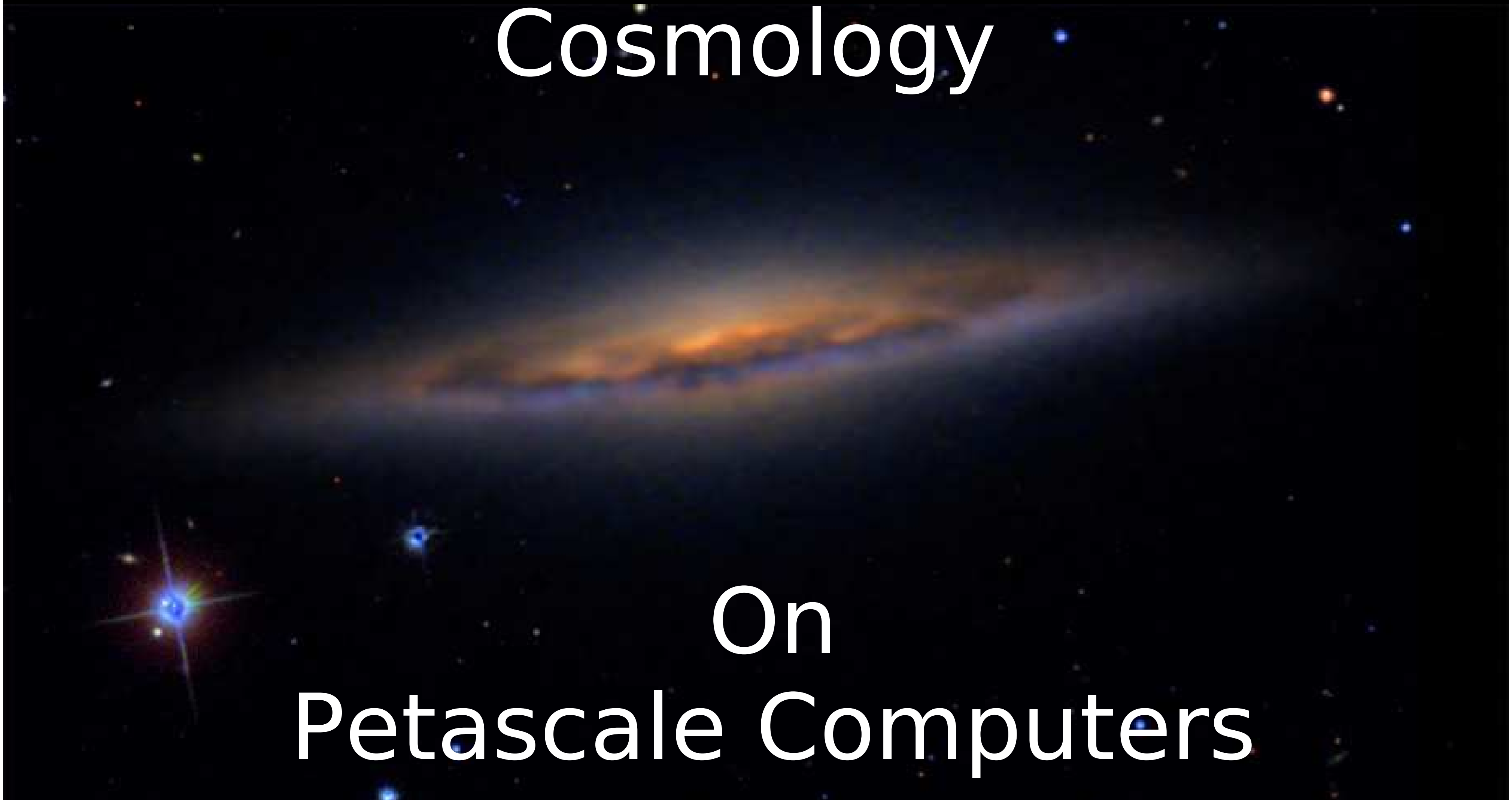
Cosmology

On
Petascale Computers



Fabio Governato, Greg Stinson, Chris Brook, Alison Brooks

Joachim Stadel, Lucio Mayer, Ben Moore, George Lake (U. ZH)

James Wadsley (McMaster)

PARALLEL PROGRAMMING LABORATORY

Department of Computer Science
University of Illinois, Urbana-Champaign

Laxmikant Kale, Celso Mendes,
Pritish Jetley, Amit Sharma
Filippo Gioachin, Orion Lawlor

Importance of Cosmology

- The great thing about cosmology is:
  - Nobody cares
  - Galaxy surveys can be used to test scaling of databases and nobody gets hurt

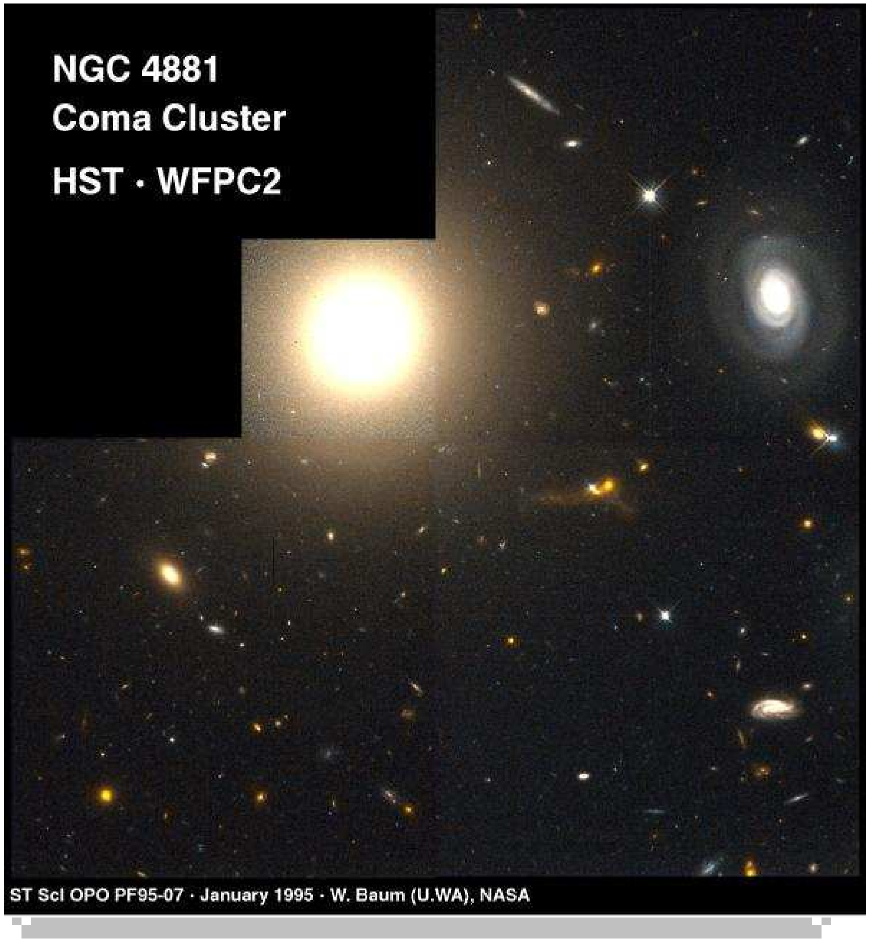
- The great thing about cosmology is:
  - Lots of people care
  - Easy to justify cosmology research to general public

Outline

- Scientific background
- How to build a Galaxy
- The importance of resolution
- Future Challenges
  - Needed Simulations
  - Technology Challenges

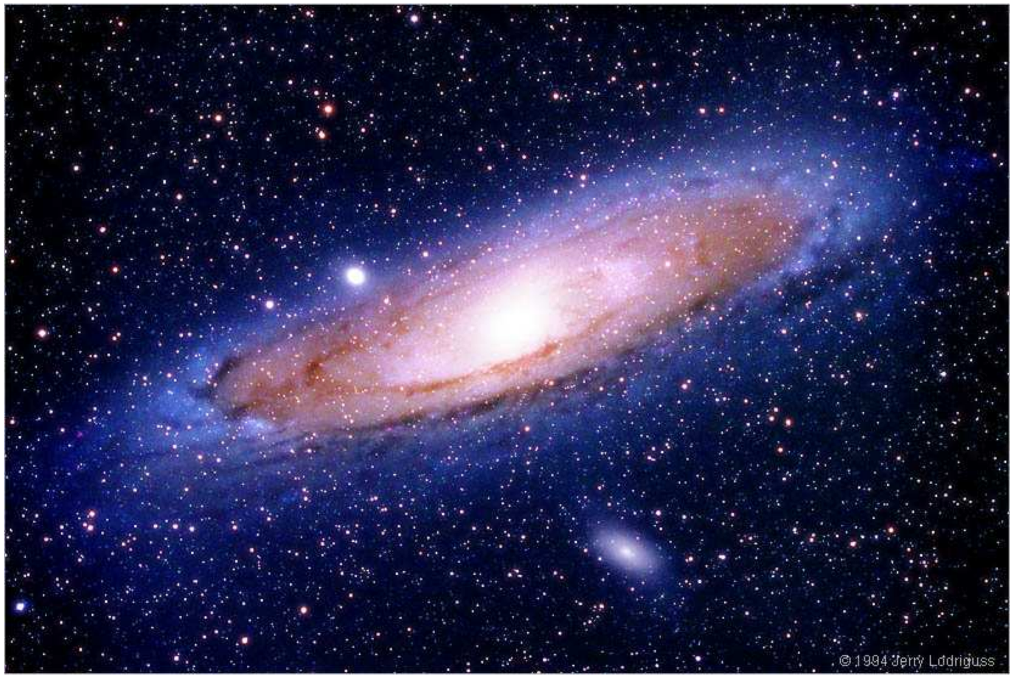
Evidence for Dark Matter

- Motions of Galaxies in Clusters
- $\langle v^2 \rangle \sim GM/R$
- $M_{total} >> M_{stars}$



Evidence for Dark Matter (II)

- Disk Galaxies:
- Rotation velocities => $M_{total} >> M_{star}$
- Simulations show distribution is spherical



The Cosmic Microwave Background

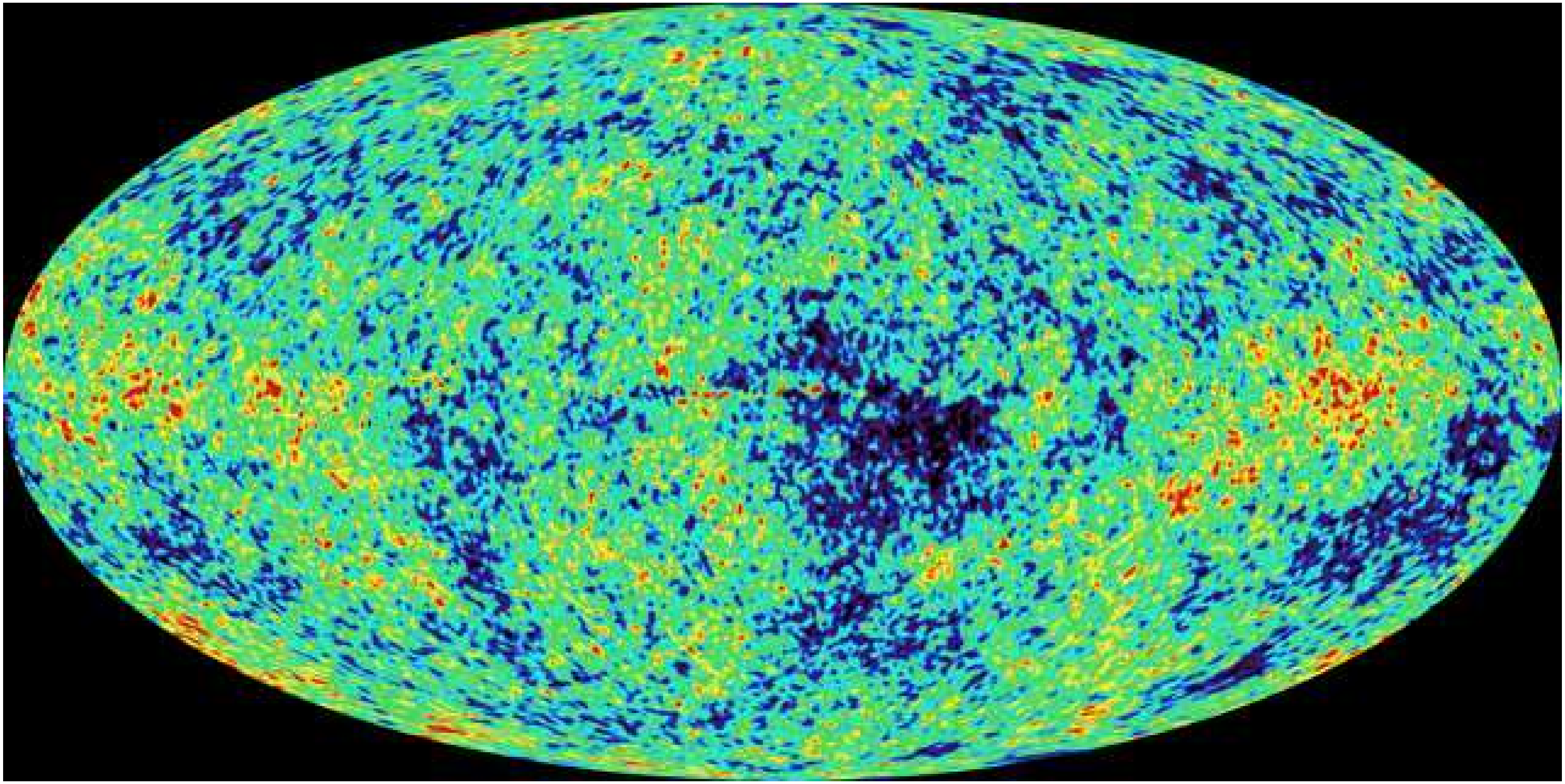
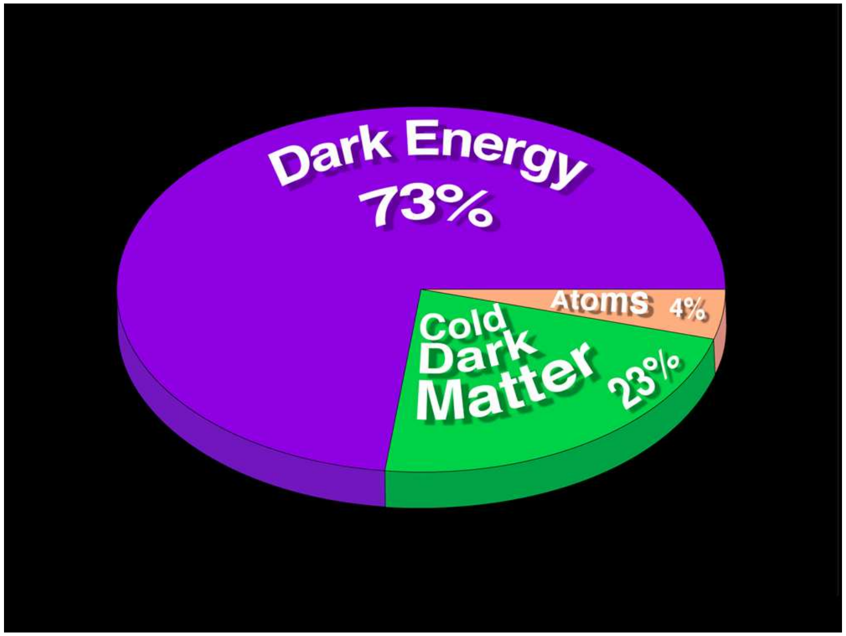
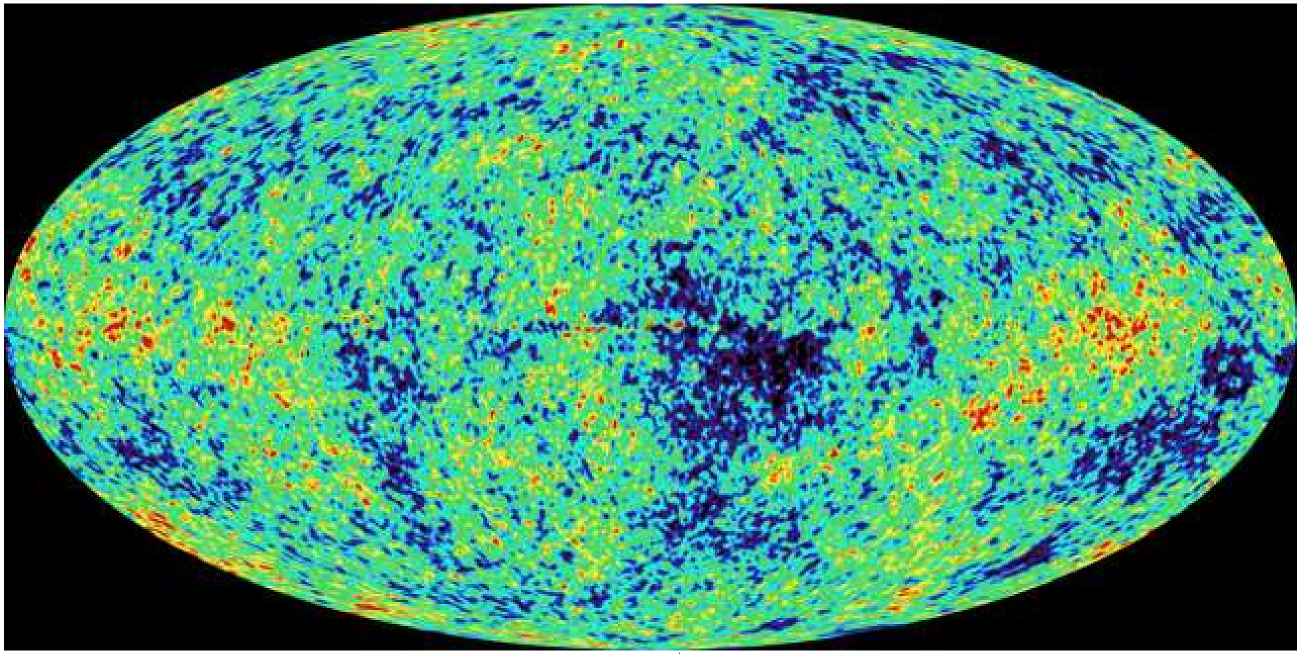


Image courtesy NASA/WMAP

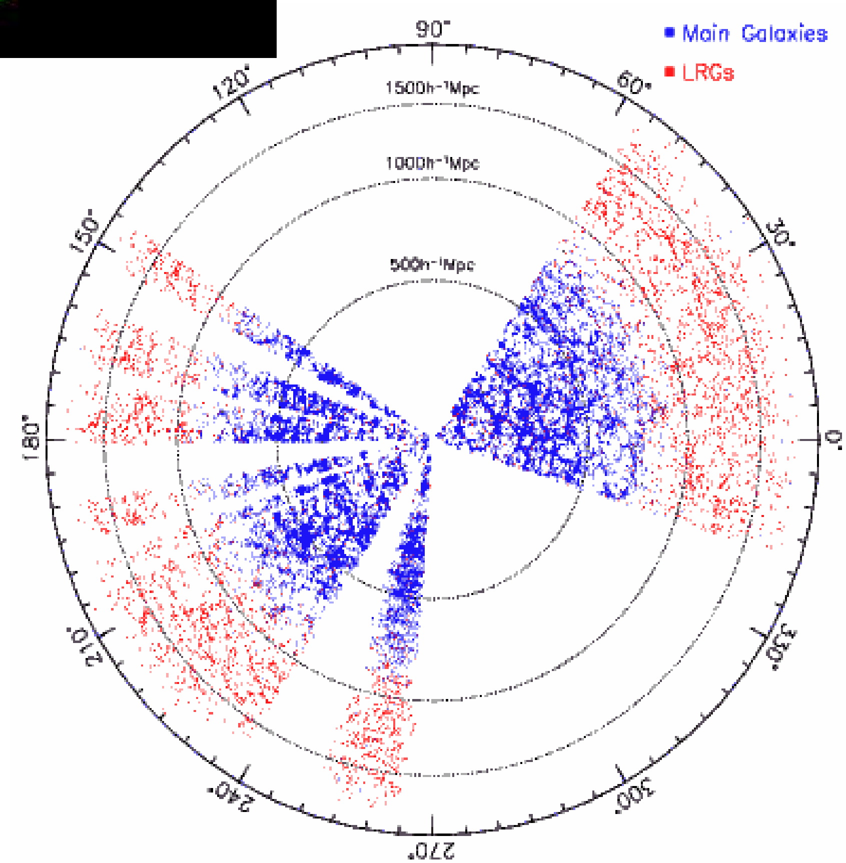
Dark Matter and Energy: What is it?

- Not baryons
- Simulations show: not known neutrinos
- Candidates:
  - Sterile Neutrinos
  - Axions
  - Lightest SUSY Particle (LSP)





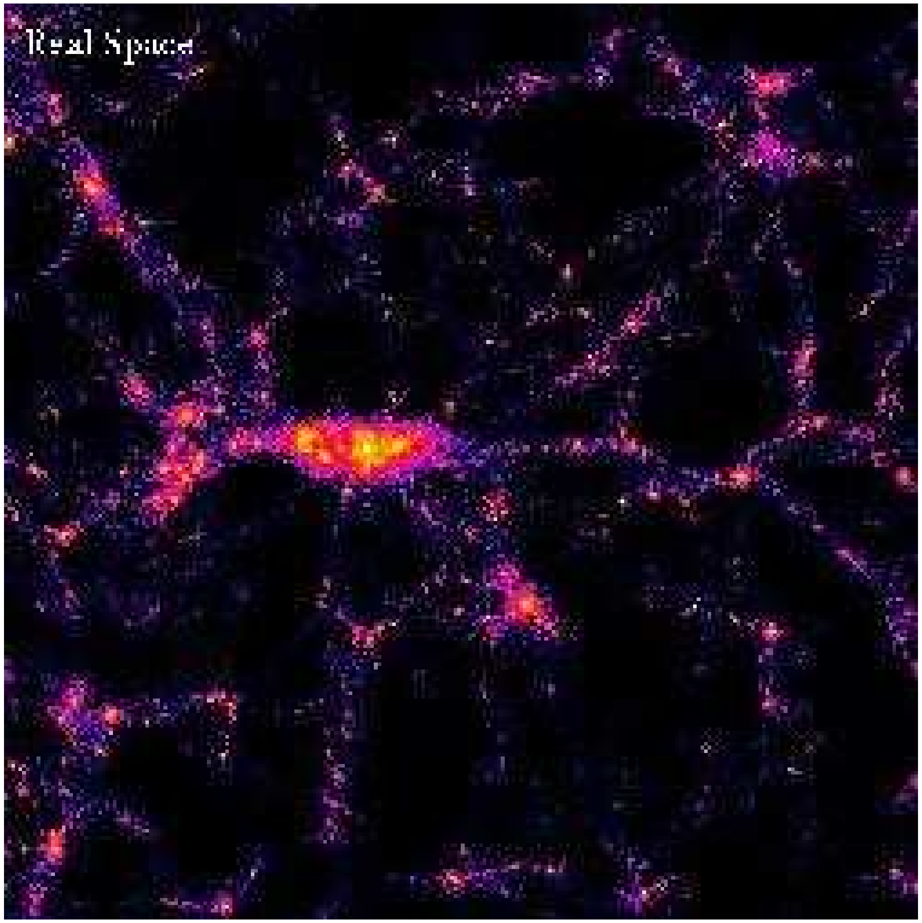
Lots of Physics



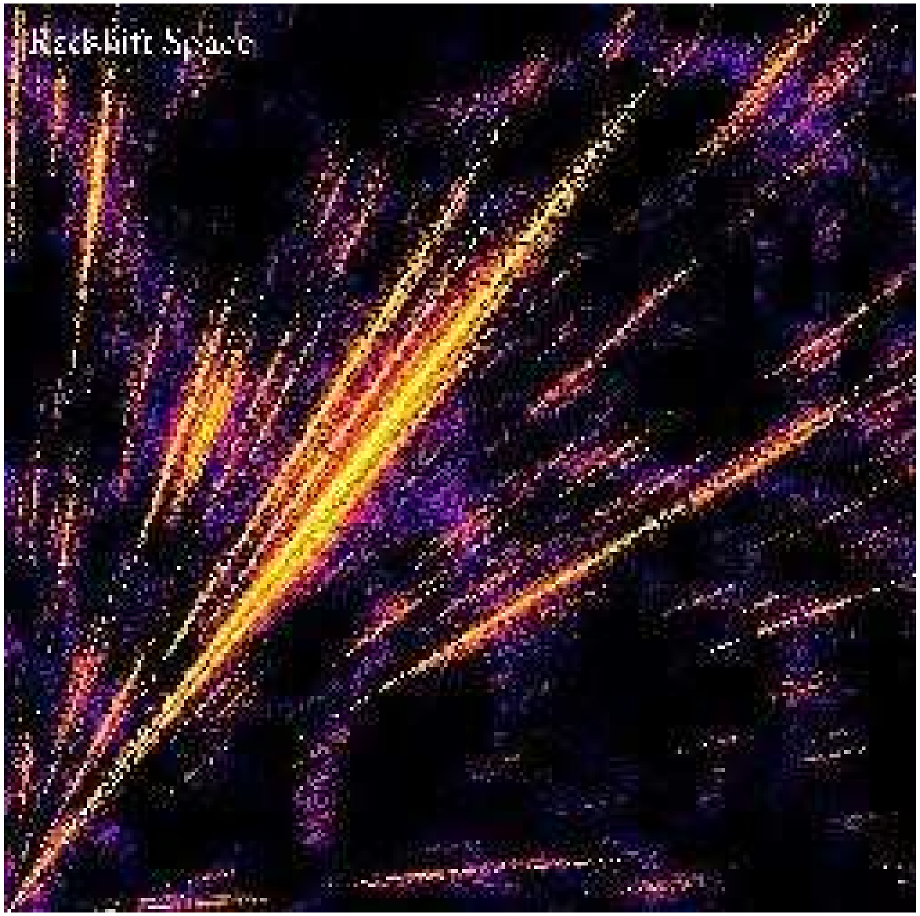
Simulation process

- Start with fluctuations based on Dark Matter properties
- Follow model analytically (good enough to get CMB)
- Create a realization of these fluctuations in particles.
- Follow the motions of these particles as they interact via gravity.
- Compare final distribution of particles with observed properties of galaxies.

Real Space

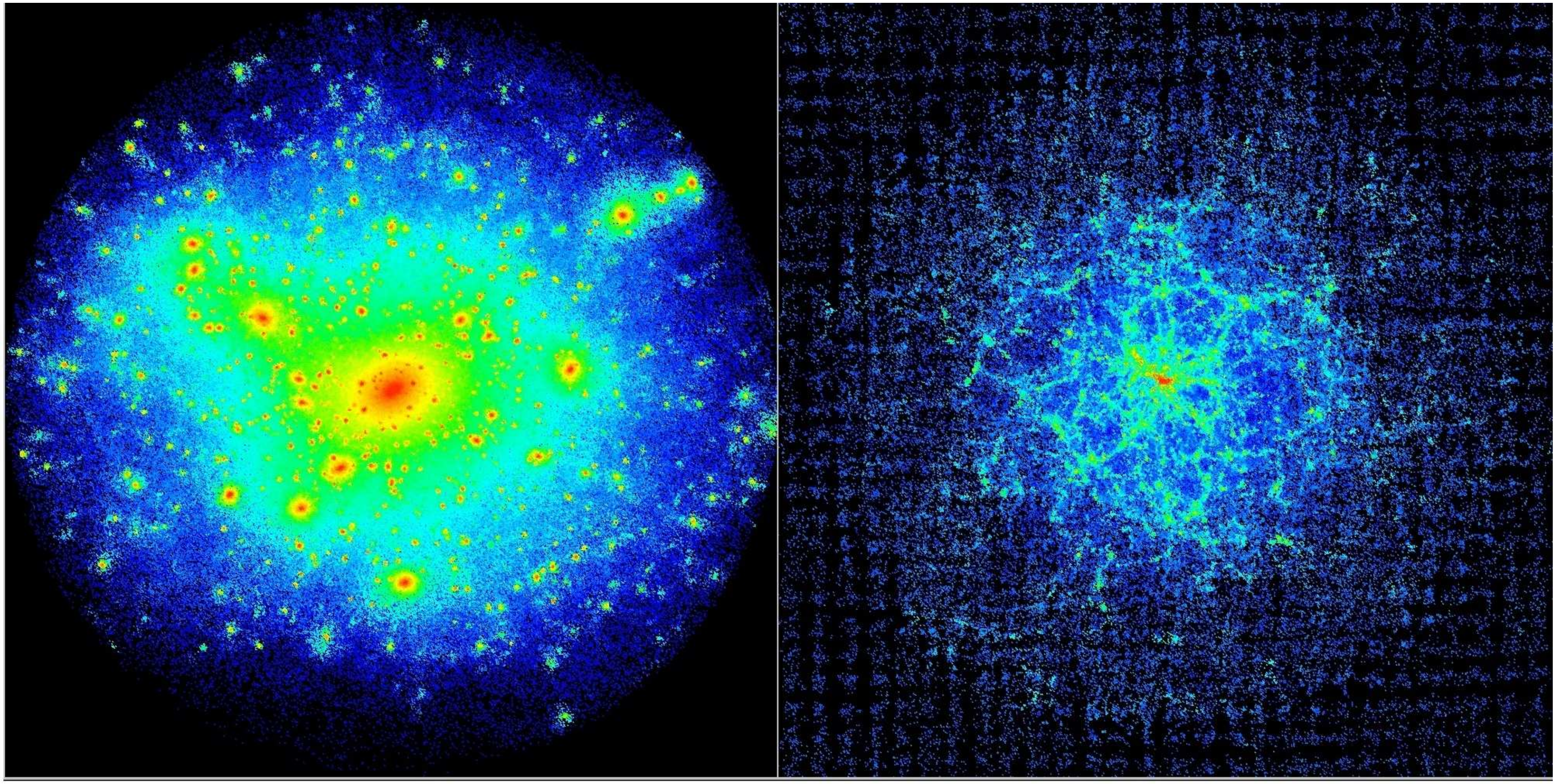


Redshift Space

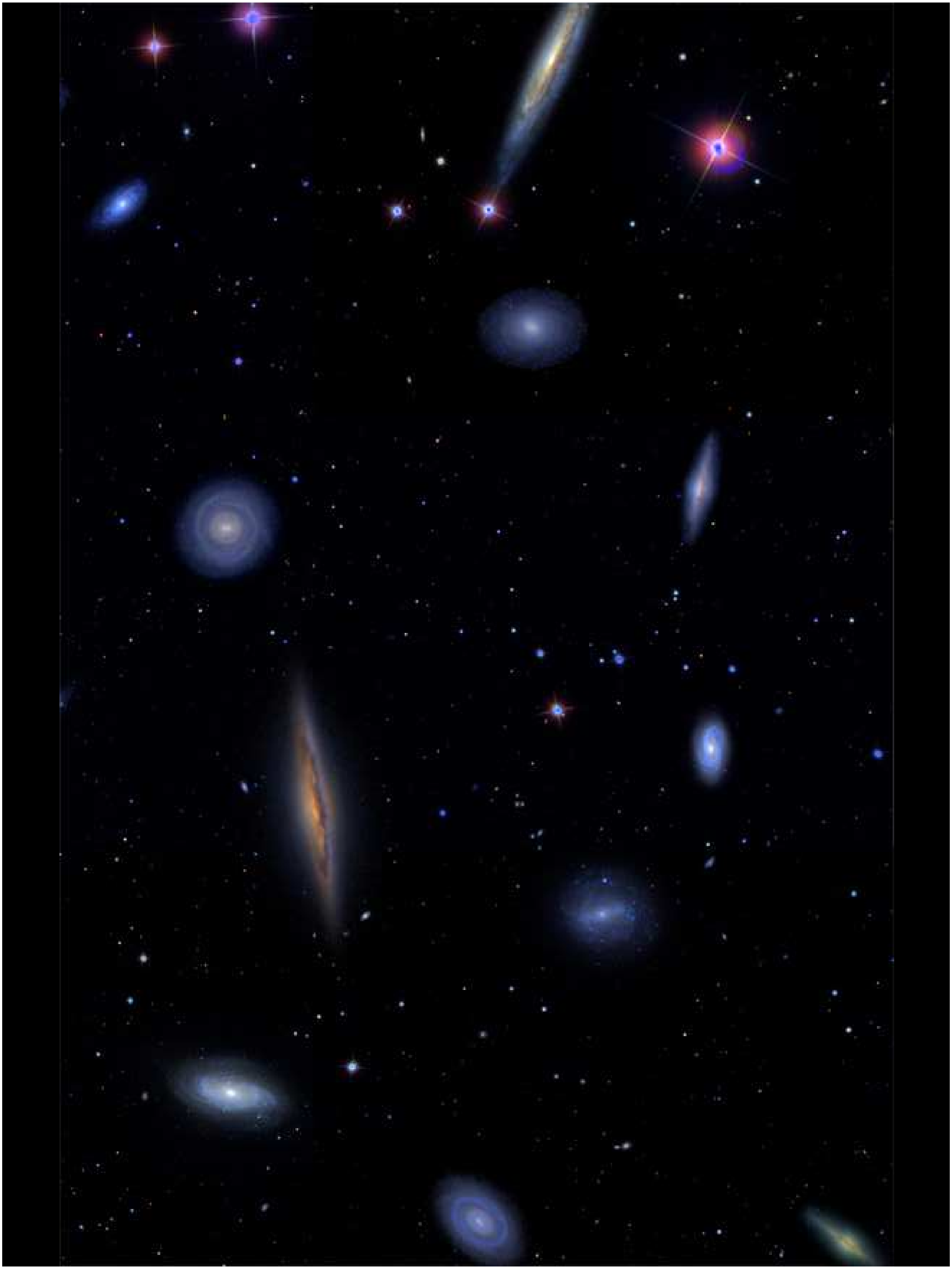


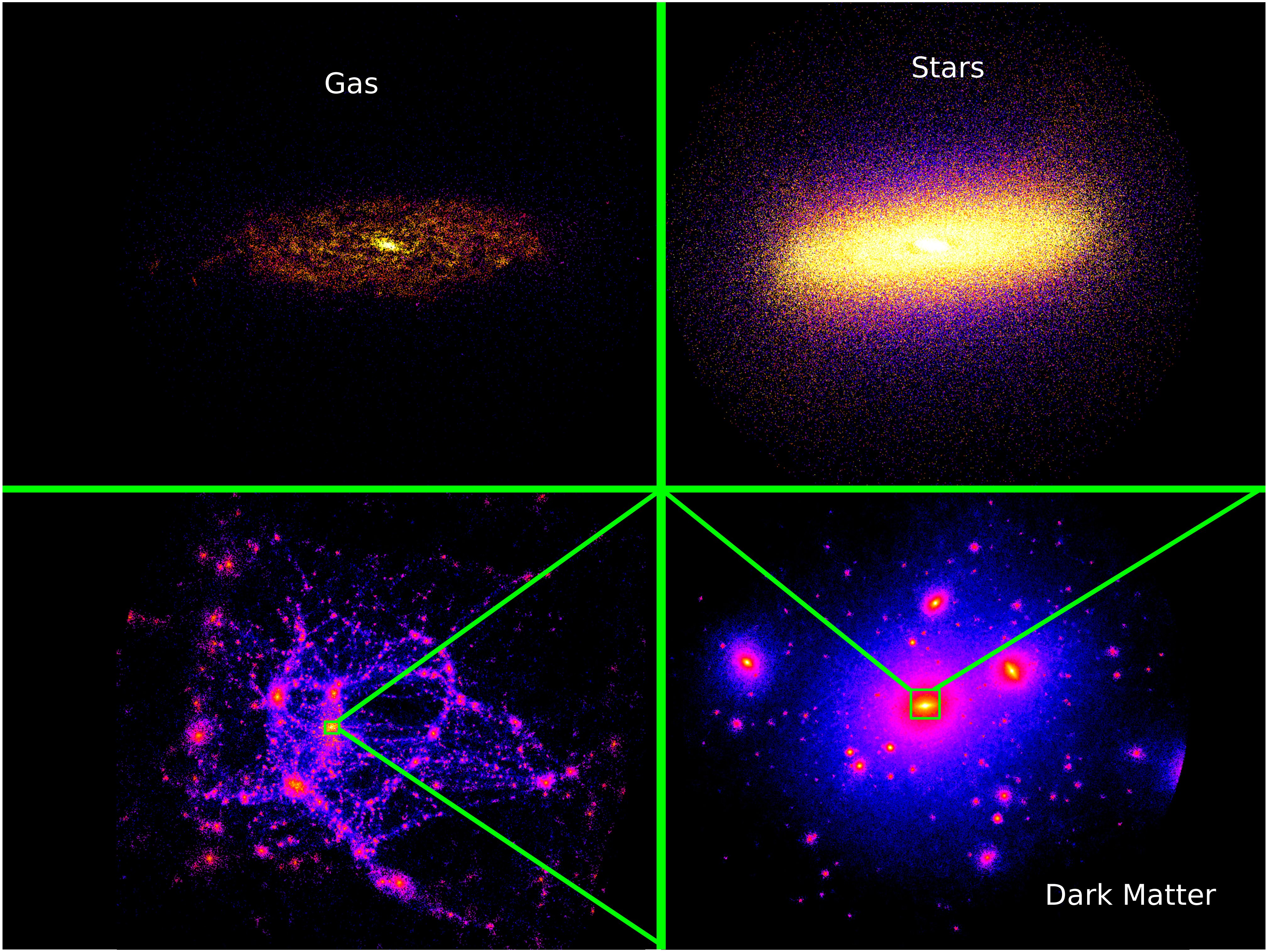
Simulating galaxies: Procedure

1. Simulate 100 Mpc volume at 10-100 kpc resolution
2. Pick candidate galaxies for further study
3. Resimulate galaxies with same large scale structure but with higher resolution, and lower resolution in the rest of the computational volume.
4. At higher resolutions, include gas physics and star formation.



<------------6 Mpc Sphere------------> <------------1000 Mpc Box------------>

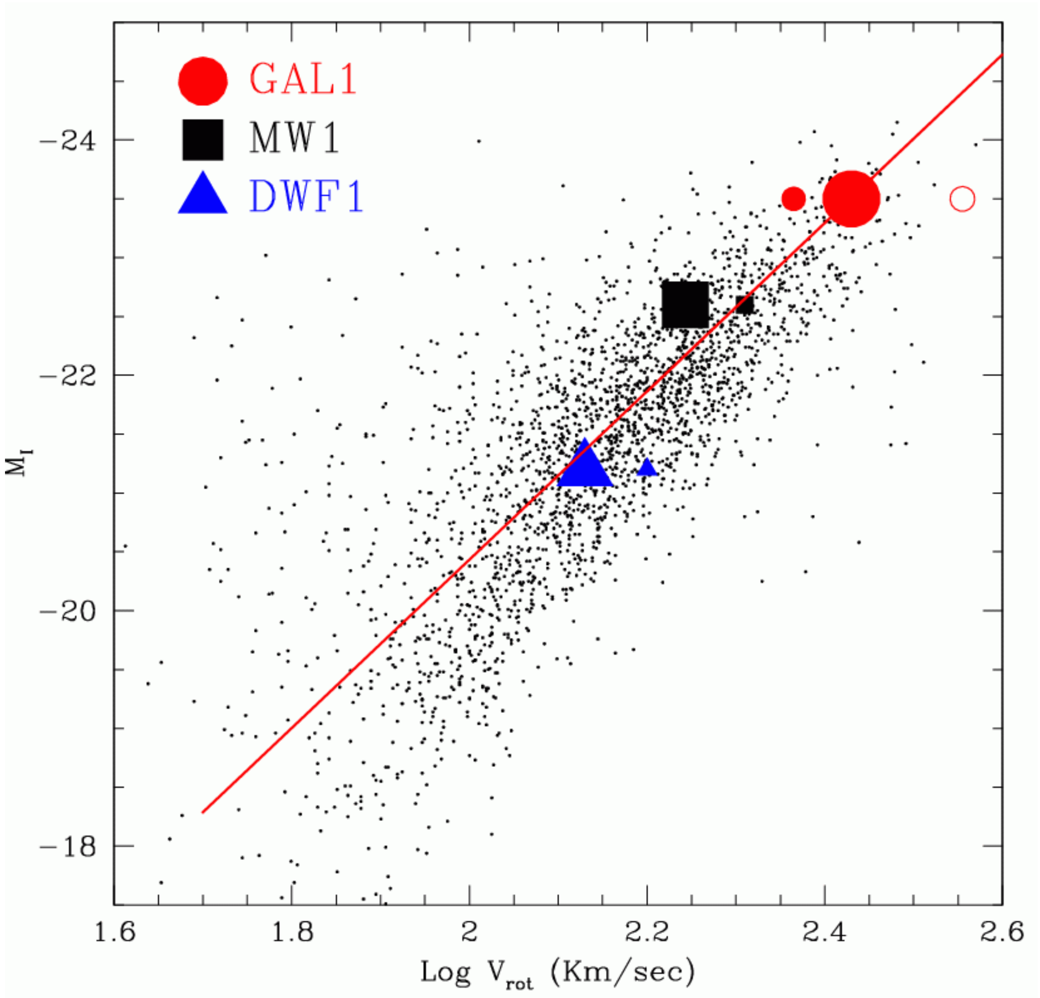



Gas
Stars
Dark Matter

Simulation successes

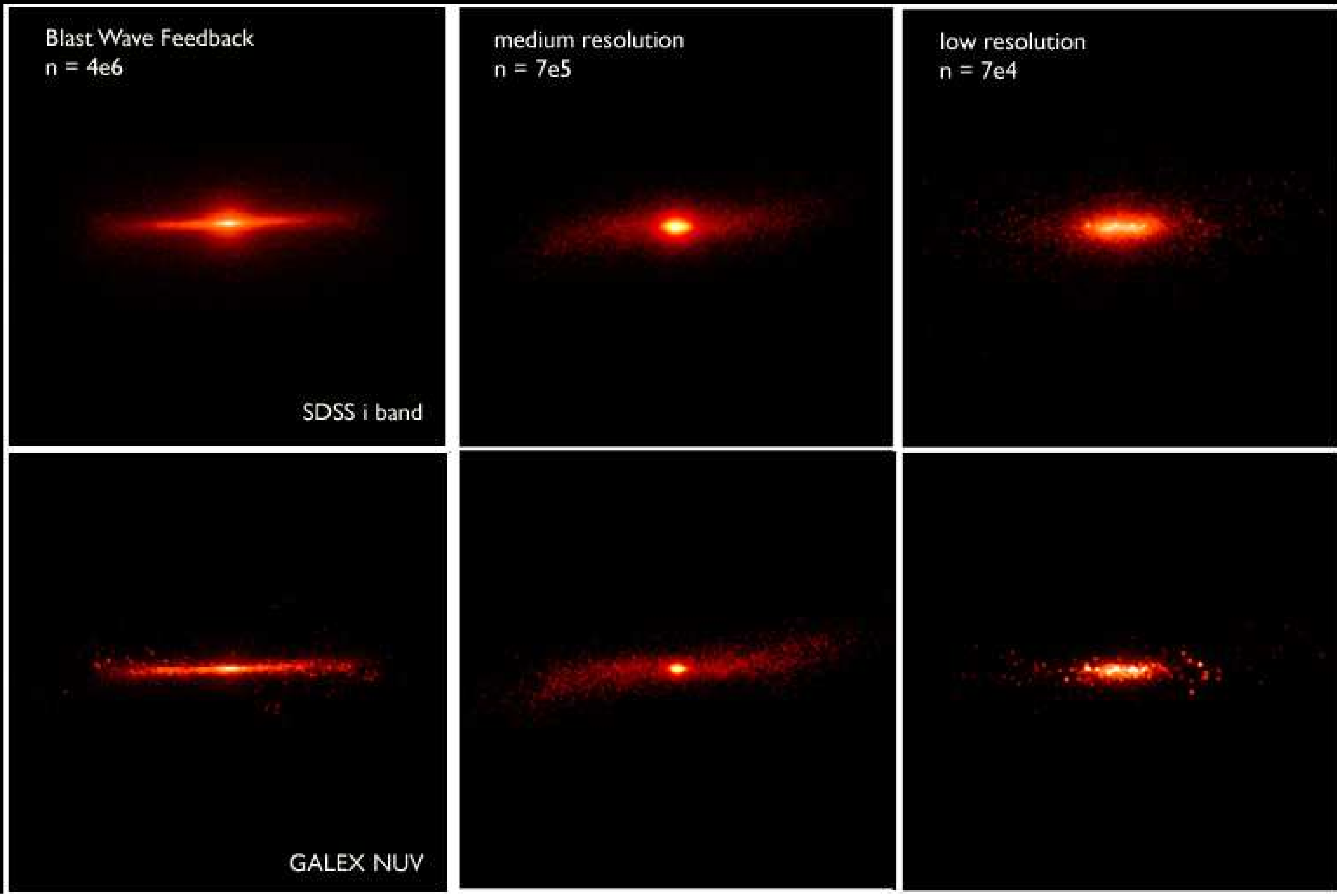
- Number and luminosity of MW satellites
- Star formation history vs. mass (downsizing)
- Galaxy Luminosity vs. Mass (Tully-Fisher)
- Disk scale length
- Chemical enrichment from Supernovea
- Correct amount of cold gas

Tully Fisher (mass-luminosity)



Giovanelli & Haynes 05

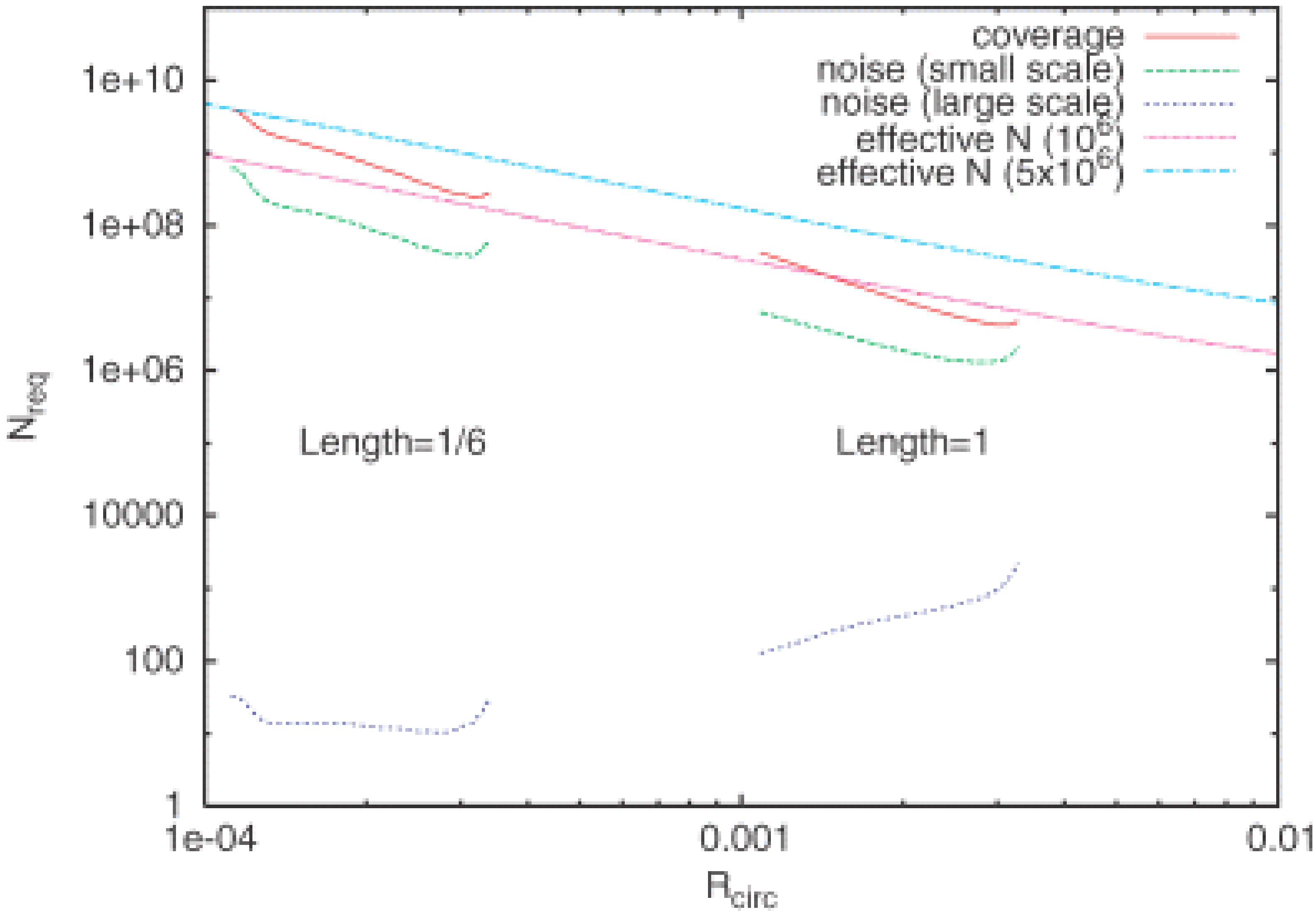
Resolution dependence



Outstanding issues

- Have we converged?
- Galaxies and their environment
- Galaxies long ago and far away.
- Models for the galaxy surveys.

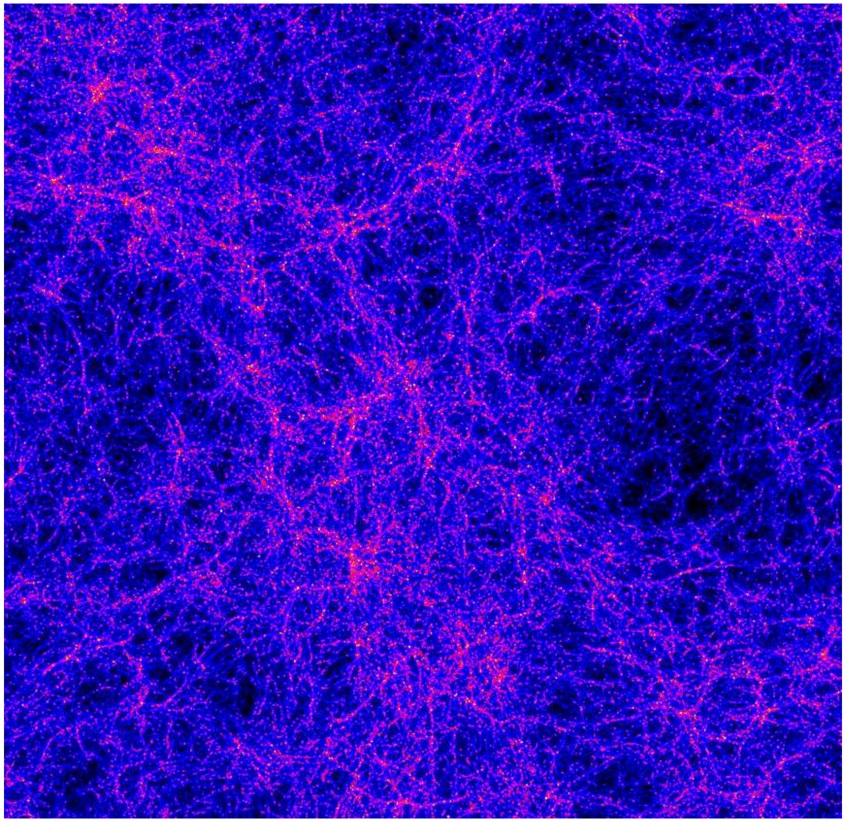
Have We converged?



Weinberg & Katz (2007)

Star Formation History: Current Simulation

- Local Universe with 800M particles
- 6 Teraflop-weeks to complete
- 60000 particles/galaxy

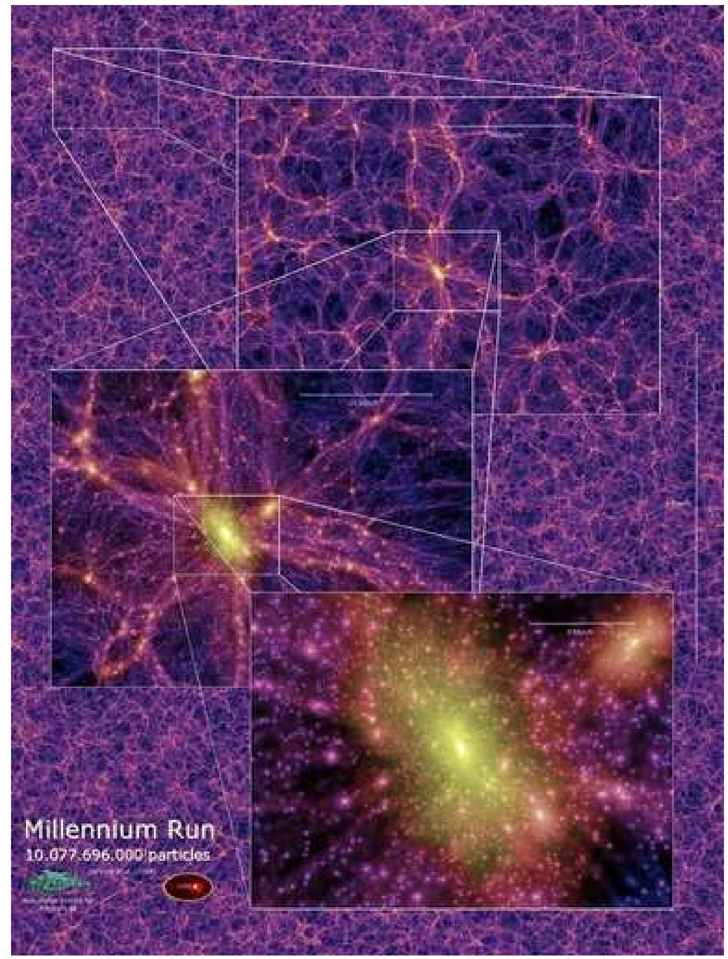


Star Formation History: What's needed

- 1 Million particles/galaxy for proper morphology
- 1 Petaflop-week of computation
- Necessary for:
  - Comparing with Hubble Space Telescope surveys of the local Universe
  - Interpreting HST images of high redshift galaxies

Large Scale Structure: Current Simulation

- “Fair sample” (700 Mpc) of Universe with 10 billion particles
- 1000 particles/galaxy
- 1 Teraflop-week to complete
- Being rerun with gas



Large Scale Structure: What’s needed

- 6.5 Gigaparsec volume
- 10 Trillion particles (1 Petabyte of RAM)
- 1 Petaflop week of computation
- Necessary for:
  - Interpreting future surveys (LSST)
  - Relating Cosmic Microwave Background to galaxy surveys

Ultimate Universe Simulation

- “Fair sample” (700 Mpc volume)
- 1 million particles/galaxy
- 10 trillion particles
- 1 Exaflop week of computation
- Poses significant algorithmic challenges

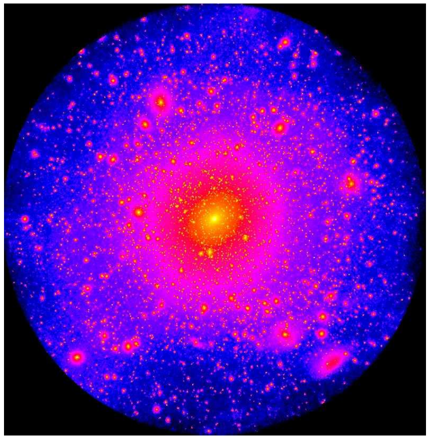
The existing code:

Gasoline

- Multi-Platform
- Massively Parallel (100s; 1000s on large sims)
- Treecode with periodic boundary conditions
- Multi-stepping (but bad load balancing)
- Hydrodynamics (via SPH) with radiative cooling
- UV background
- Star Formation
- Supernovae feedback into thermal energy

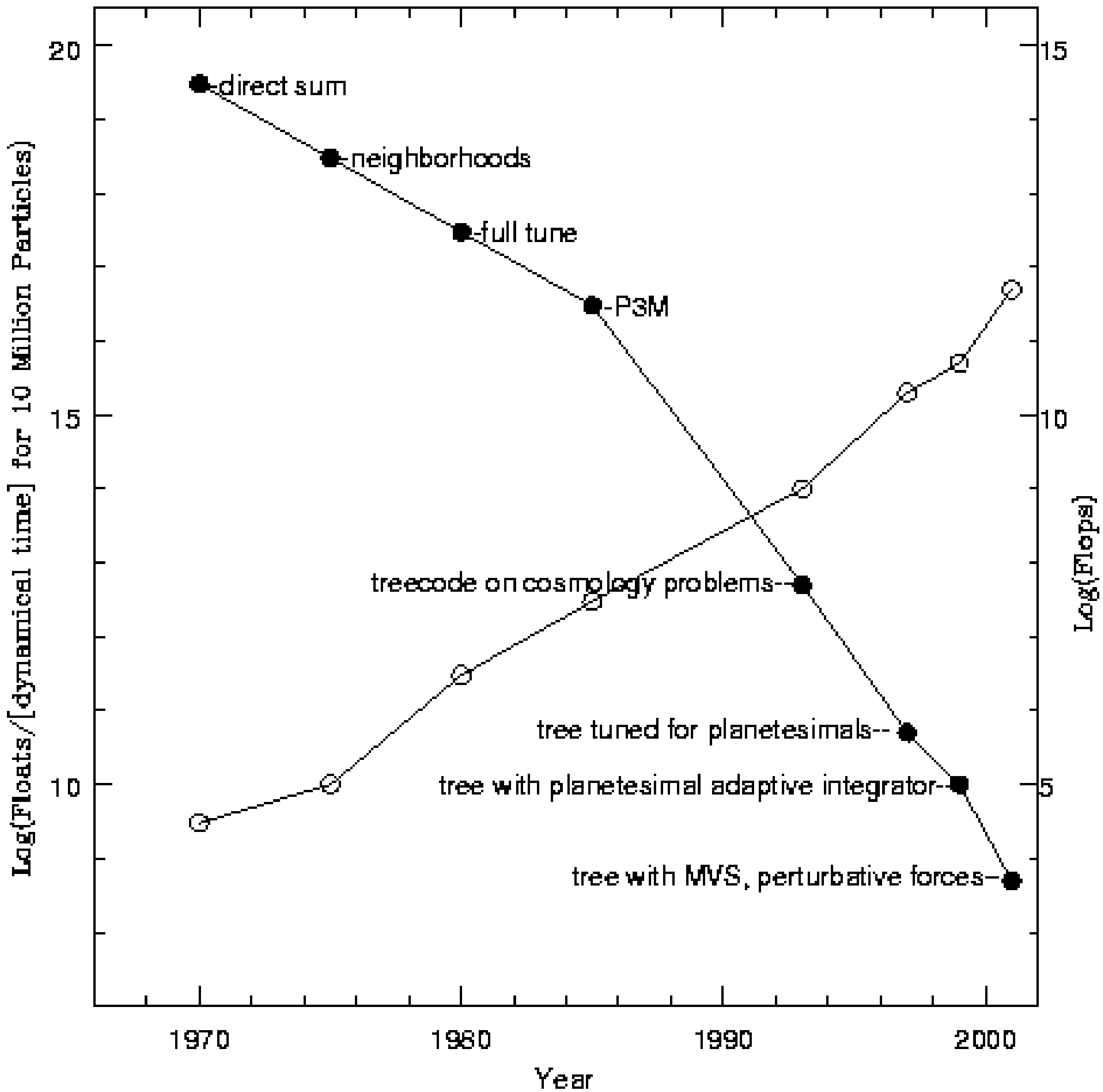
ChaNGa:
CHArm N-body GrAvity

- Chares are “Tree Pieces”
- “TreeCache” for amortizing interprocessor communication
- Periodic boundaries
- Multiple timesteps
- Dynamic load balancing with choice of strategies



Timestepping Challenges

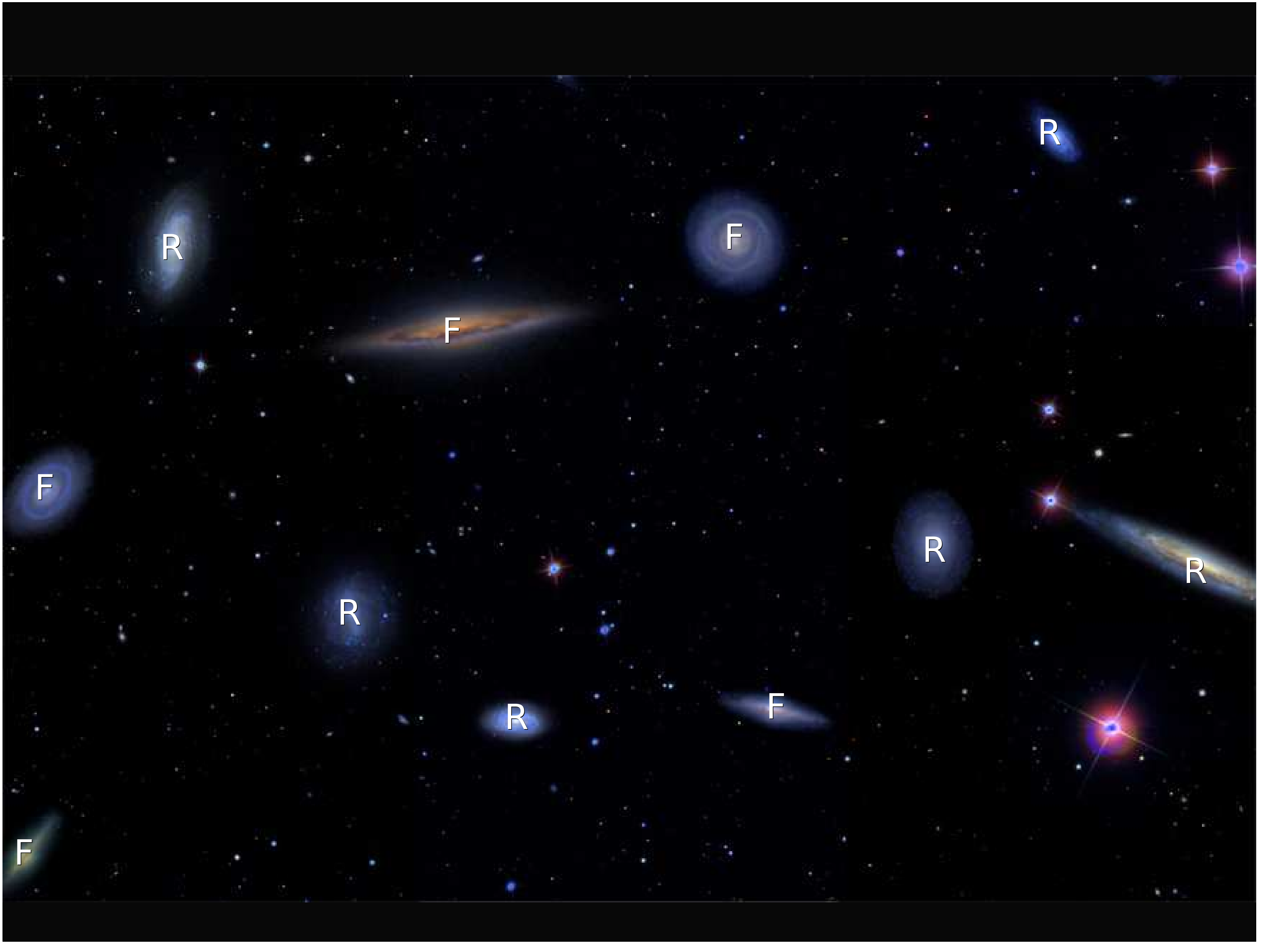
- 1/*m* particles need *m* times more force evaluations
- Naively, simulation cost scales as N^(4/3)ln(N)
  - This is a problem when N ~ 1e9 or greater
- If each particle an individual timestep scaling reduces to N (ln(N))^2
- A difficult dynamic load balancing problem



-direct sum
-neighborhoods
-full tune
-P3M
treecode on cosmology problems--
tree tuned for planetesimals--
tree with planetesimal adaptive integrator--
tree with MVS, perturbative forces--
Log(Floats/[dynamical time] for 10 Million Particles)
Log(Flops)
Year
20
15
10
5
1970
1980
1990
2000

Computing Challenge Summary

- The Universe is big => we will always be pushing for more resources
- New algorithm efforts will be made to make efficient use of the resources we have
  - Efforts made to abstract away from machine details
  - Parallelization efforts need to depend on more automated processes.


R
R
F
F
F
F
R
R
R
R
F
F