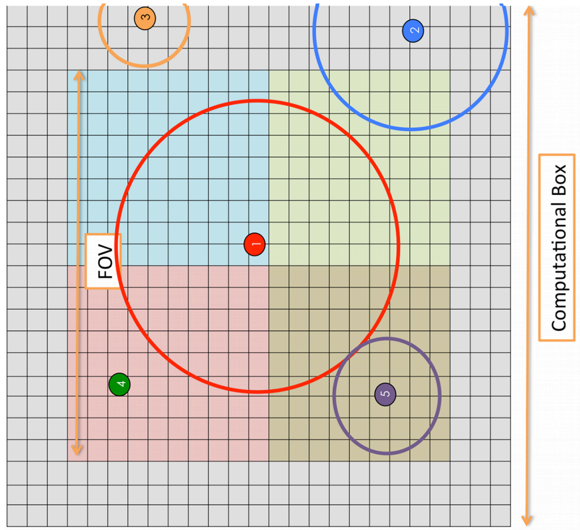
- Offloading particles to the GPU
- Rasterization: *one-thread-per-particle* approach
- Rendering: algorithm refactoring was necessary to address
  - unbalanced distribution of particles and
  - race conditions
- Copying back the final images to the CPU

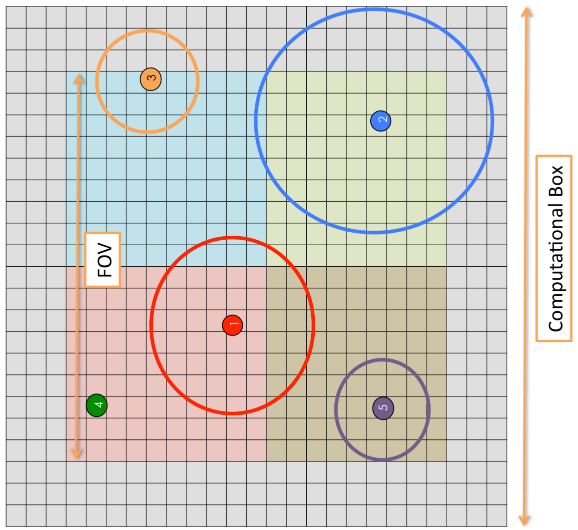


Rendering GPU implementation

Based on a particle classification :

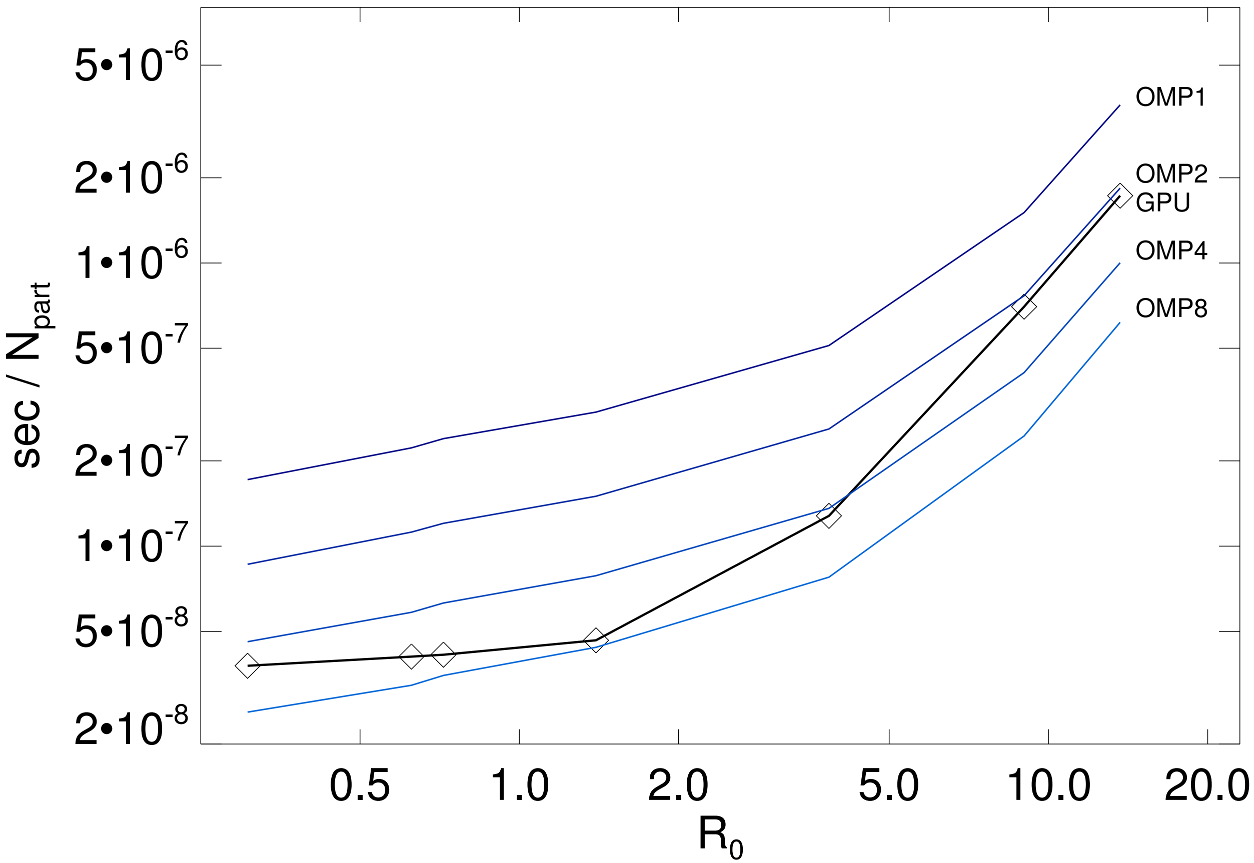
- **Small**
  - Processed on the GPU
  - Affect single pixel
- **Medium**
  - Processed on the GPU
  - Based on tiling of the image
  - Each particle influences only ONE tile
  - Additional algorithmic components necessary (“extra computation”)
- **Big**
  - Would affect more than one tile
  - Processed on the CPU, synchronously to the GPU
  - Data copy also overlapped to GPU computation





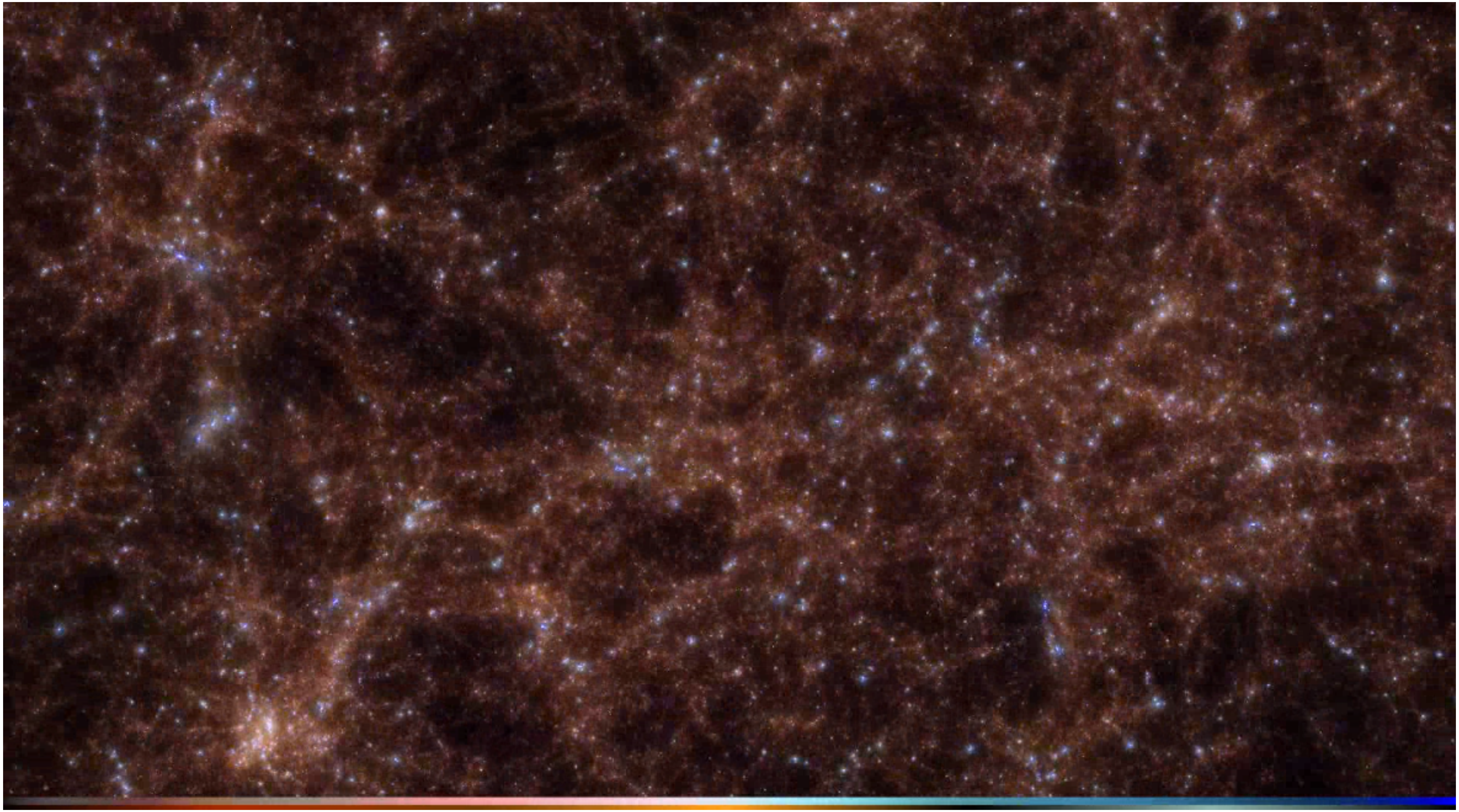
PERFORMANCE TESTS

Single Node Tests



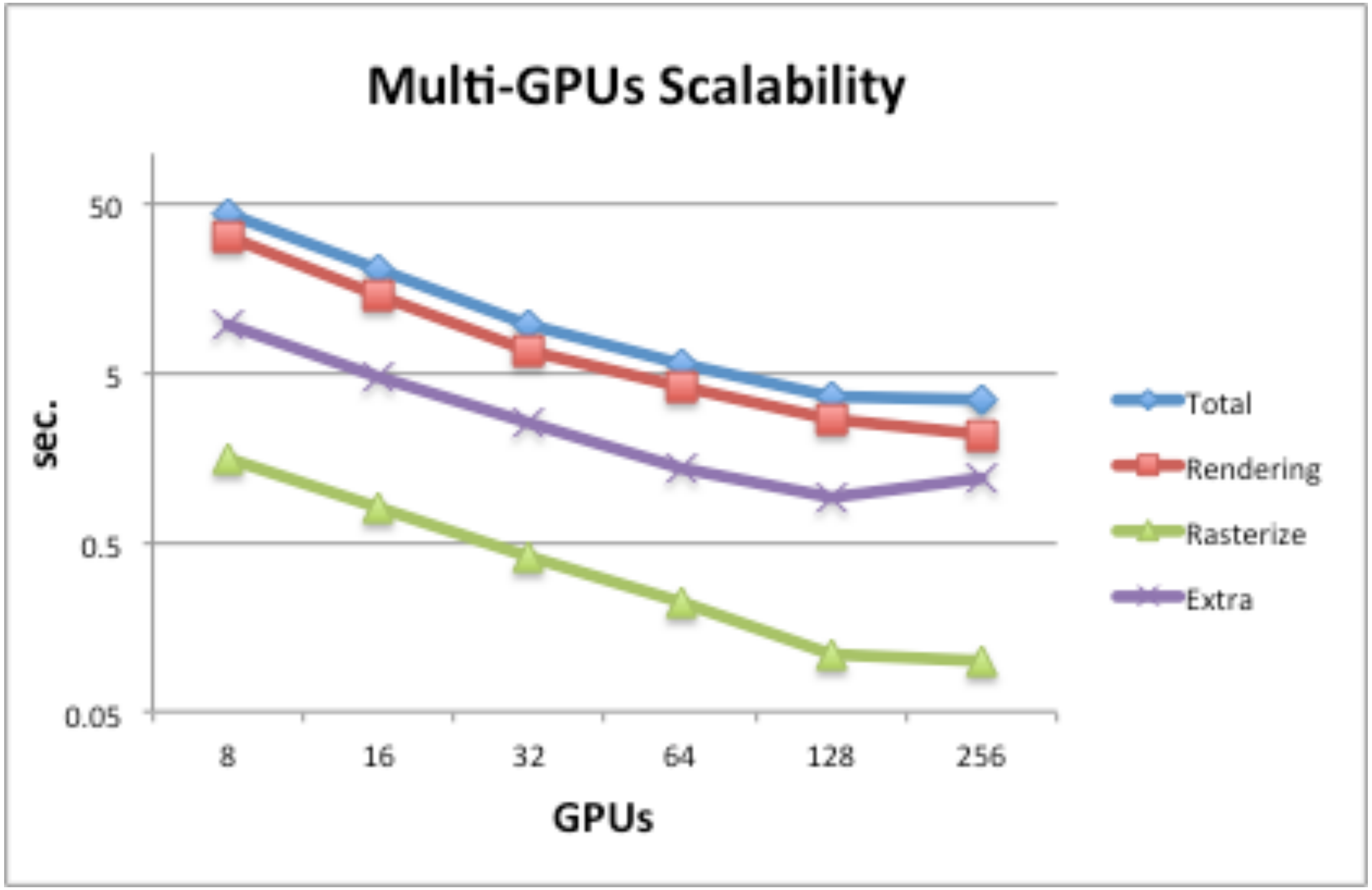
- GPU and CPU comparison in relation to the particles' radius
- **Black curve** represents GPU computing times.
- **Blue curves** are CPU times ranging from 8 cores (light blue, OMP8) to 1 core (dark blue, OMP1)

The Magneticum Visualization



http://www.mpa-garching.mpg.de/~kdolag/Simulations/

Scalability (Magneticum)



- 6 billion particles
- Overall processing times ranging from 50 to ~5 secs

MPS

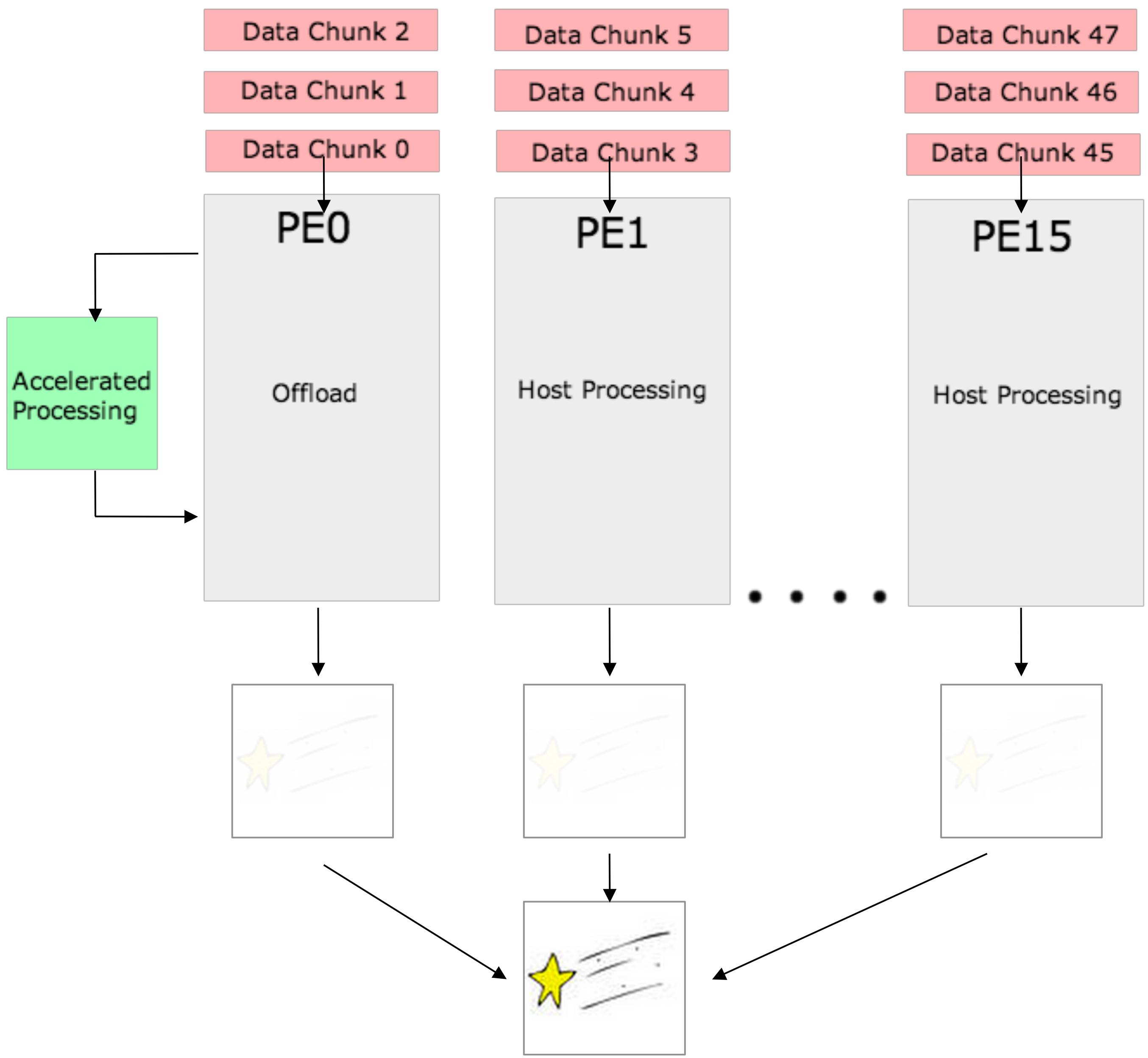
- enabling multiple MPI processes to run concurrently on the GPU
  - up to 32 MPI processes with variable workloads
- Each MPI process can be assigned to a different hardware work queue
  - maximizing GPU utilization and increasing overall performance

Cores	Computing Times	Speed-up
1	10.74	
2	7.654	1.40
4	6.841	1.57

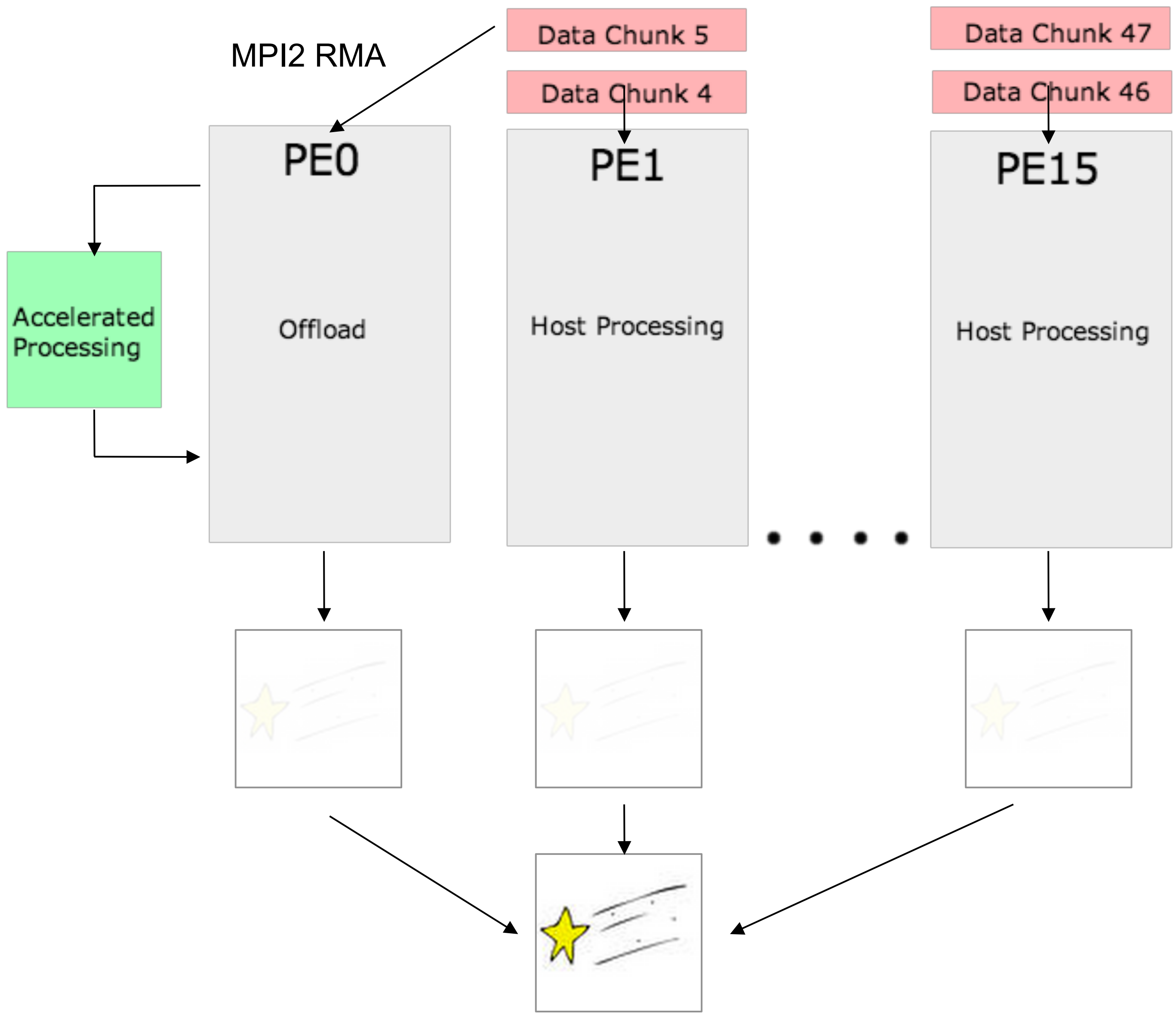
our results on a 200M particle simulation

FUTURE DEVELOPMENTS

Generic Offloading Model - 1



Generic Offloading Model - 2



References

K.Dolag, M.Reinecke, C.Gheller, S.Imboden, Splotch: Visualizing Cosmological Simulations, *arXiv:0807.1742 [astro-ph.IM]*

Z.Jin, M.Krokos, M.Rivi, C.Gheller, K.Dolag, M. Reinecke, High-Performance Astrophysical Visualization using Splotch, *Procedia Computer Science* 1 (2010) pp. 1775-1784

M.Rivi, C.Gheller, T. Dykes, M.Krokos, K.Dolag, GPU Accelerated Particle Visualization with Splotch, *arXiv: 1309.1114 [astro-ph.IM]*

https://github.com/splotchviz/splotch