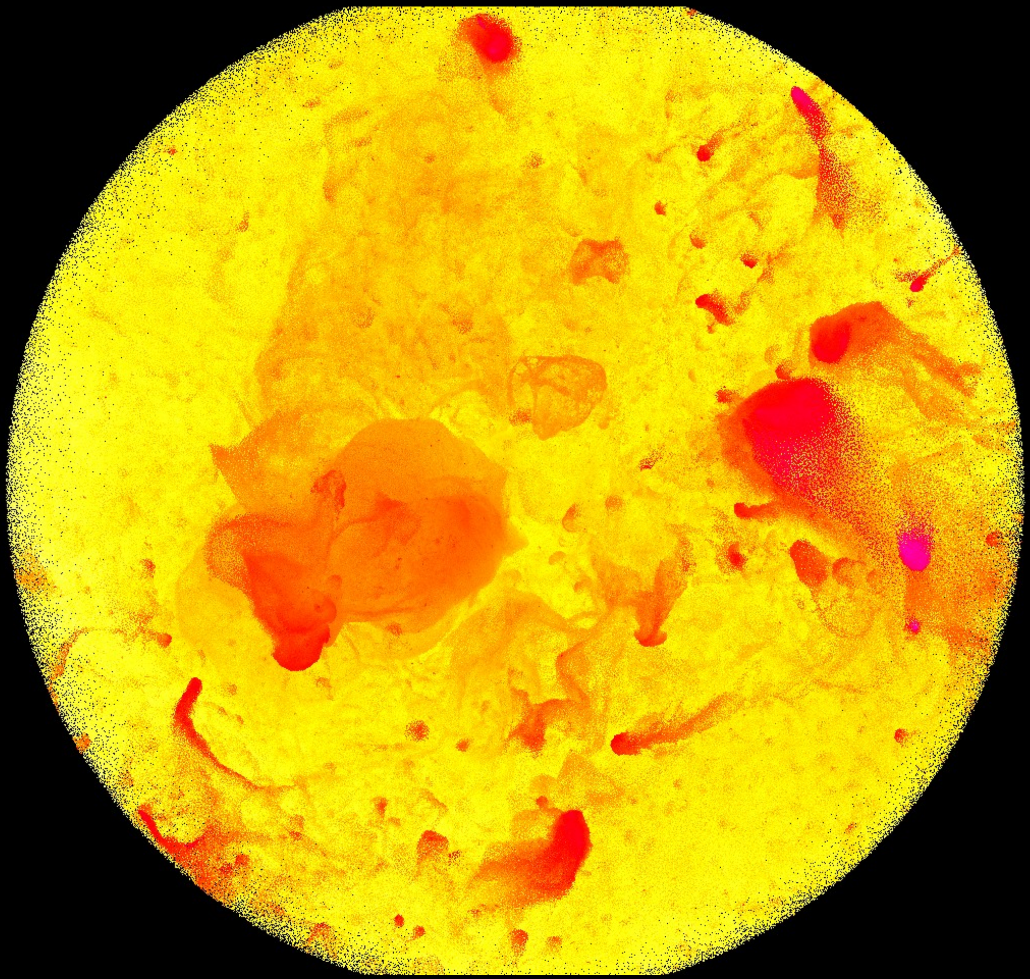
ChaNGa



CHArm N-body GrAvity

PPL PARALLEL PROGRAMMING LAB
DEPT OF COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS

Laxmikant Kale
Filippo Gioachin
Pritish Jetley
Celso Mendes
Amit Sharma
Lukasz Wesolowski
Edgar Solomonik
Orion Lawlor

Outline

- Scientific background
  - Cosmology and fundamental questions
  - Galaxy catalogs and simulations
  - Simulation Challenges
- Charm++ and those challenges
  - Previous state of the art: Gasoline
  - AMPI and Gasoline
  - Charm++ and the send paradigm
  - CkCache, etc.
- Future Challenges

Cosmology at 130,000 years

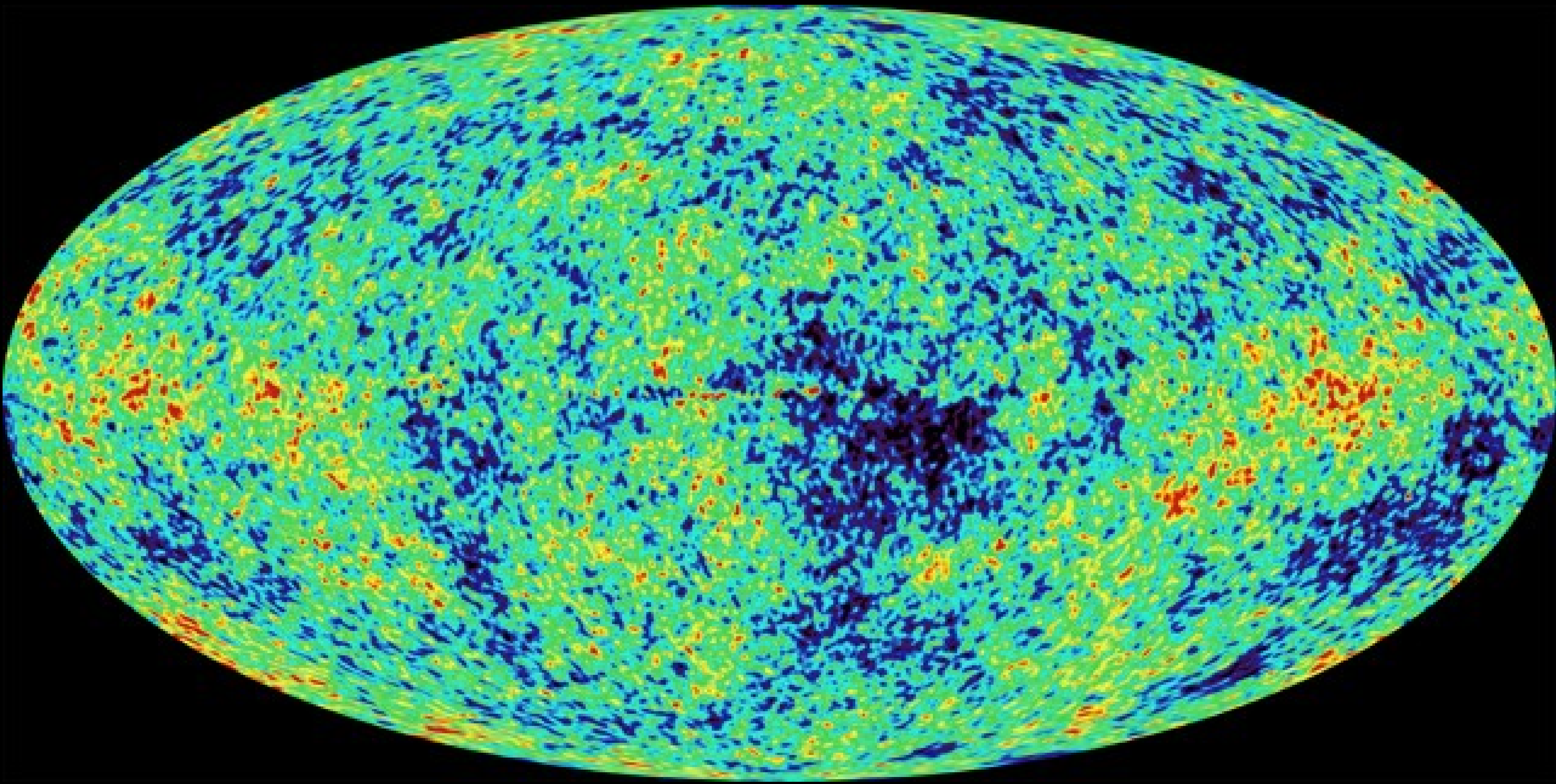
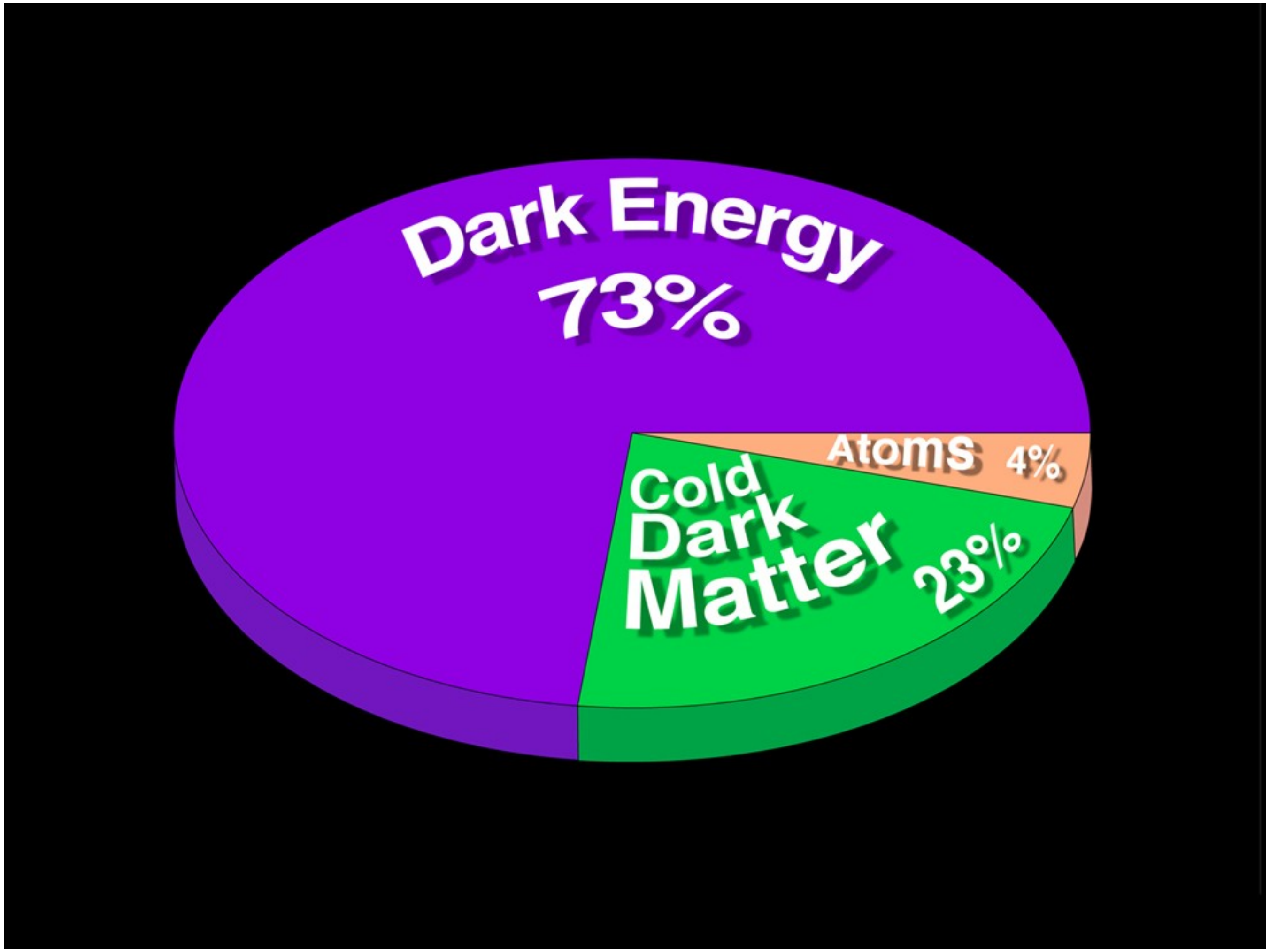
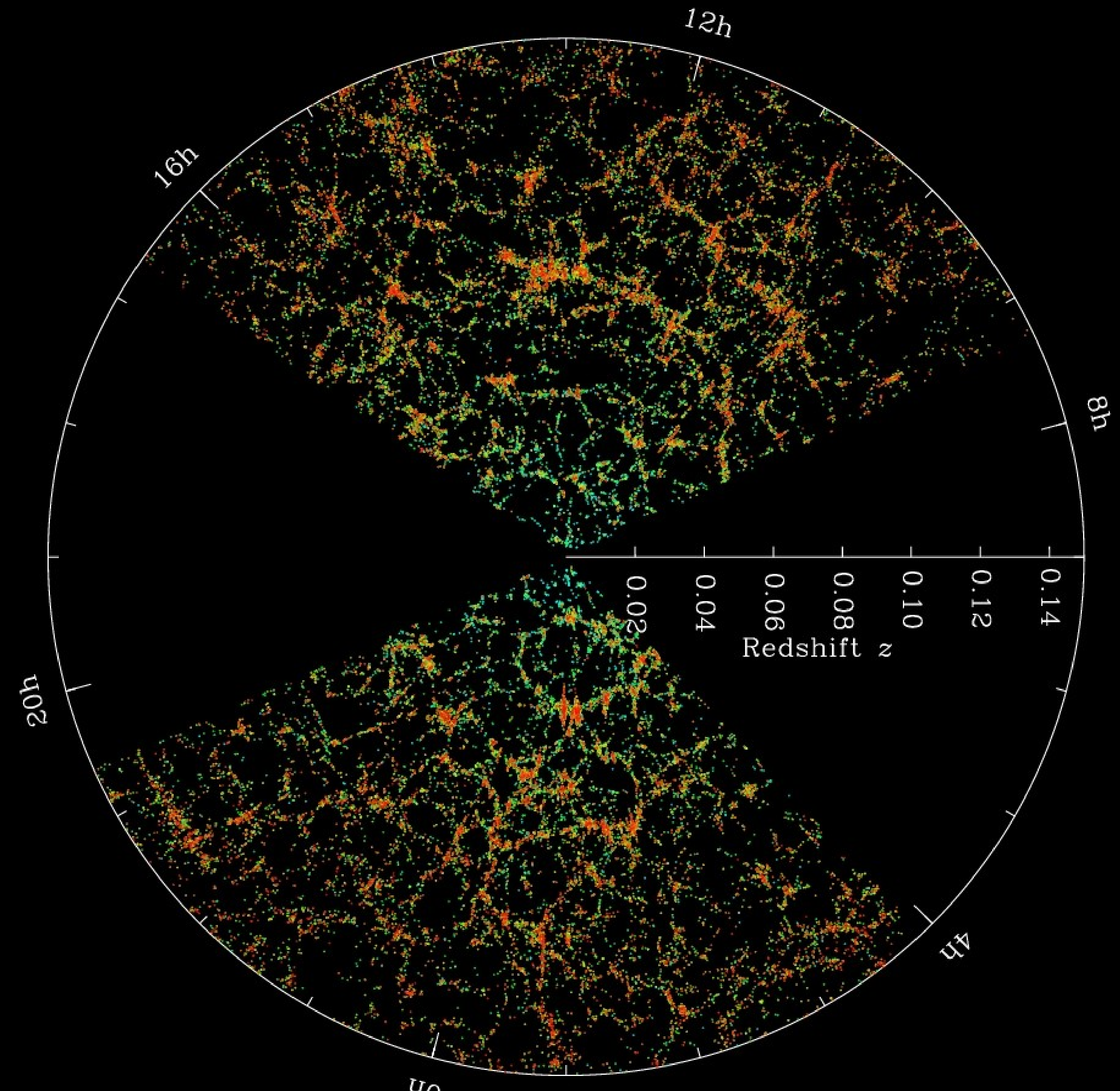


Image courtesy NASA/WMAP


Dark Energy
73%
Atoms 4%
Cold
Dark
Matter
23%


12h
16h
8h
20h
4h
0.02
0.04
0.06
0.08
0.10
0.12
0.14
Redshift z

… is not so simple

Computational Cosmology

CMB has fluctuations of 1e-5

Galaxies are overdense by 1e7

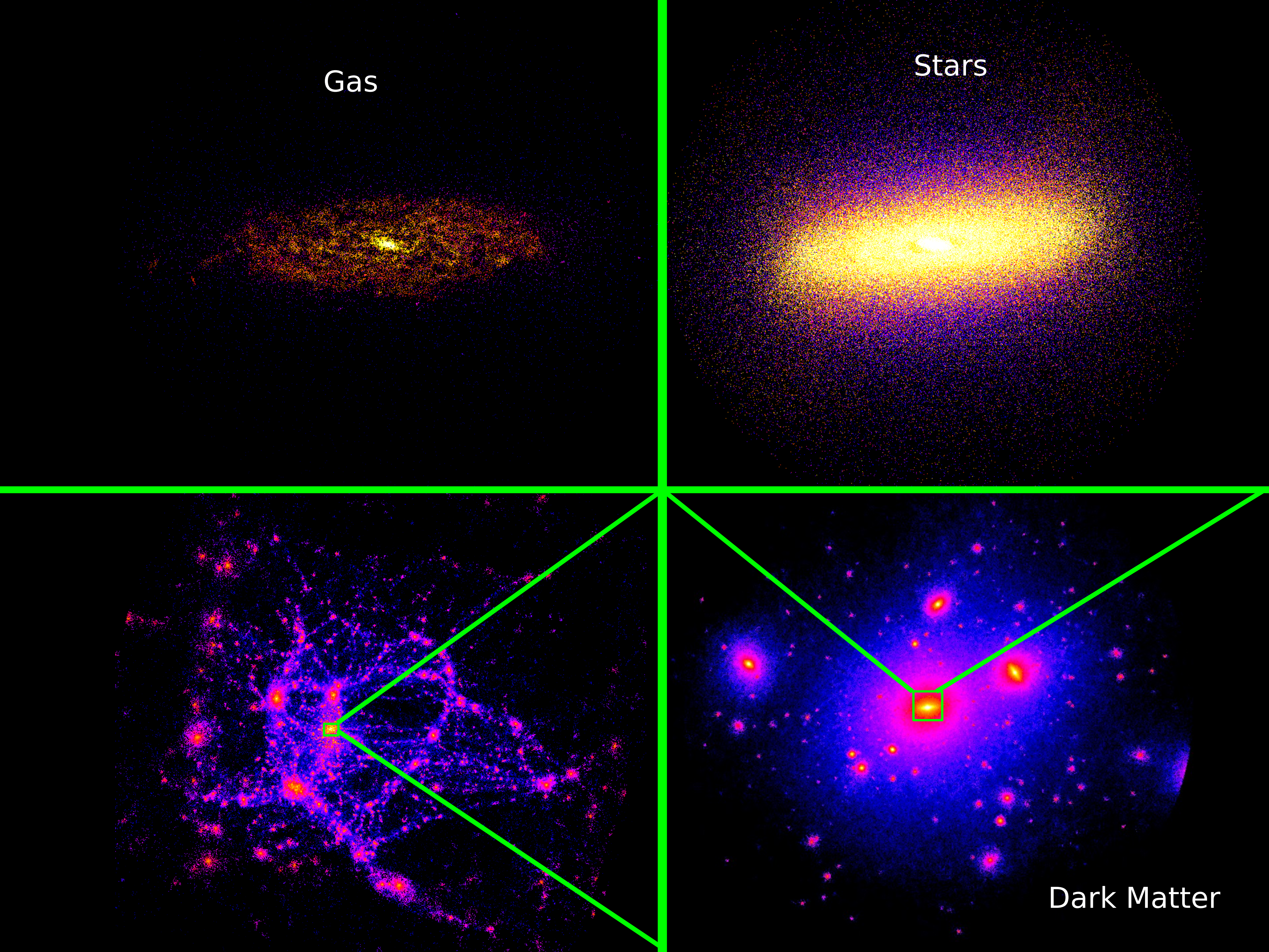
It happens (mostly) through **Gravitational Collapse**

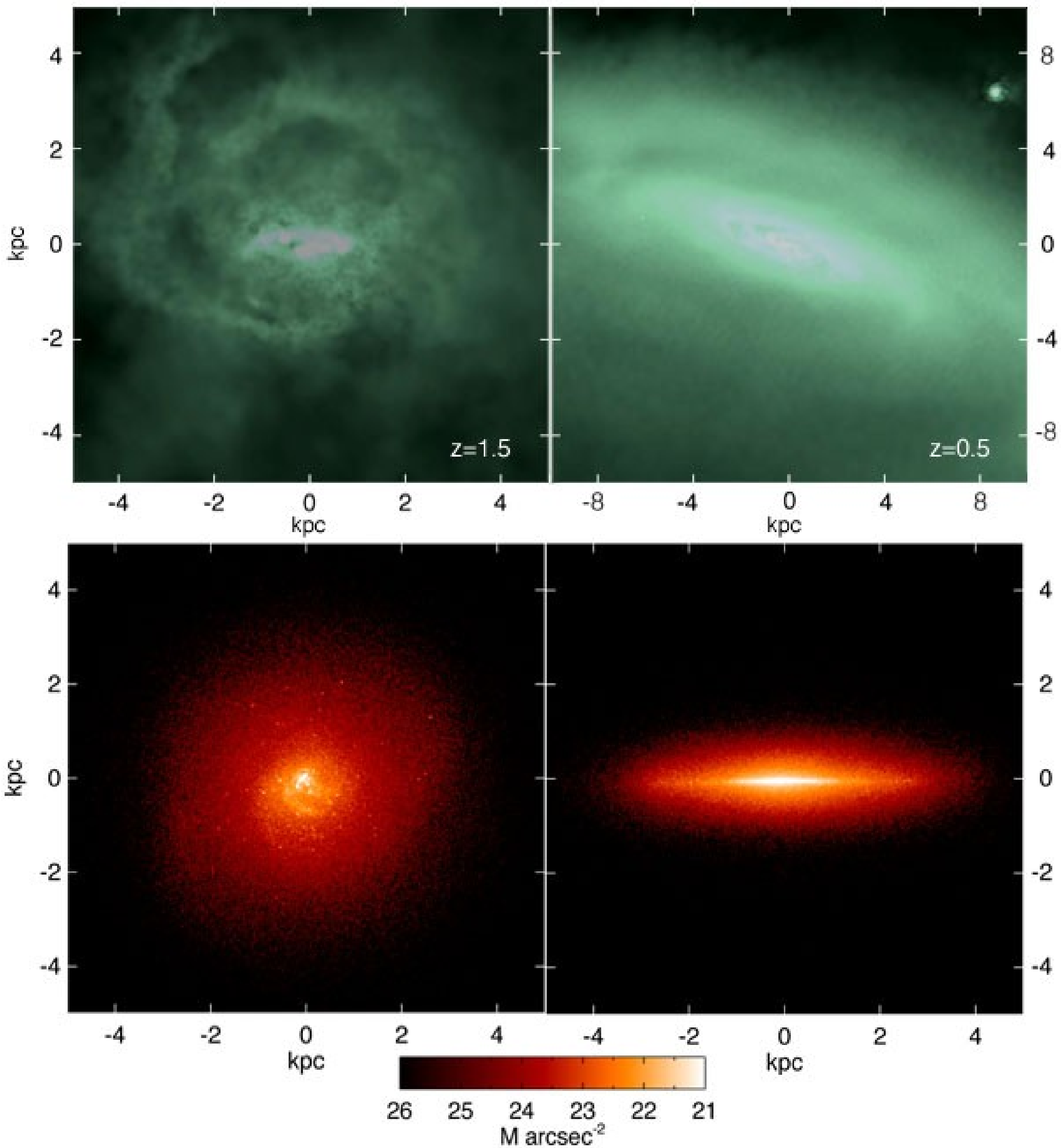
Making testable predictions from a cosmological hypothesis requires

- Non-linear, dynamic calculation
- e.g. **Computer simulation**

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.


Gas
Stars
Dark Matter



z=1.5
z=0.5
kpc
M arcsec^{-2}
26 25 24 23 22 21

Galactic structure in the local Universe: What’s needed

- 1 Million particles/galaxy for proper morphology/heavy element production
- 800 M core-hours
- Necessary for:
  - Comparing with Hubble Space Telescope surveys of the local Universe
  - Interpreting HST images of high redshift galaxies

Large Scale Structure: What's needed

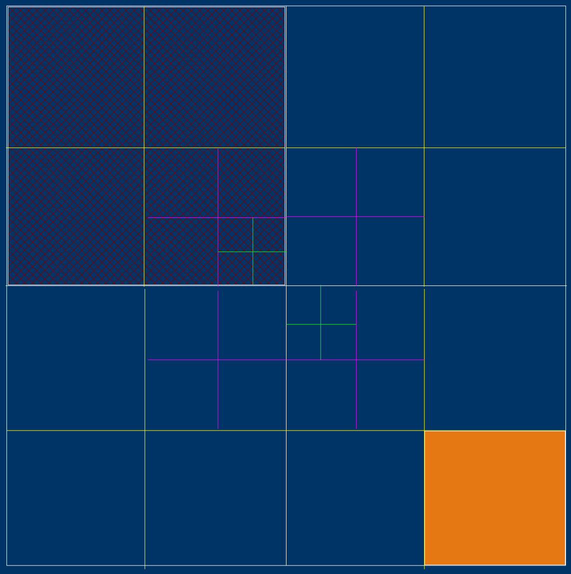
- 700 Megaparsec volume for "fair sample" of the Universe
- 18 trillion core-hours (~ exaflop year)
- Necessary for:
  - Interpreting future surveys (LSST)
  - Relating Cosmic Microwave Background to galaxy surveys
- Cmp. Exaflop example from P. Jetley:
    - 200 Mpc volume

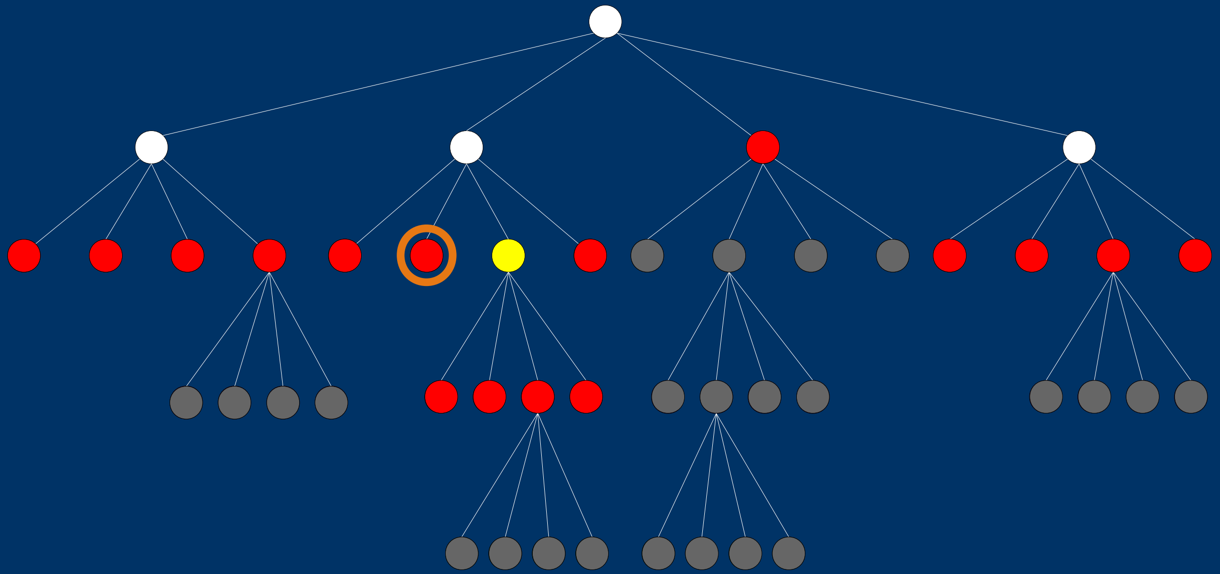
Computational Challenges

- Large spacial dynamic range: > 100 Mpc to < 1 kpc
  - Hierarchical, adaptive gravity solver is needed
- Large temporal dynamic range: 10 Gyr to < 1 Myr
  - Multiple timestep algorithm is needed
- Gravity is a long range force
  - Hierarchical information needs to go across processor domains

Basic Gravity algorithm ...

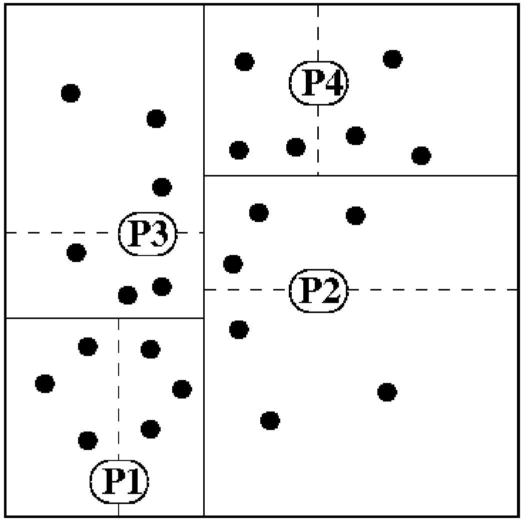
- Newtonian gravity interaction
  - Each particle is influenced by all others: $O(n^2)$ algorithm
- Barnes-Hut approximation: $O(n\log n)$
  - Influence from distant particles combined into center of mass

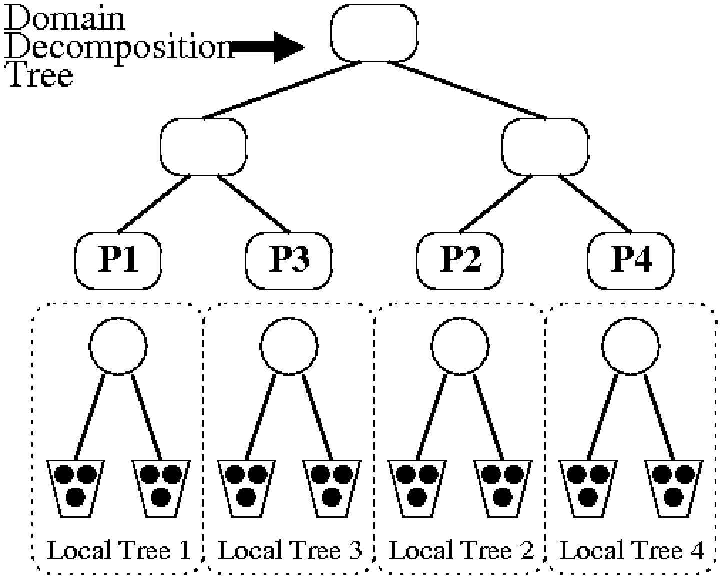




Legacy Code: PKDGRAV/GASOLINE

- Originally implemented on KSR2
  - Ported to: PVM, pthreads, MPI, T3D, CHARM++
- KD tree domain decomposition/load balancing
- Software cache: latency amortization

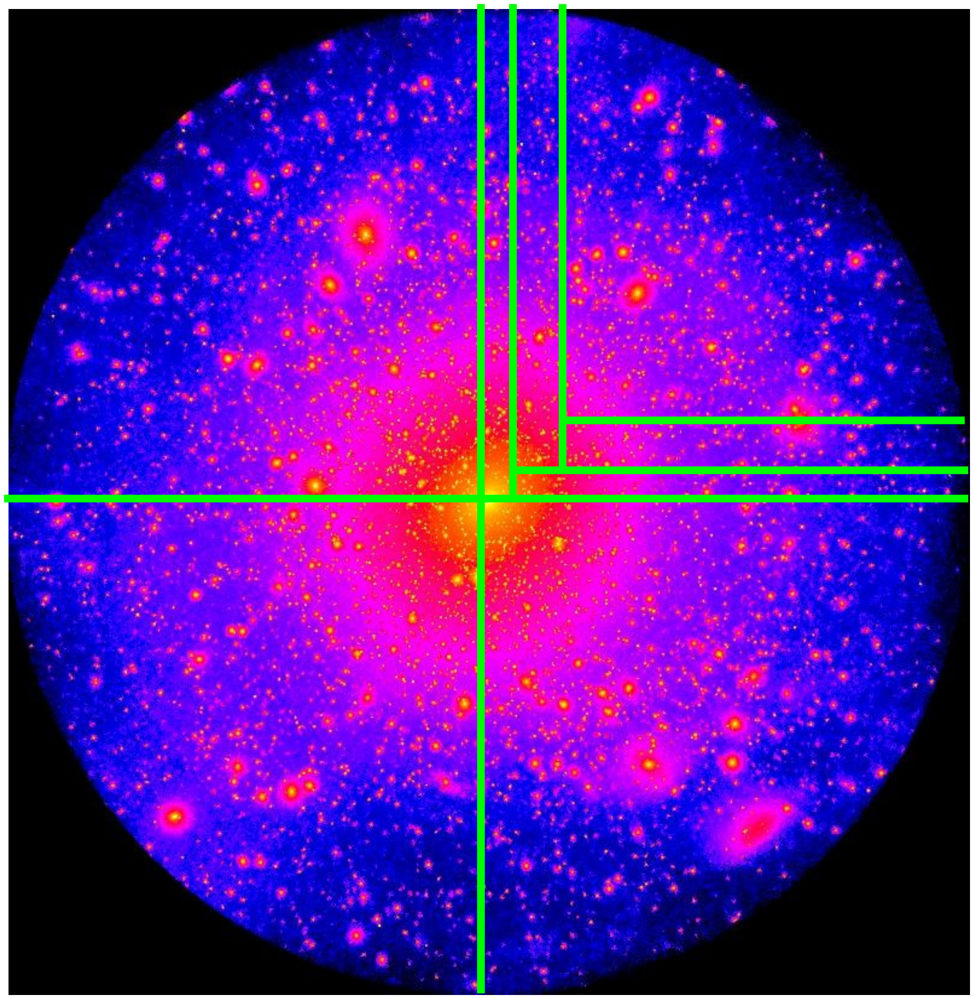




PKDGRAV/GASOLINE Issues

- Load balancing creates more work, systematic errors.
- Multistep domain decomposition
- Latency amortization, but not hiding via software cache
  - Fast network is required
  - SPH scaling is poor
- Porting: MPI became the standard platform

Clustering and Load Balancing



Charm++ features

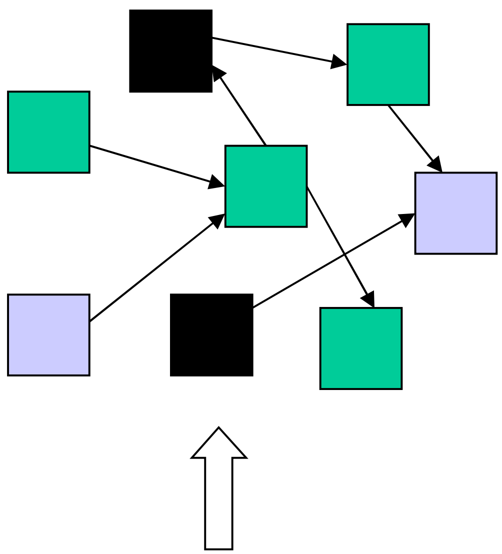
- “Automatic”, measurement-based load balancing.
- Natural overlap of computation and communication
- Not hardwired to a given data structure.
- Object Oriented: reuse of existing code.
- Portable
- NAMD: molecular dynamics is similar.
- Approachable group!

Building a Treecode in CHARM++: Porting GASOLINE

- AMPI port of GASOLINE
  - Very straightforward
  - Adding Virtual Processors gave poor performance: separate caches increased communication
- CHARM++ port of GASOLINE
  - Good match to RMI design
  - Charm++ allowed some minor speed improvements
  - Still, more than one element/processor does not work well

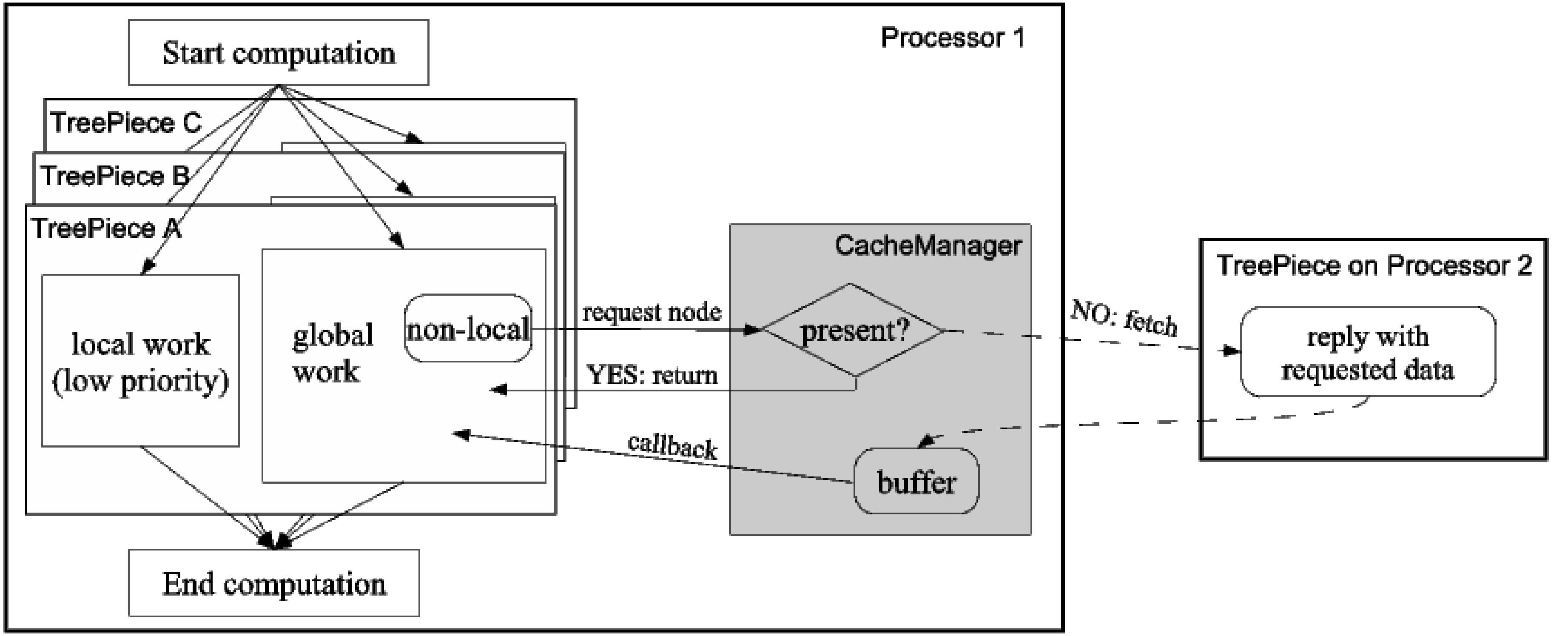
Building a Treecode in CHARM++: Starting afresh

- Follow Charm++ paradigm: send particle data as walk crossed boundaries
- Very large number of messages.
- Back to software cache



User View

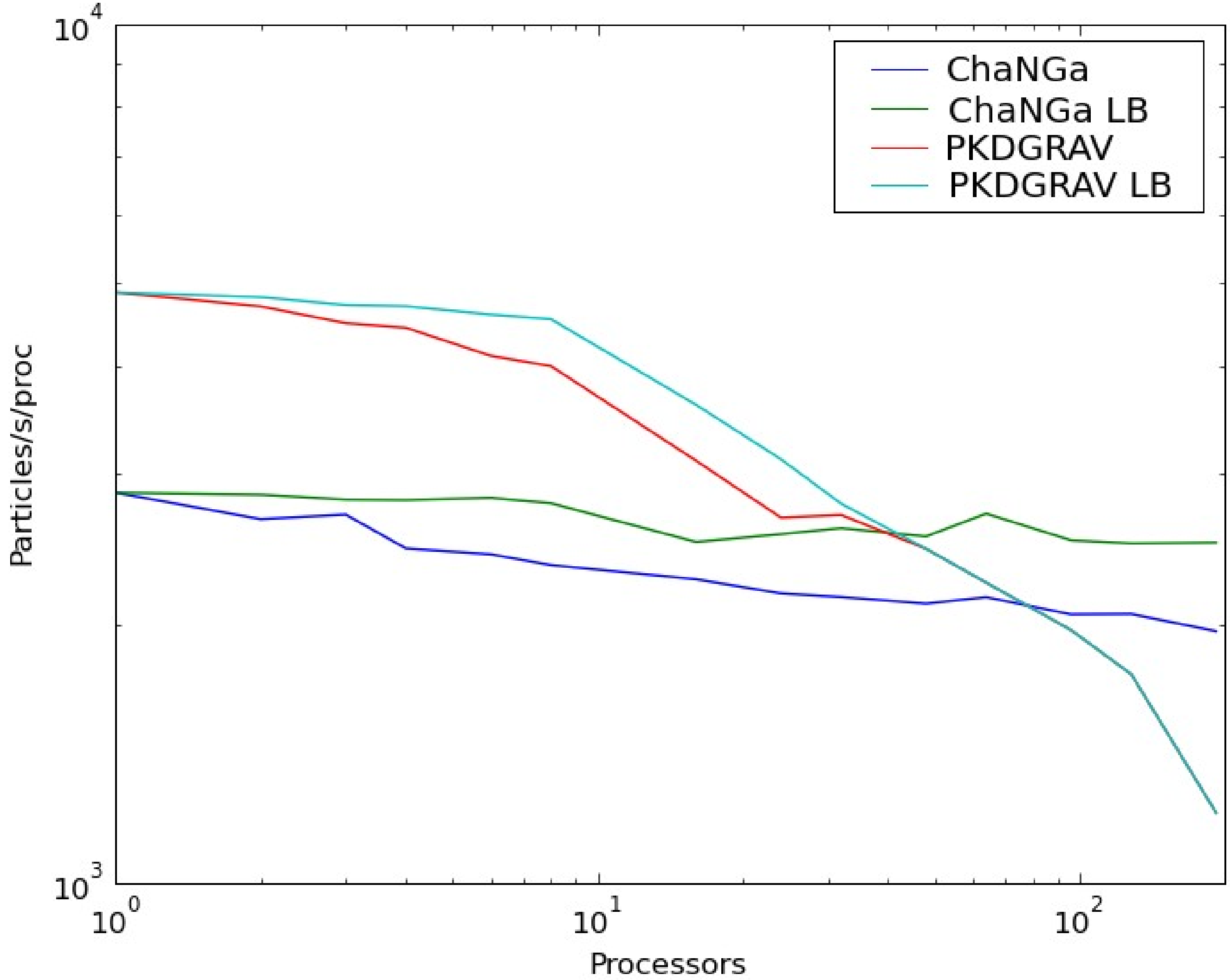
Overall Algorithm



ChaNGa Features

- Tree-based gravity solver
- High order multipole expansion
- Periodic boundaries (if needed)
- Individual multiple timesteps
- Dynamic load balancing with choice of strategies
- Checkpointing (via migration to disk)
- Visualization

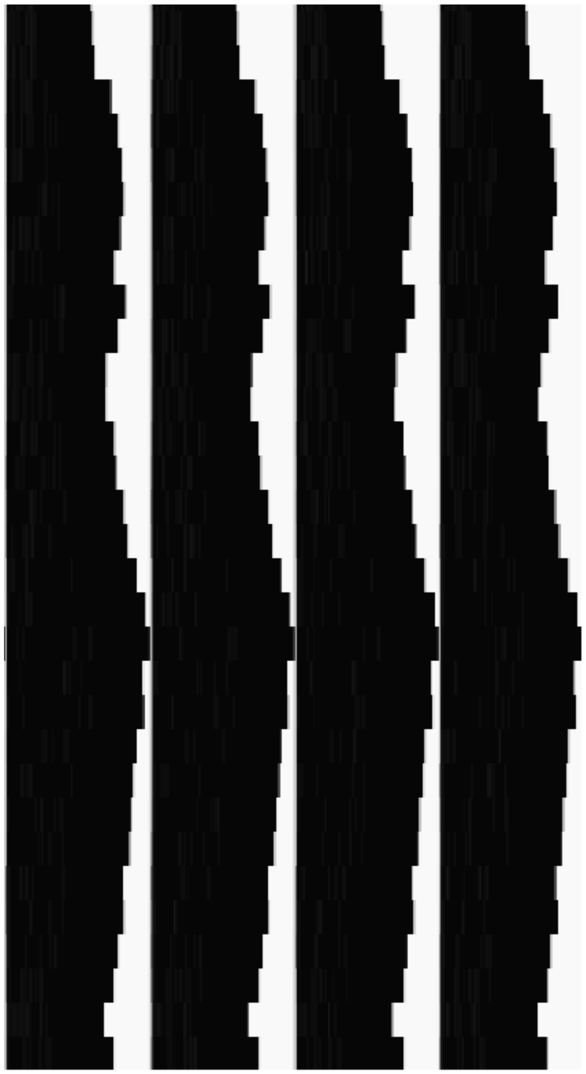
Zoom-in Scaling



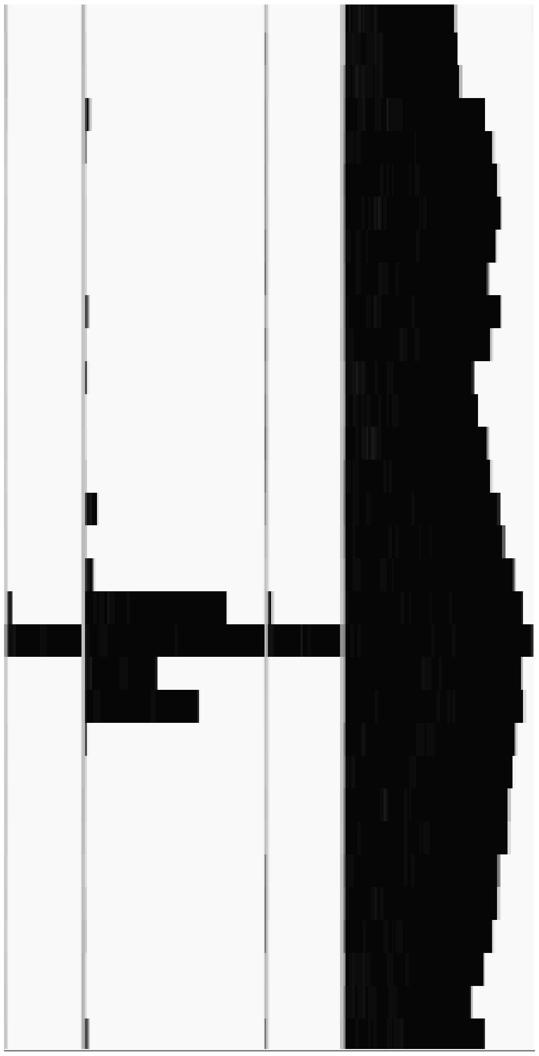
Multistep Loadbalancer

- Use Charm++ measurement based load balancer
- Modification: provide LB database with information about timestepping.
  - “Large timestep”: balance based on previous large step
  - “Small step” balance based on previous small step
  - Maintains principle of persistence

Results on 3 rung example



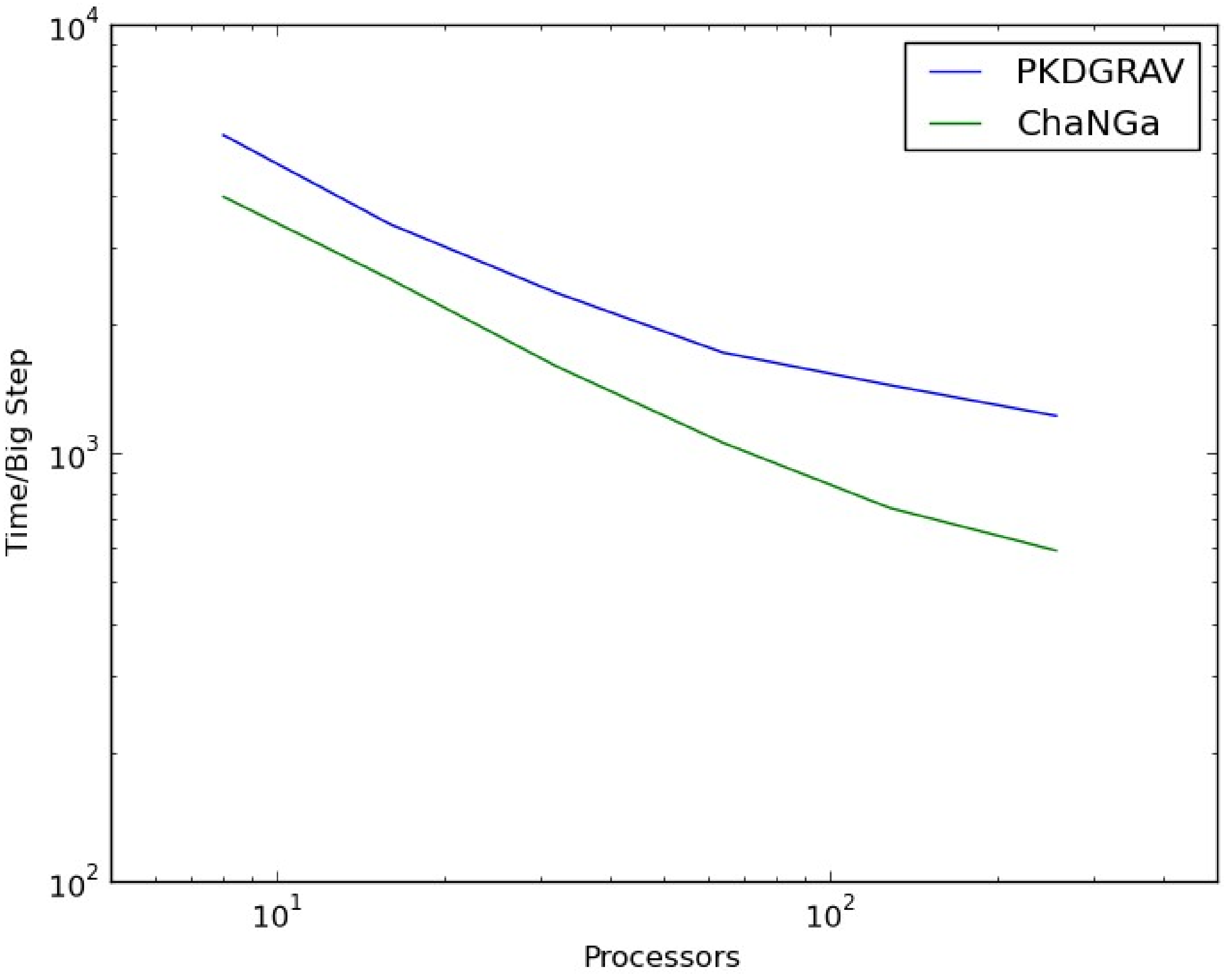
613s



429s

228s

Multistep Scaling



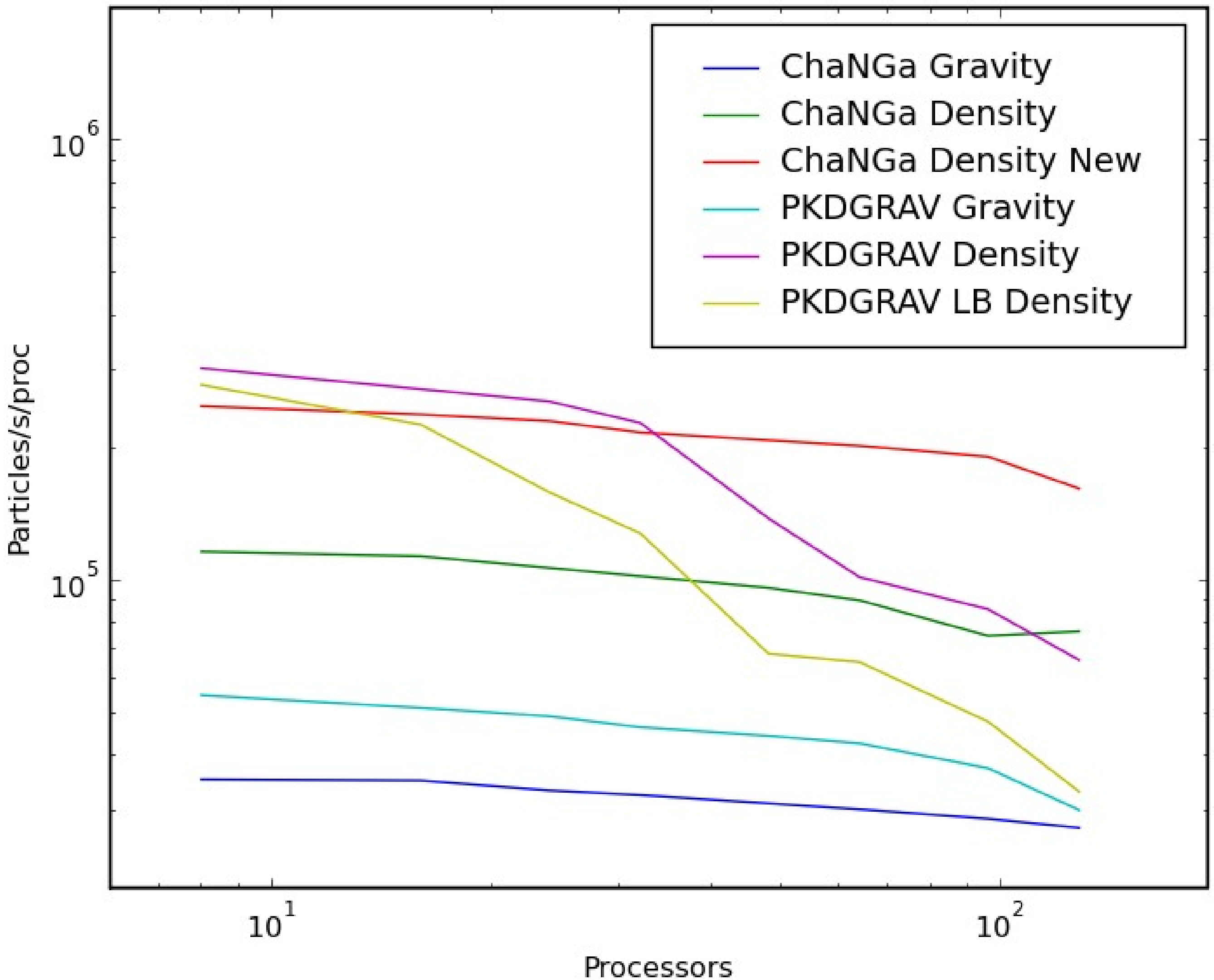
Smooth Particle Hydrodynamics

- Making testable predictions needs Gastrophysics
    - High Mach number
    - Large density contrasts
- Gridless, Lagrangian method
- Galilean invariant
- Monte-Carlo Method for solving Navier-Stokes equation.
- Natural extension of particle method for gravity.

SPH Challenges

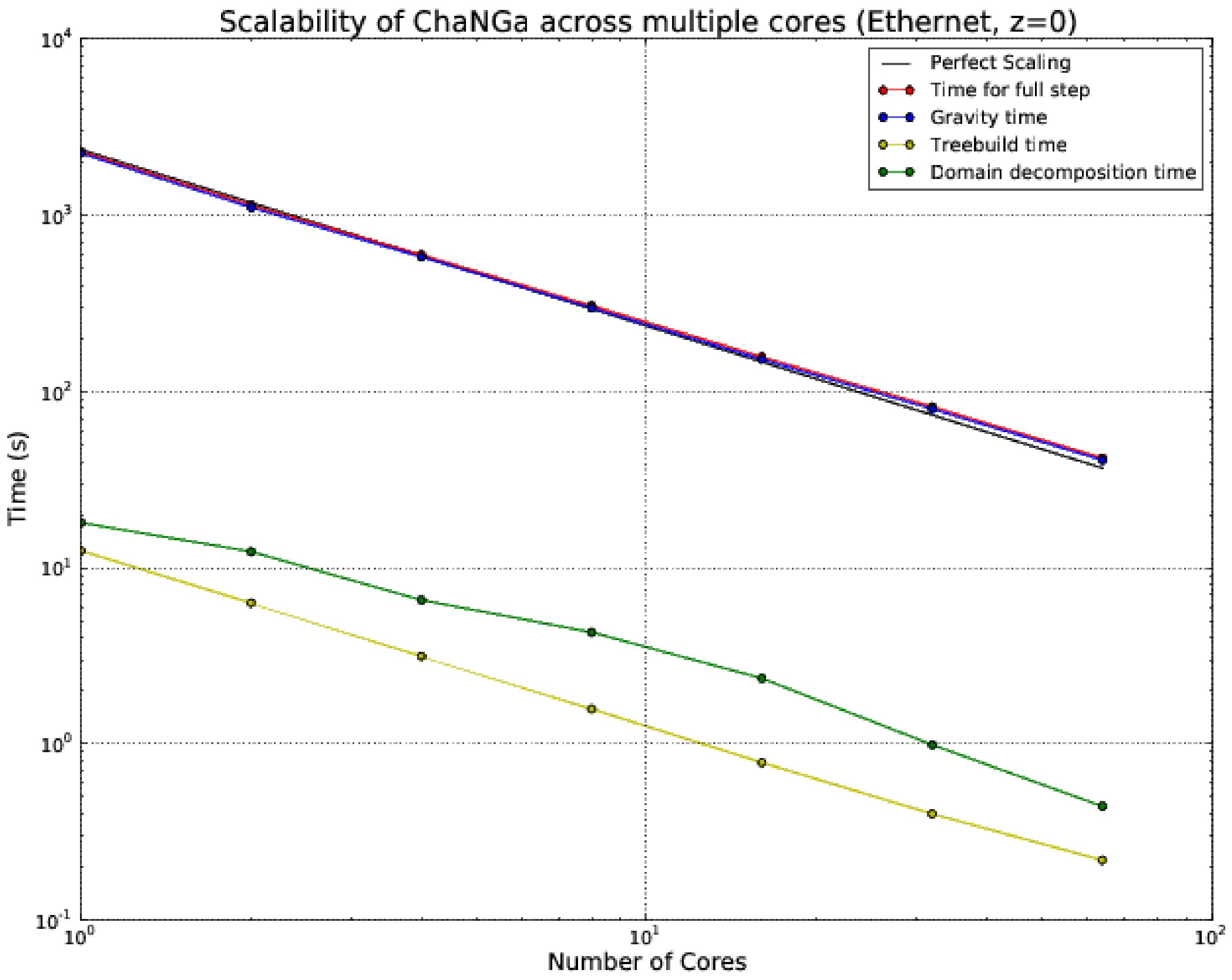
- Increased density contrasts/time stepping.
- K-nearest neighbor problem.
  - Trees!
- More data/particle than gravity
- Less computation than gravity
- Latency much more noticable

SPH Scaling



ChaNGa Gravity
ChaNGa Density
ChaNGa Density New
PKDGRAV Gravity
PKDGRAV Density
PKDGRAV LB Density
Particles/s/proc
10^6
10^5
10^1
10^2
Processors

Ethernet scaling



Current uses

- Large scale structure
  - Dynamics of gas in galaxy clusters
  - Galaxy formation in the local Universe
- Galactic dynamics
  - Formation of nuclear star clusters
  - Disk heating from substructure
- Protoplanetary disks
  - Thermodynamics and radiative transfer

Future

- More Physics
  - Cooling/Star formation recipes
  - Charm++ allows reuse of PKDGRAV code
- Better gravity algorithms
  - New domain decomposition/load balancing strategies
  - Multicore/heterogeneous machines
- Other Astrophysical problems
  - Planet formation
  - Planetary rings

Charm++ features: reprise

- “Automatic”, measurement-based load balancing.
  - But needs thought and work
- Migration to GPGPU and SMP
- Object Oriented: reuse of existing code.
- Approachable group
  - Enhance Charm++ to solve our problems.

Summary

- Cosmological simulations provide a challenges to parallel implementations
  - Non-local data dependencies
  - Hierarchical in space and time
- ChaNGa has been successful in addressing this challenges using Charm++ features
  - Message priorities
  - New load balancers