Max-Planck-Institut für Astrophysik

MAX-PLANCK-GESELLSCHAFT

Understanding Galaxy Formation with the help of Peta-scale computing

NCSA

Ludwig Oser (Columbia)

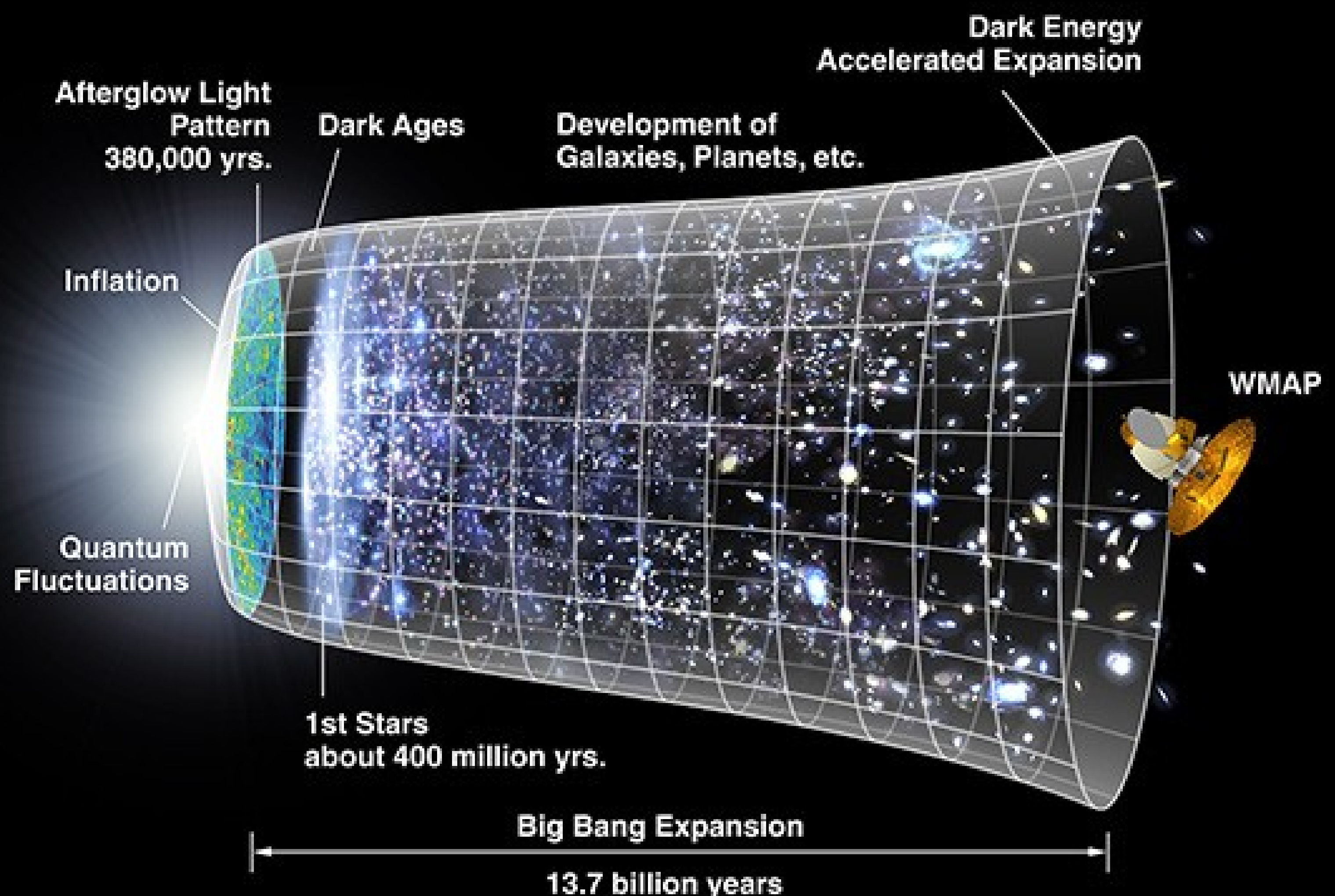
05/13/2014

Ken Nagamine (UNLV), Jeremiah P. Ostriker (Princeton), Thorsten Naab (MPA), Manisha Gajbe (NCSA), Greg Bryan (Columbia), Renyue Cen (Princeton)

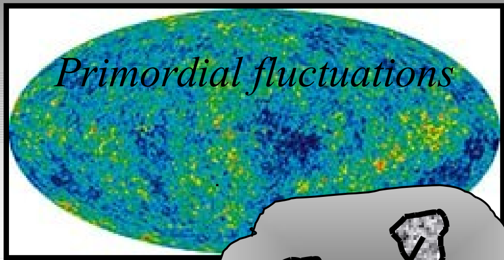
NSF

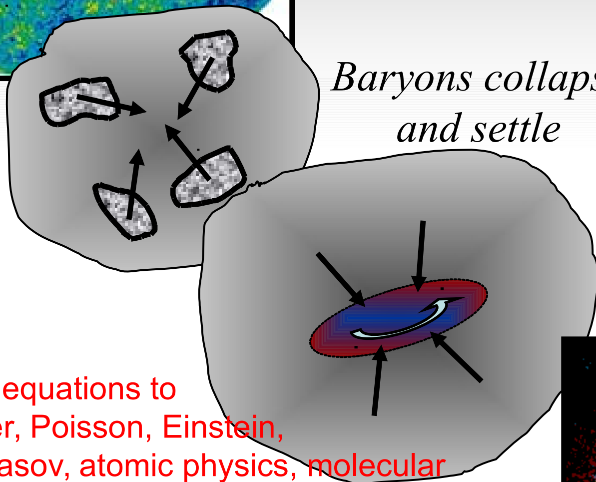
CAGE

Goal: simulating the nonlinear universe



Galaxy formation and evolution: a multi-scale, multi-physics, multi-astrophysics problem

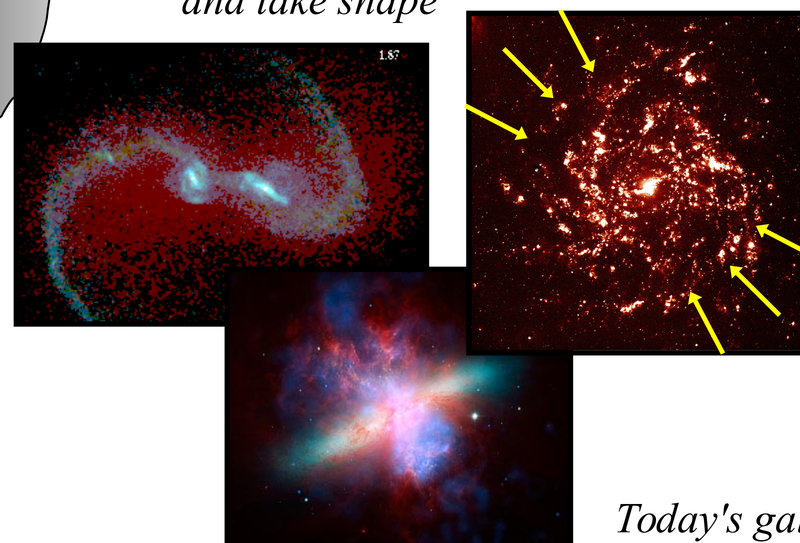




Baryons collapse and settle

From stars to large-scale cosmic web: 17 decades in mass dynamic range, 16 decades in spatial dynamic range

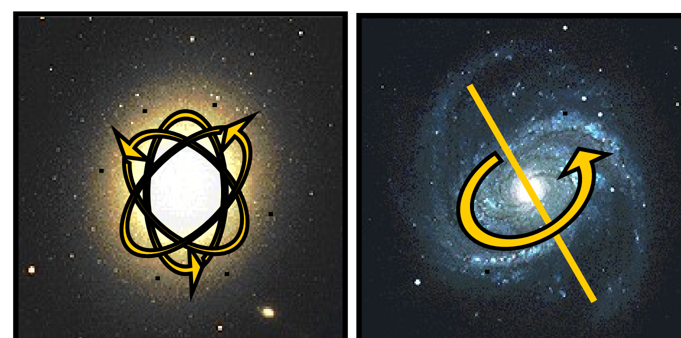
Galaxies assemble and take shape



A bunch of equations to Solve: Euler, Poisson, Einstein, Newton, Vlasov, atomic physics, molecular physics, …

- Baryonic mass accretion?
- Angular momentum?
- Timescales/mergers?
- Star formation efficiency?
- Inside-out galaxy formation?
- Connection between bulge and disk formation?
- Feedback AGN, SNII, SNIa etc.? Dissipation?
- Environment? Evolution as function of mass?
- Relation between dark matter and baryons?
- Assembly of galaxies?

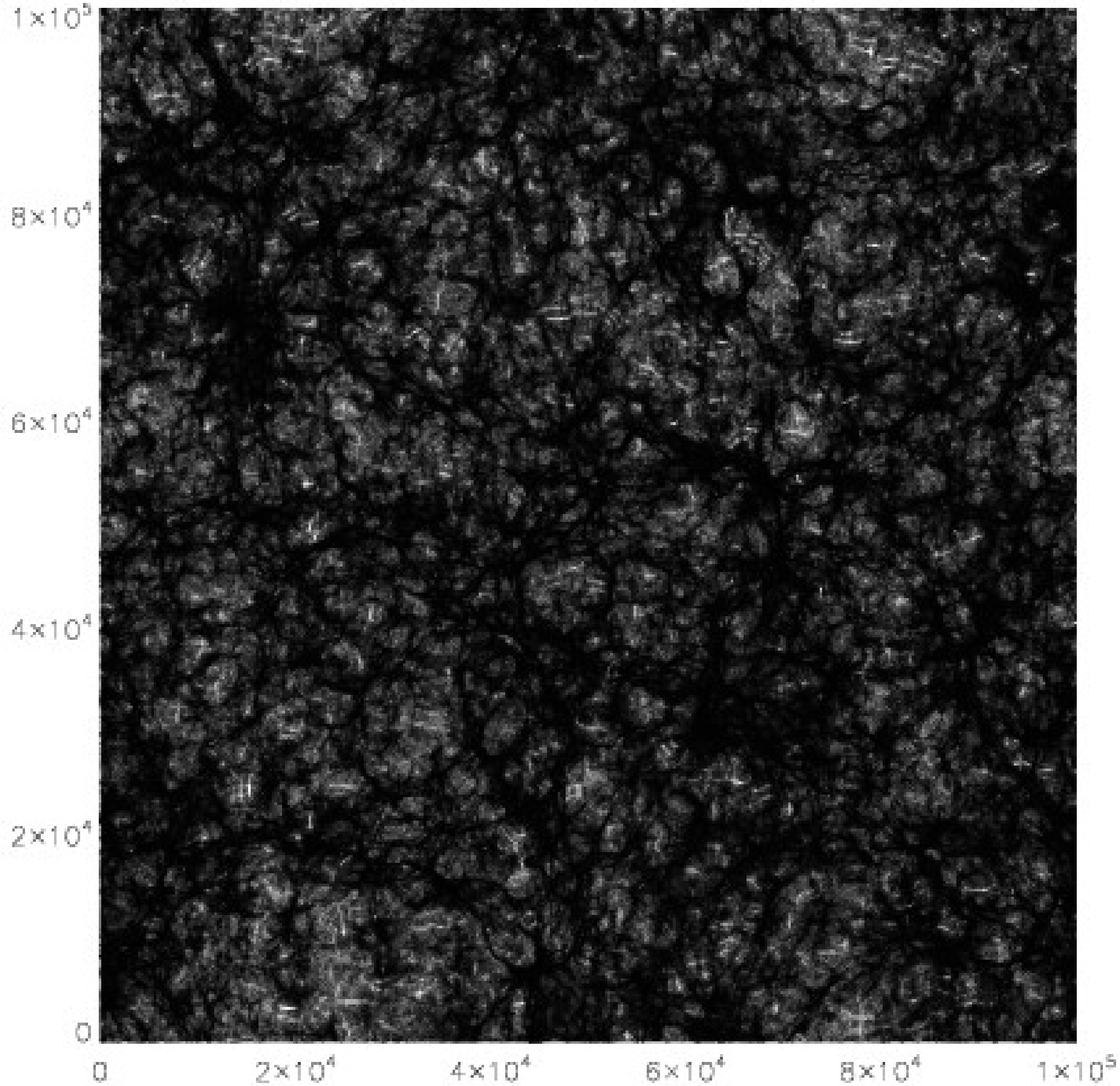
Today's galaxies



Images from NASA/WMAP Science Team and Hubble Heritage

z=6

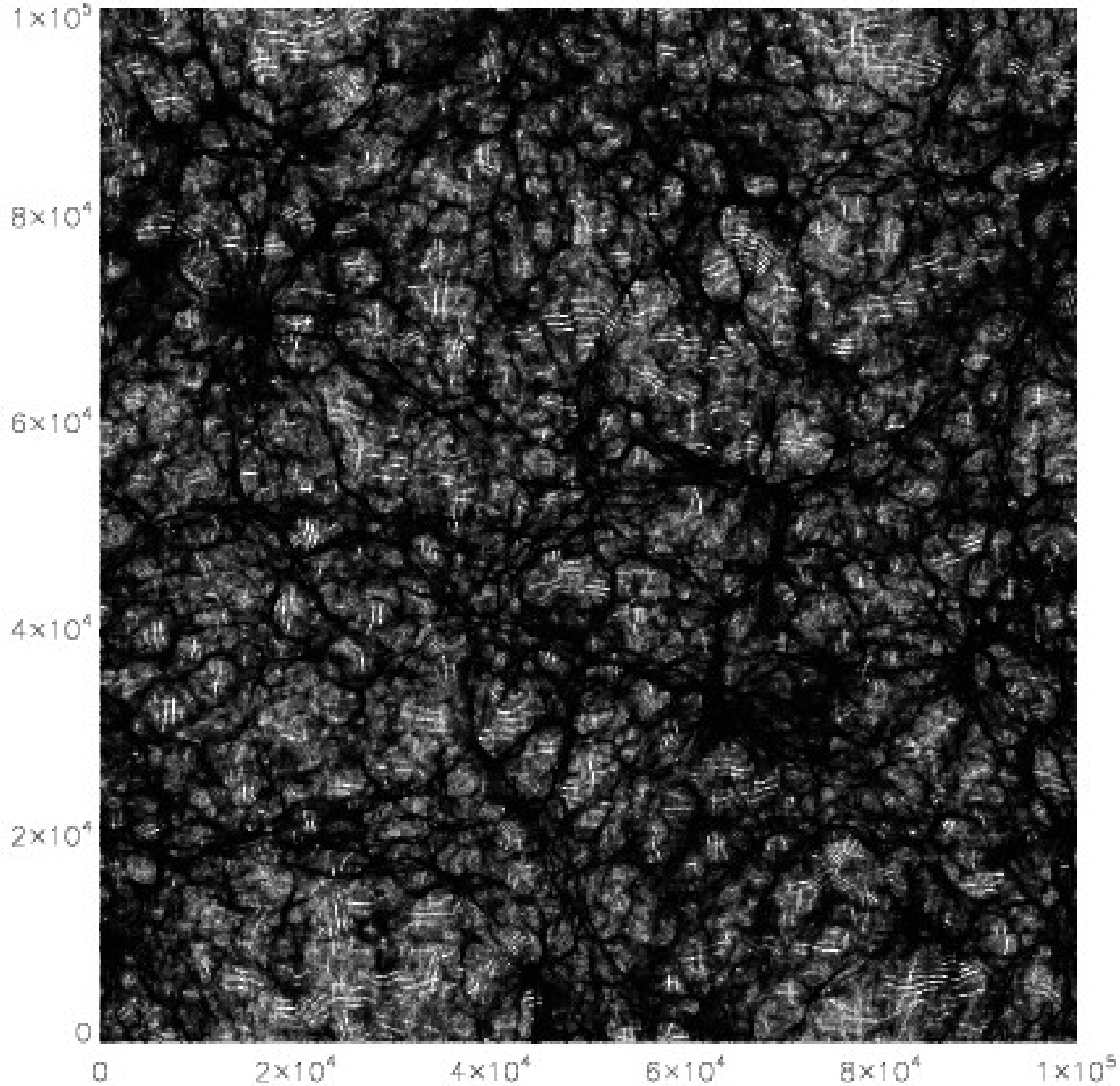
12.7 Gyr



100 Mpc = 3.3 * 10^8 ly

z=4

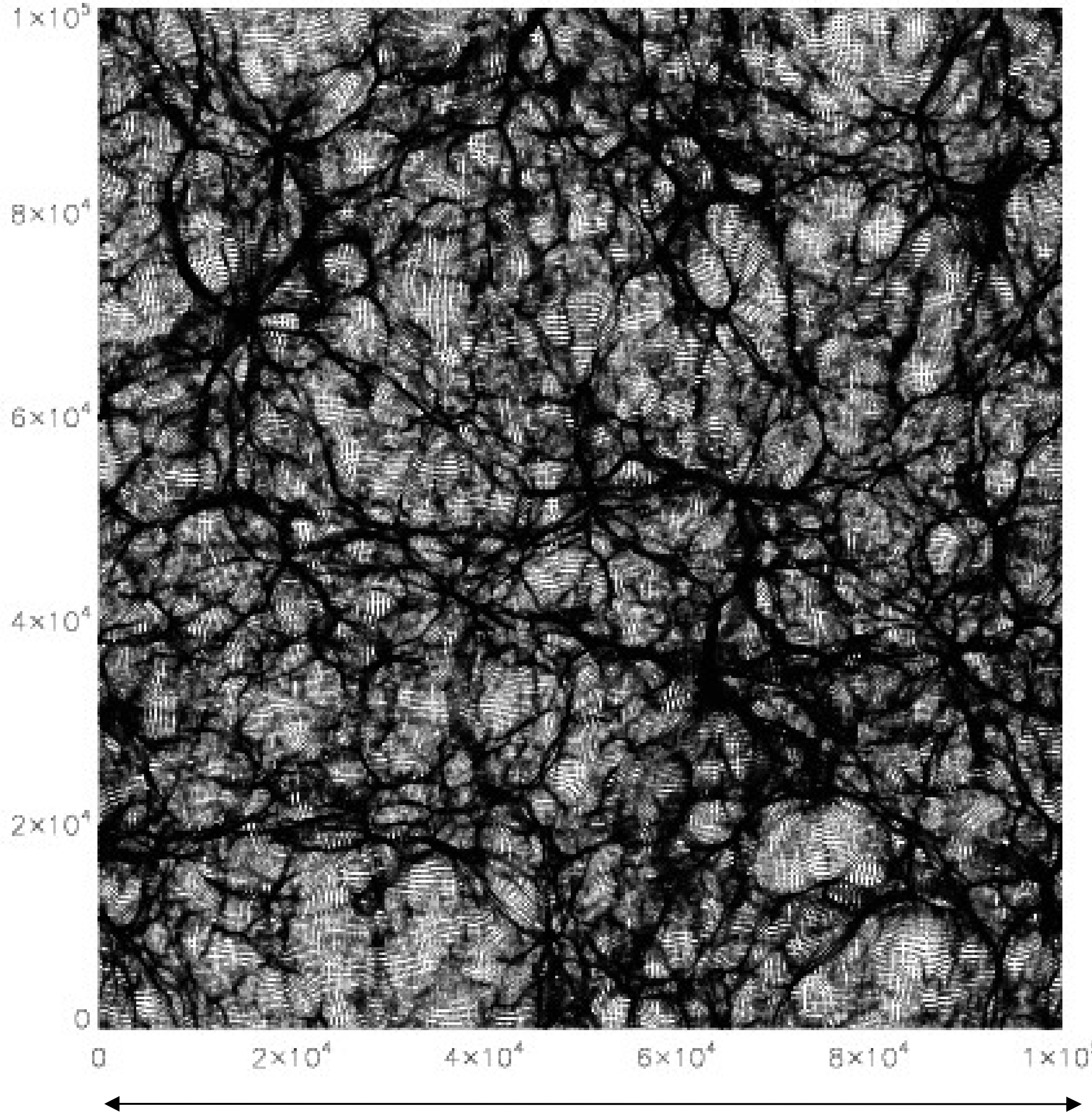
12.0 Gyr



100 Mpc = 3.3 * 10^8 ly

z=2

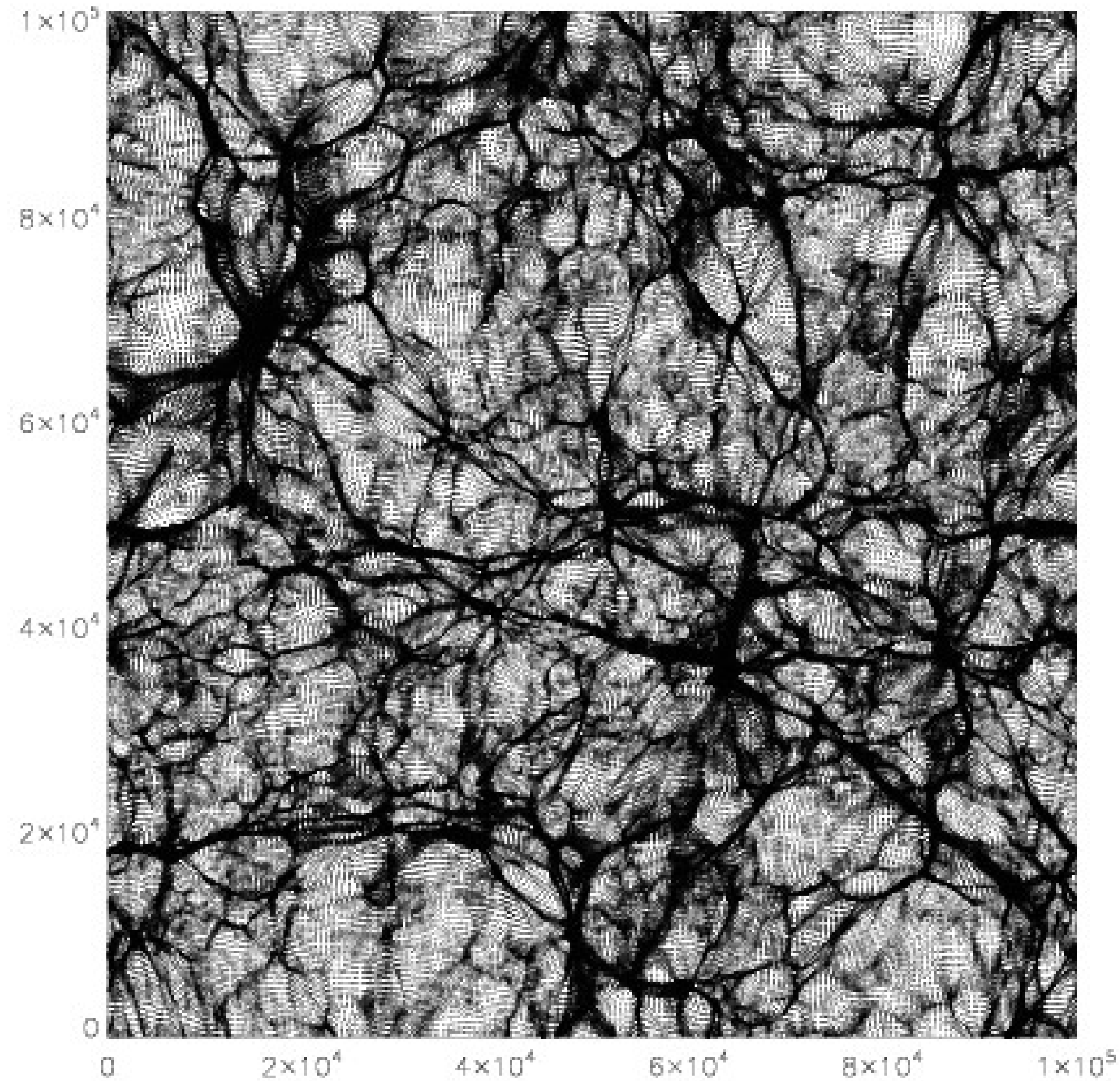
10.2 Gyr



100 Mpc = 3.3 * 10^8 ly

z=1

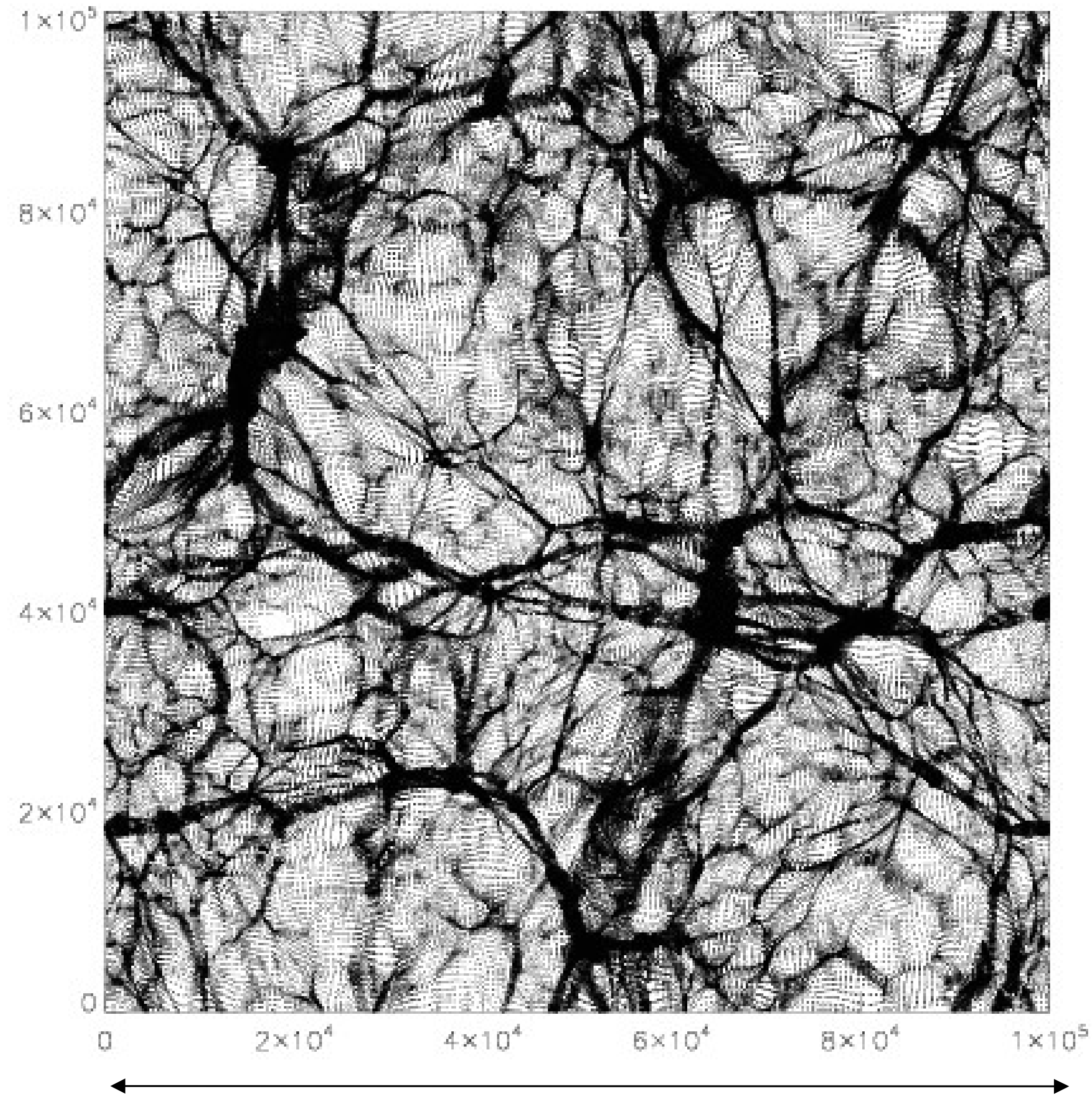
7.7 Gyr



100 Mpc = 3.3 * 10^8 ly

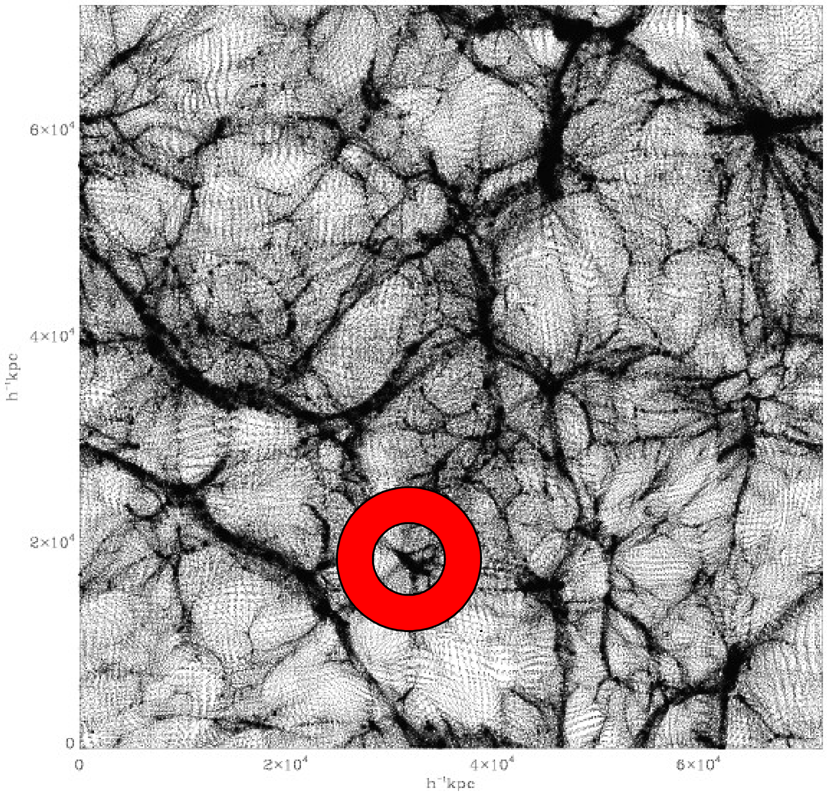
z=0

today



100 Mpc = 3.3 * 10^8 ly

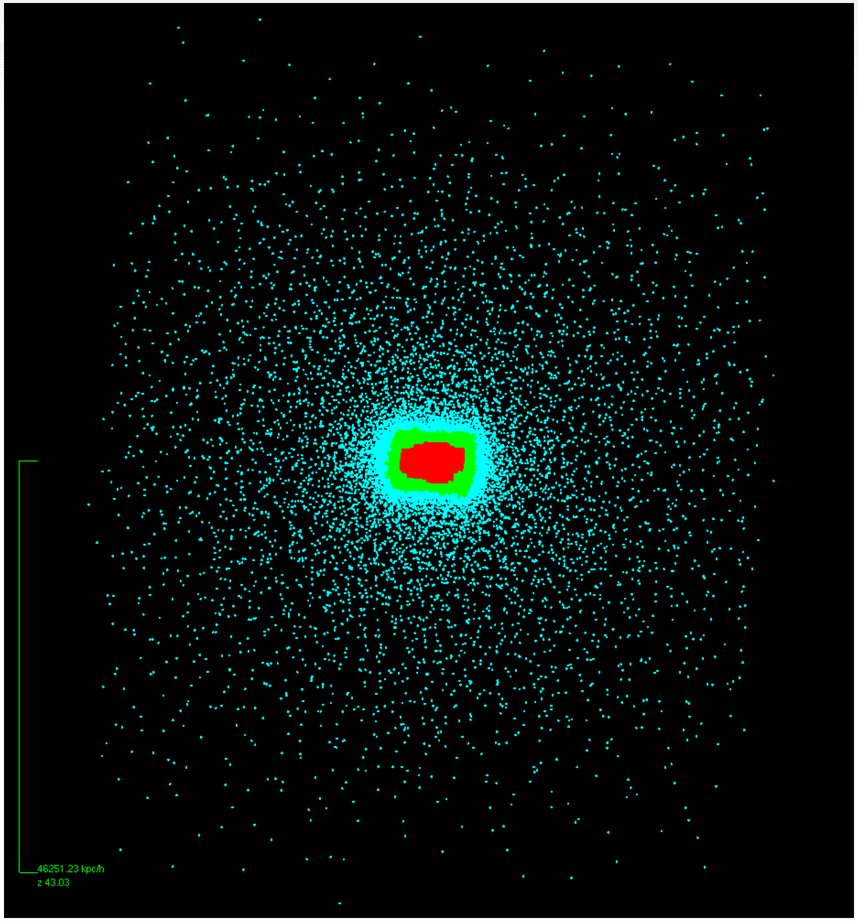
Zoom-in Resimulations



100^3 Mpc, 512^3 particles dark matter only, 100 snapshots (WMAP3: Ω_m = 0.26, Ω_Λ = 0.74, h = 0.72)

Trace back particles that will form a gravitationally bound structure at the present day

Zoom-in Resimulations

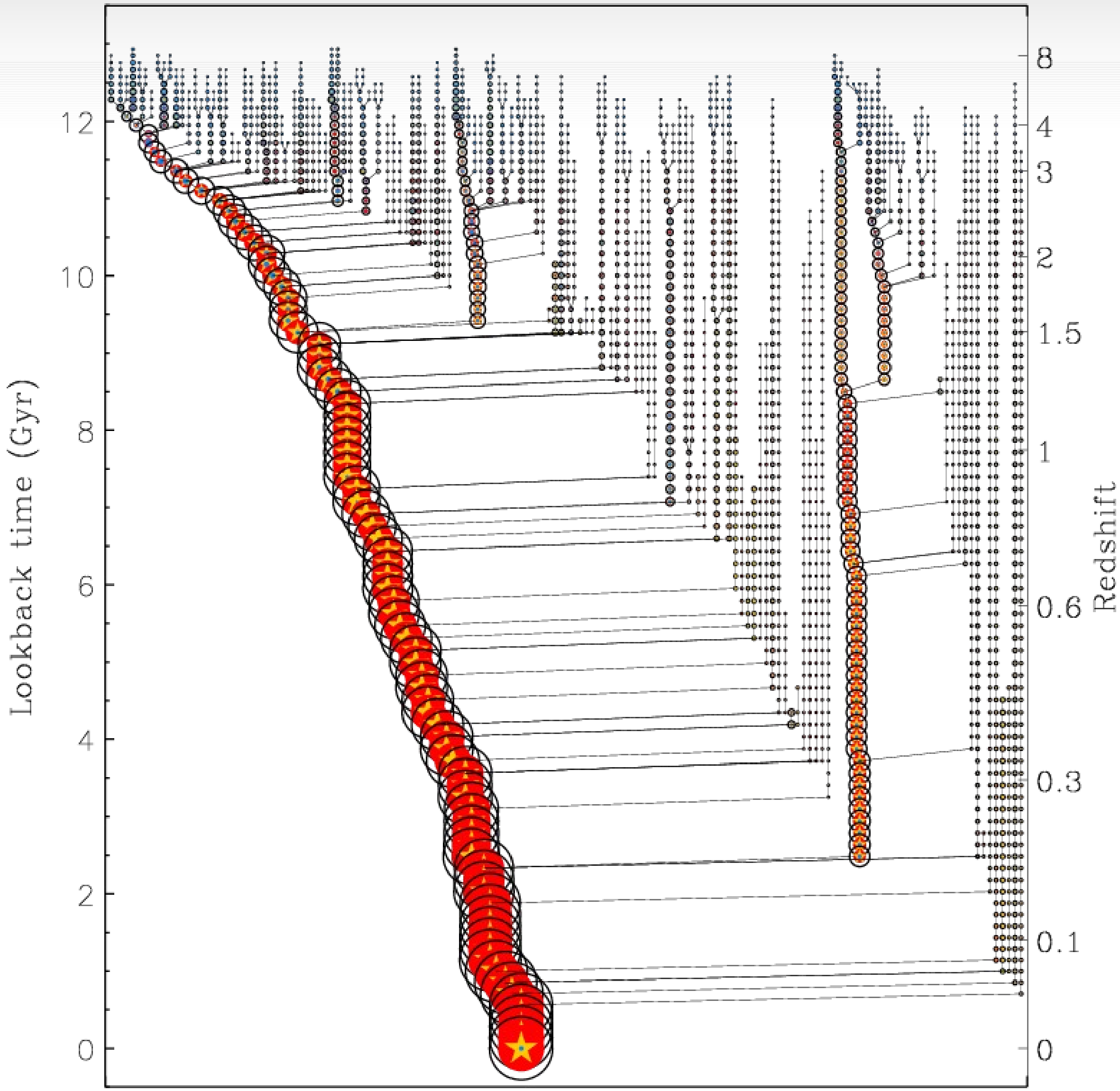


Particles are replaced with gas and dark matter particles at a higher resolution level

Simulation is redone including radiative cooling and star formation

Cosmological context is preserved!

Assembly history



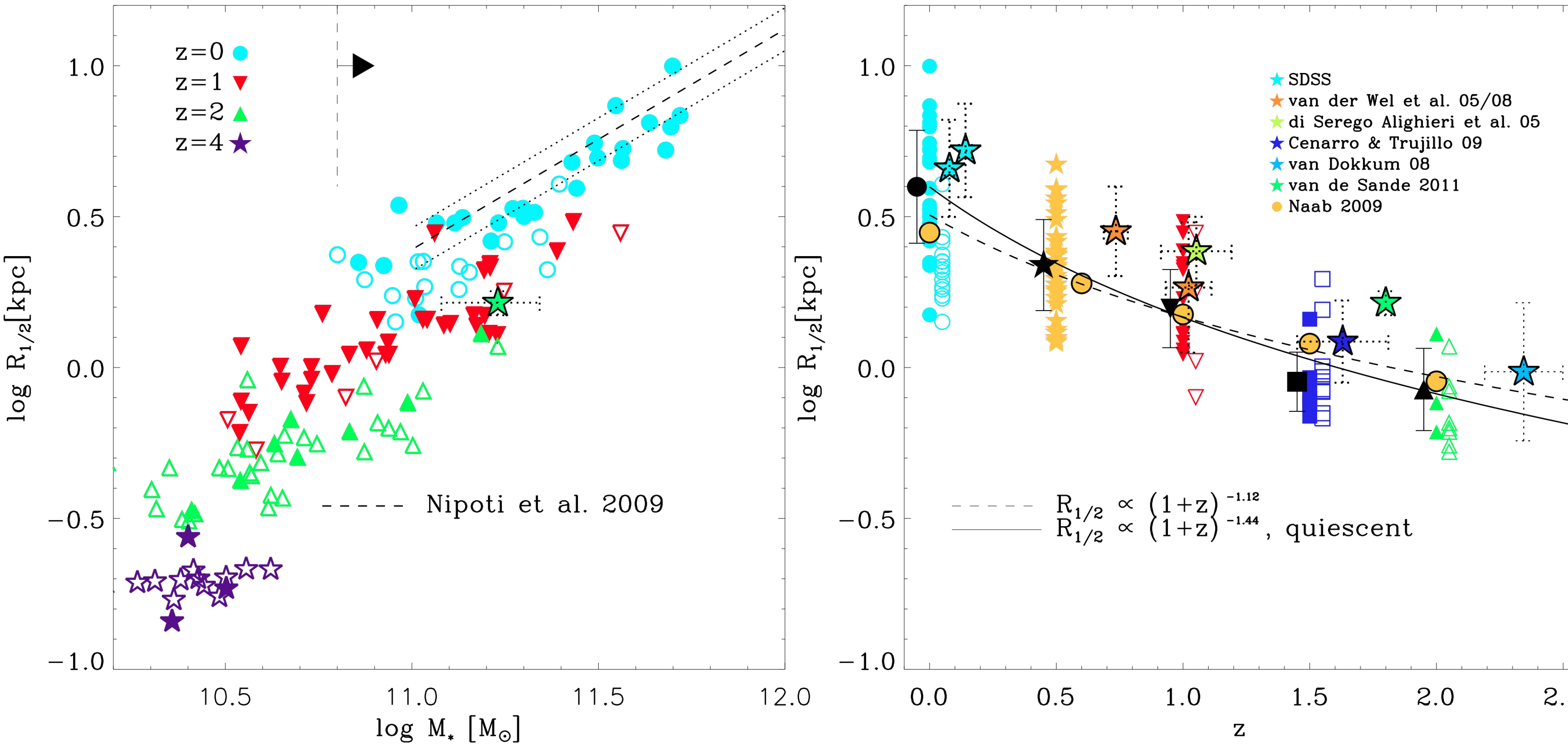
Intricate formation history

Hirschmann et al. 2012

Zoom-in

- Largest sample of cosmological zoom-in simulations so far (up to 10^{14} M_{SUN})
- Successful in explaining present-day properties of galaxies (Sizes, LOSVD, age distribution of stars, kinematics…)
- Still limited number count when compared to observations (difficult to compare scatter or subsamples)

Example: Size evolution



(Oser et al. 2012)

Statistics

- SDSS: spectra of nearly one million local galaxies
- Need to find agreement in statistical properties of the galaxy population
- Two approaches:

Full-box simulation <> ensemble computing

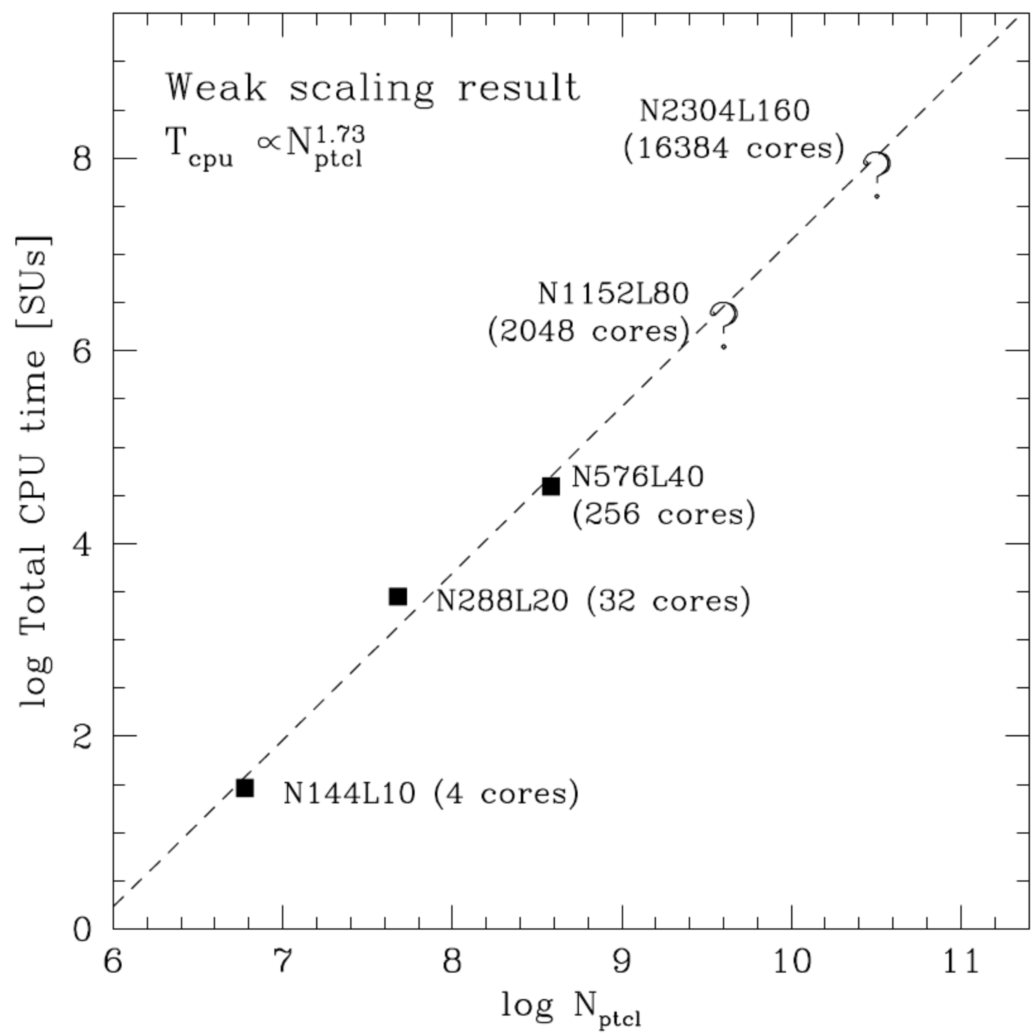
Pros and Cons

- Cons: No “sight lines” through simulation volume
- Long-range baryonic effects, e.g. reionization of the universe
- Cons: Much higher resolution possible in zoom-in simulations
- Different models can be tested (SNe, AGN)

Full-box scaling

- Including SPH
- Full-box simulations:

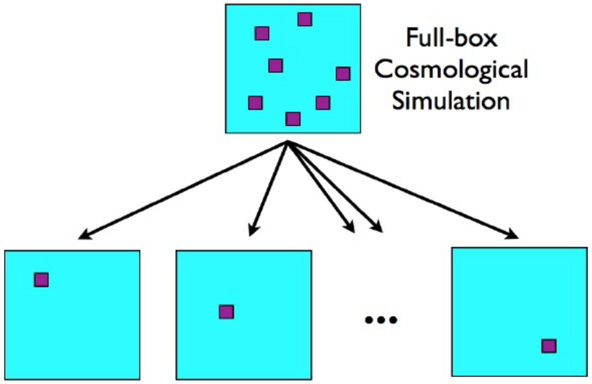
  $T_{cpu} \sim N^{1.73}$



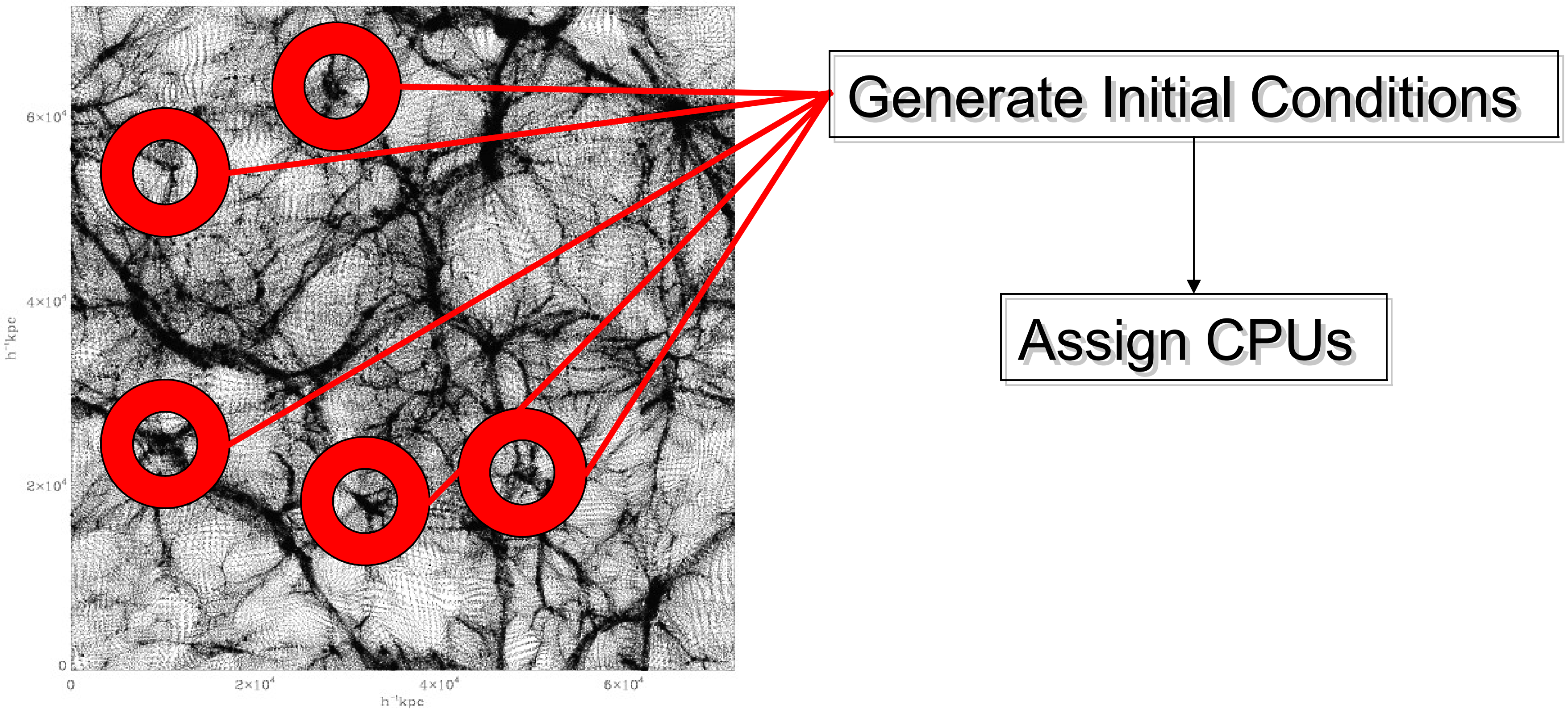
HECA

- Hierarchical Ensemble Computing Algorithm
- Embarrassingly parallel problem:

  Instead of increasing the number of processors with the problem size, the number of simulations is increased, i.e. $T_{CPU} \sim N^1$
- Overhead for having to resimulate the background is negligible

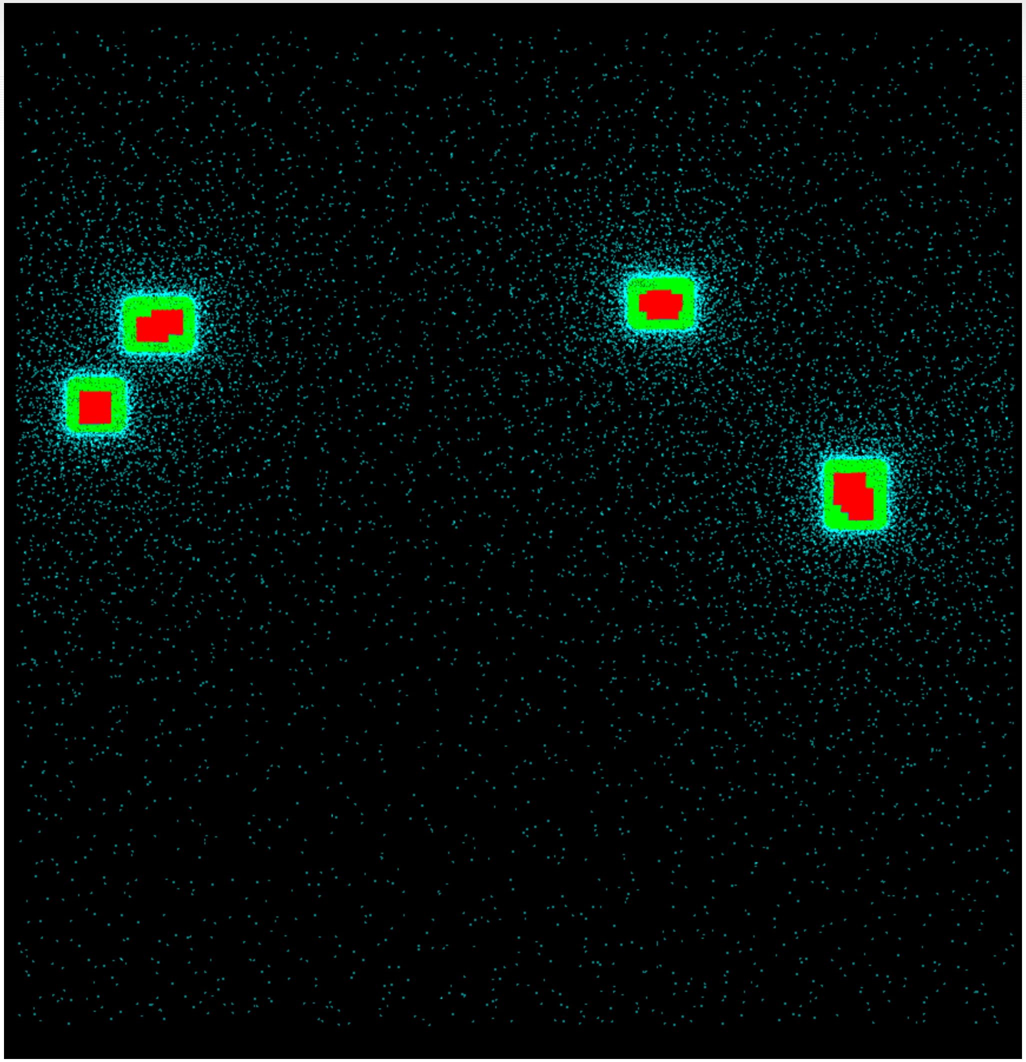


Scheduler



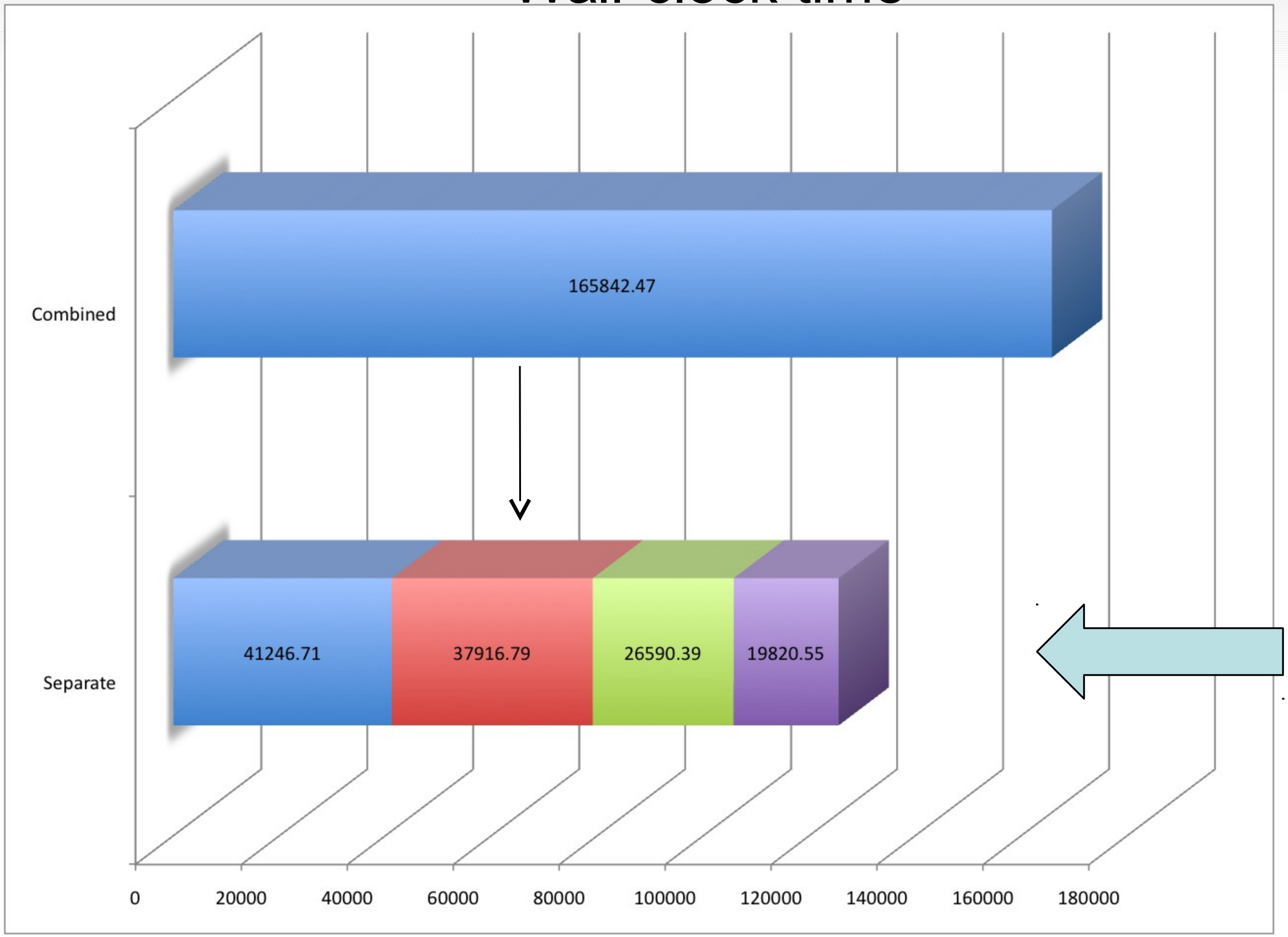
Generate Initial Conditions
Assign CPUs
6×10⁴
4×10⁴
2×10⁴
0
h⁻¹kpc

Scalings



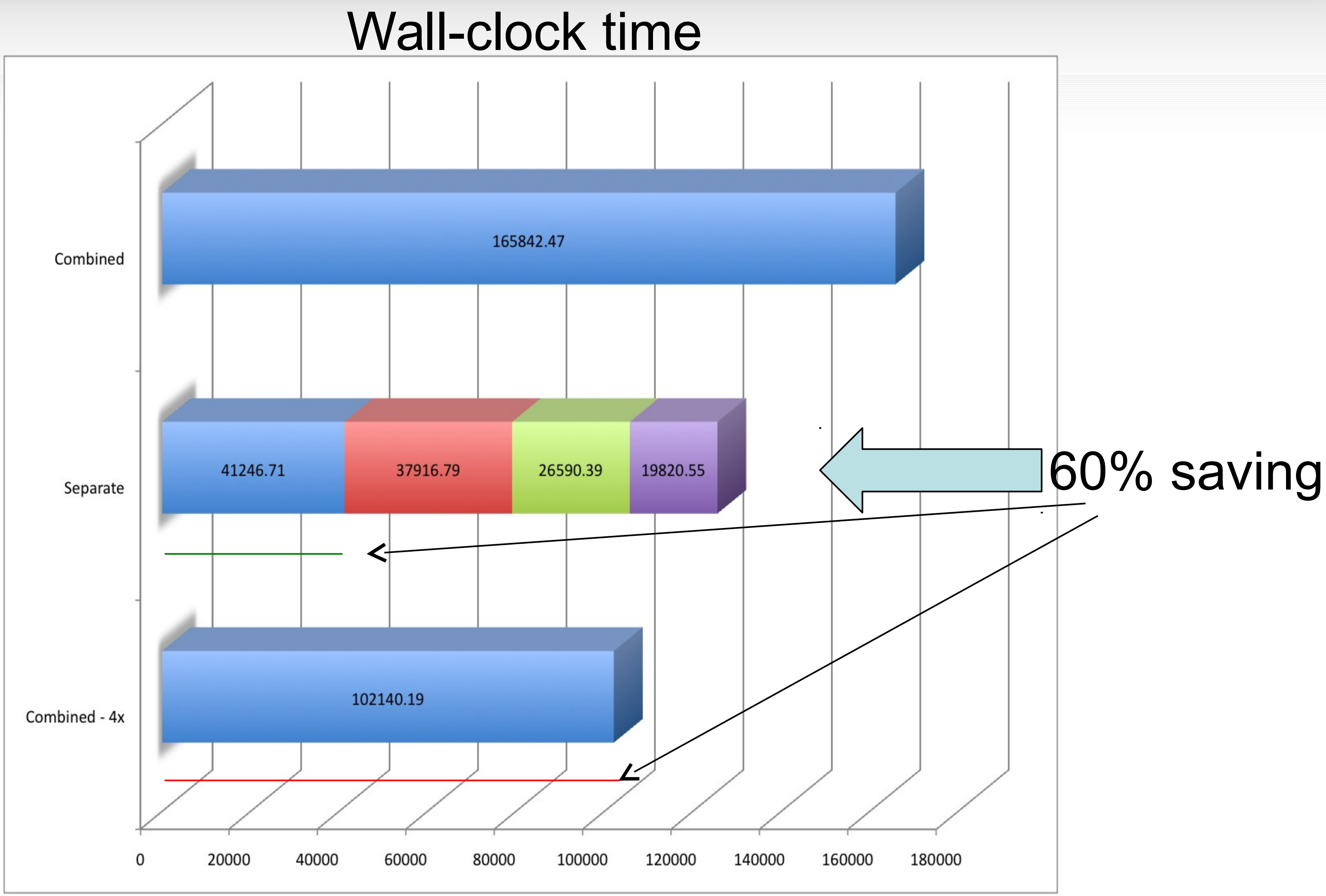
Scalings:separation saving

Wall-clock time

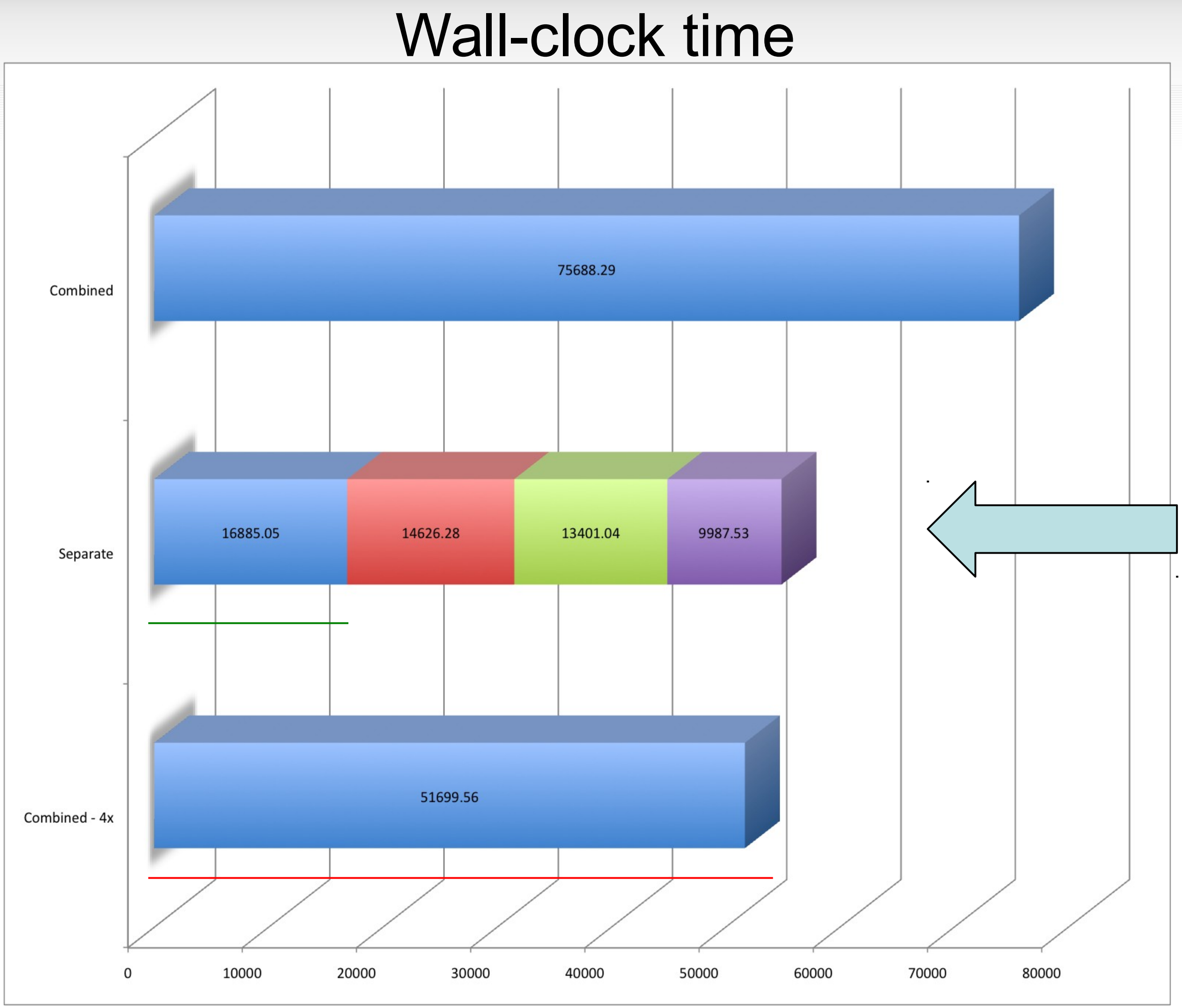


30% saving

Scalings: more saving /w more cores

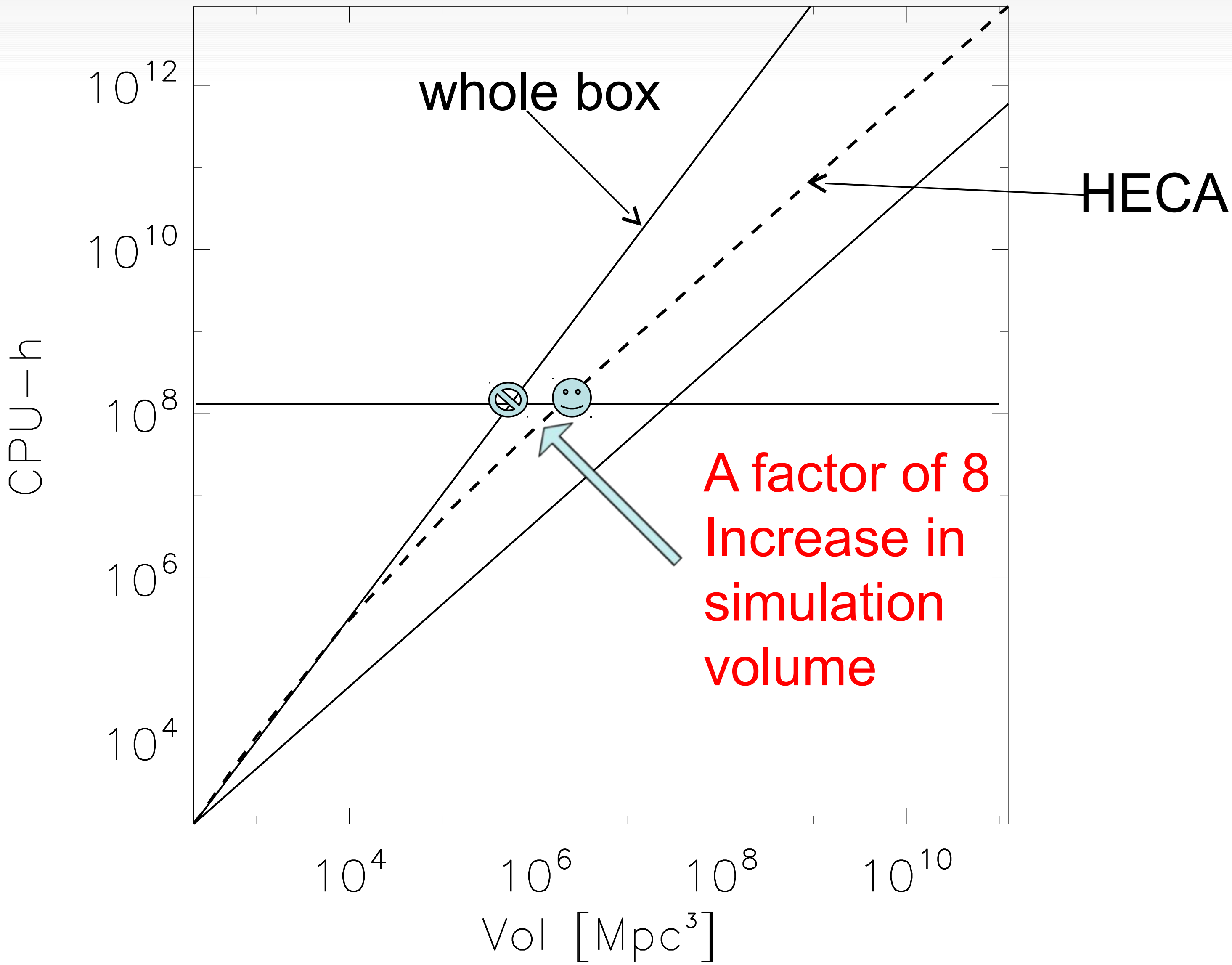


Scalings: more saving with high res



70% saving
good news

Scalings with 0.1 Gigahours



Outlook

- Hybrid (OpenMP + MPI) approach
- Implementation of FTI library
- Some 'physics' are still missing
- Convert ICs for grid based codes (ENZO, AMR, TVD)

Conclusions

- Higher resolution possible in HECA than in full-box simulations
- Scalable up to arbitrarily large processor counts

  ➔ Statistical relevant sample of galaxies at high resolution
- Different physical models can be implemented and tested

Thank You

Acknowledgement:

This research is part of the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (award number OCI 07-25070) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana-Champaign, its National Center for Supercomputing Applications, Cray, and the Great Lakes Consortium for Petascale Computation.