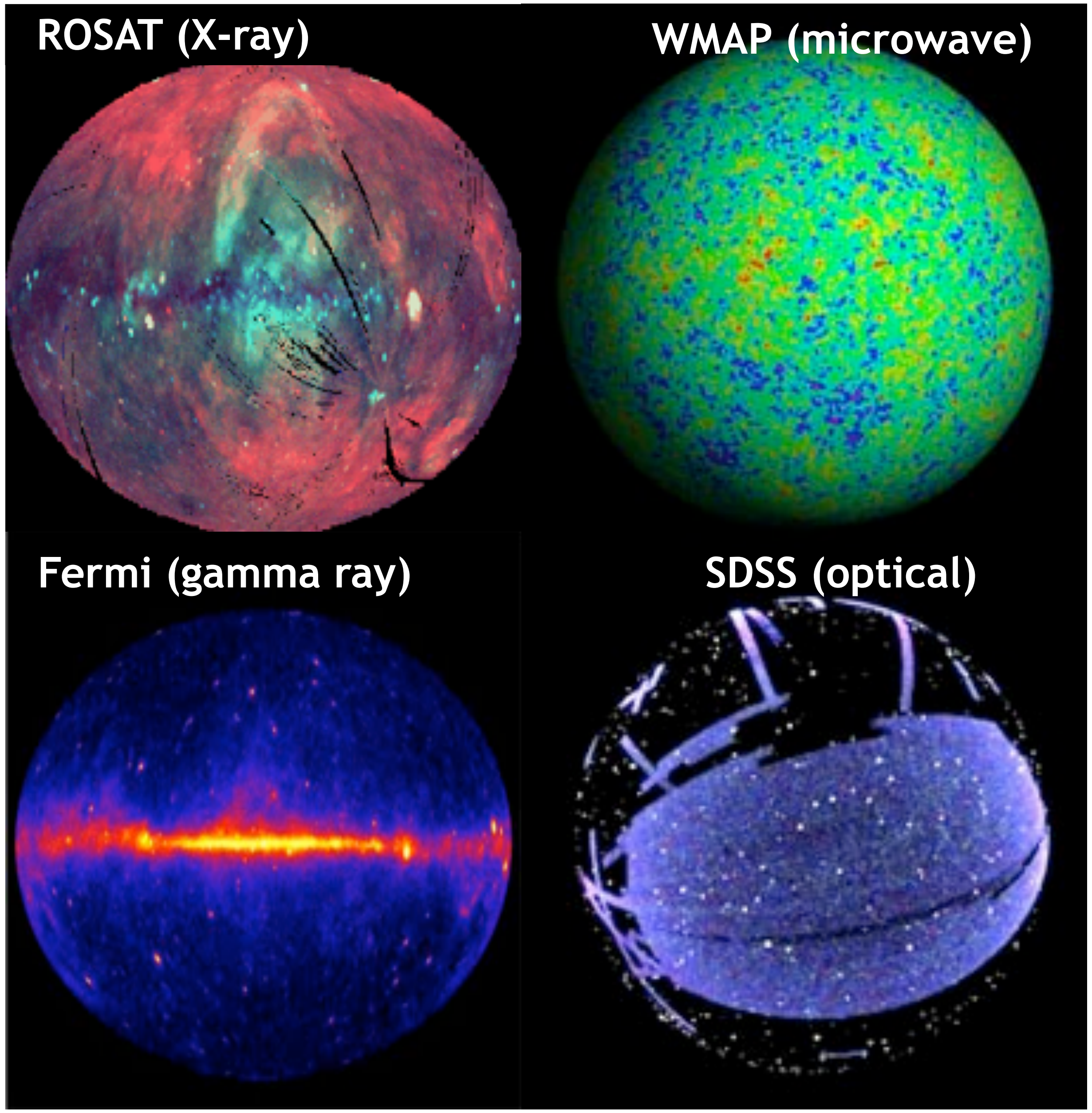
Computing the Dark Universe

Salman Habib
High Energy Physics Division
Mathematics & Computer Science Division
Argonne National Laboratory

Collaborators:
ANL: S. Bhattacharya, H. Finkel, K. Heitmann, A. Pope
LANL: J. Ahrens, D. Daniel, P. Fasel, N. Frontiere, P. McCormick, P. Sathre, J. Woodring
UC Berkeley/LBNL: J. Carlson, Z. Lukic, M. White

Modern Cosmology and Sky Maps

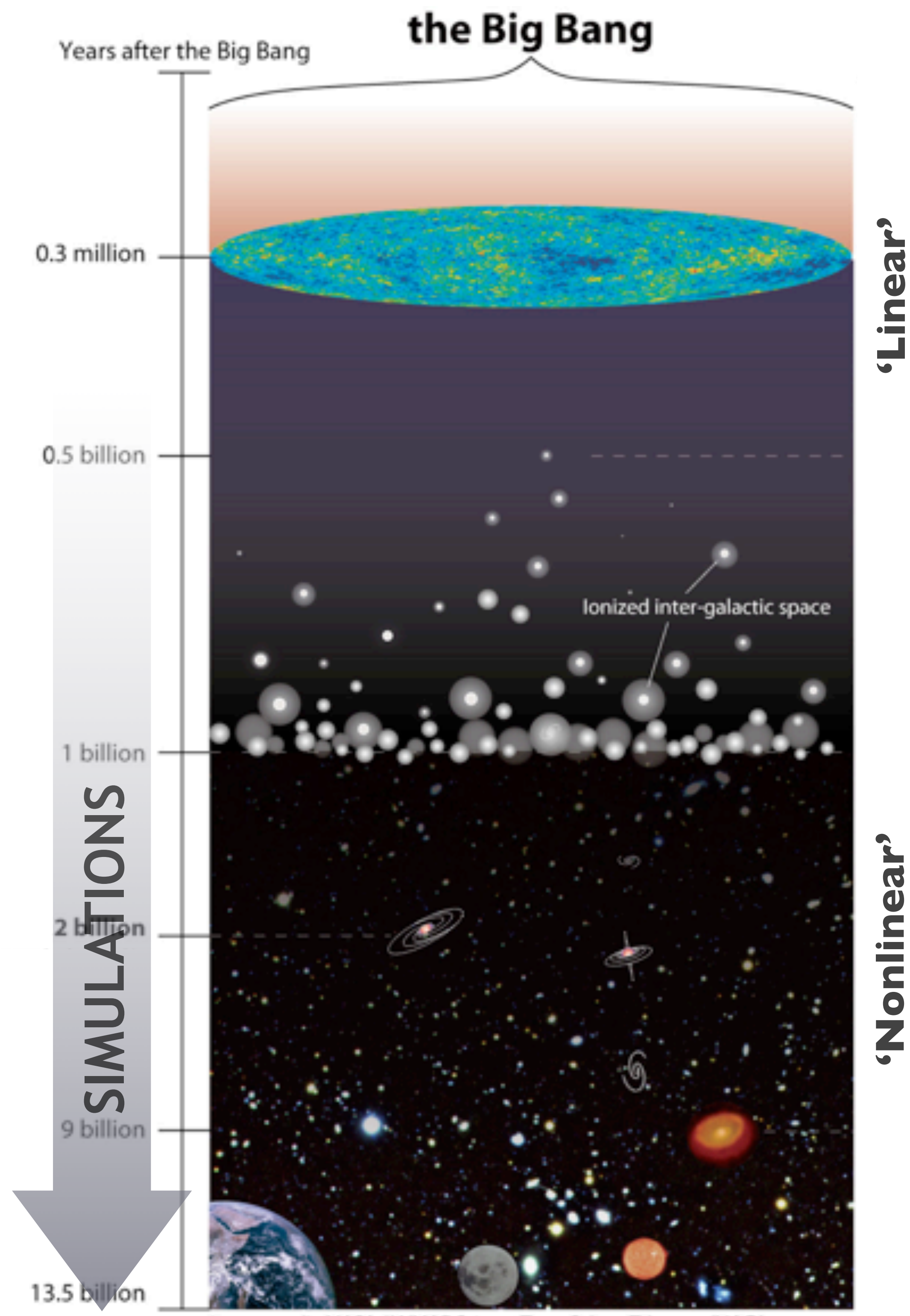
- Modern cosmology is the story of mapping the sky in multiple wavebands
- Maps cover measurements of objects (stars, galaxies) and fields (CMB temperature)
- Maps can be large (SDSS has~200 million galaxies, many billions for LSST)
- Statistical analysis of sky maps
- All precision cosmological analyses constitute a statistical inverse problem: from sky maps to scientific inference
- Therefore: No cosmology without (large-scale) computing



Explosion of information from sky maps: Precision Cosmology

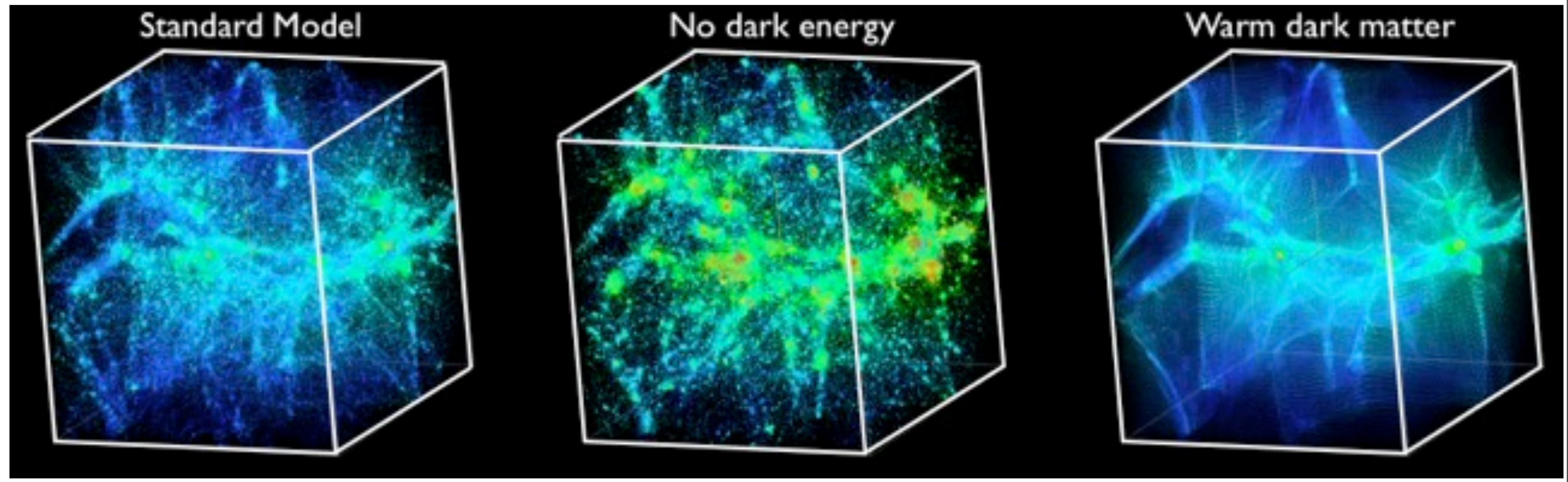
Structure Formation: The Basic Paradigm

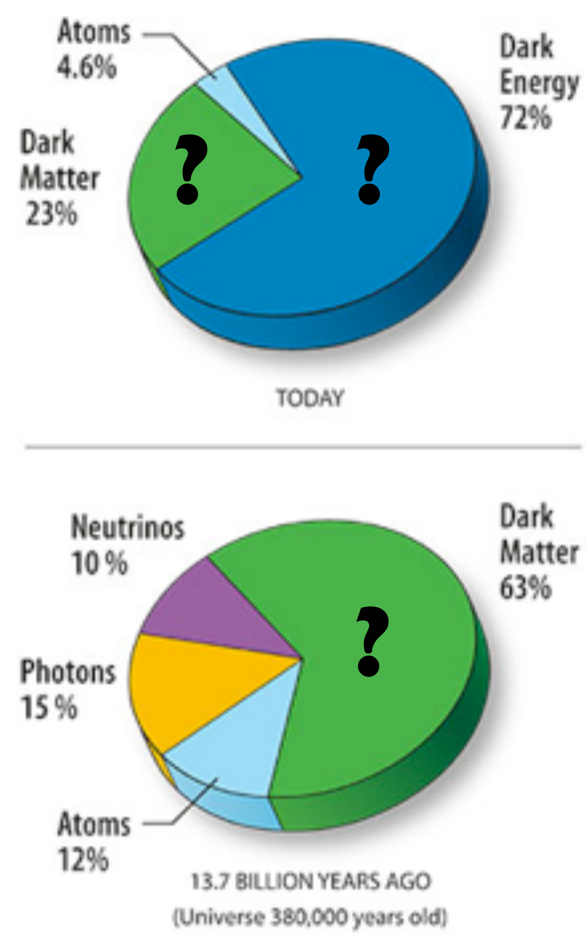
- **Solid understanding of structure formation; success underpins most cosmic discovery**
  - Initial conditions laid down by inflation
  - Initial perturbations amplified by gravitational instability in a dark matter-dominated Universe
  - Relevant theory is gravity, field theory, and atomic physics ('first principles')
- **Early Universe:** Linear perturbation theory very successful (CMB)
- **Latter half of the history of the Universe:** Nonlinear domain of structure formation, impossible to treat without large-scale computing



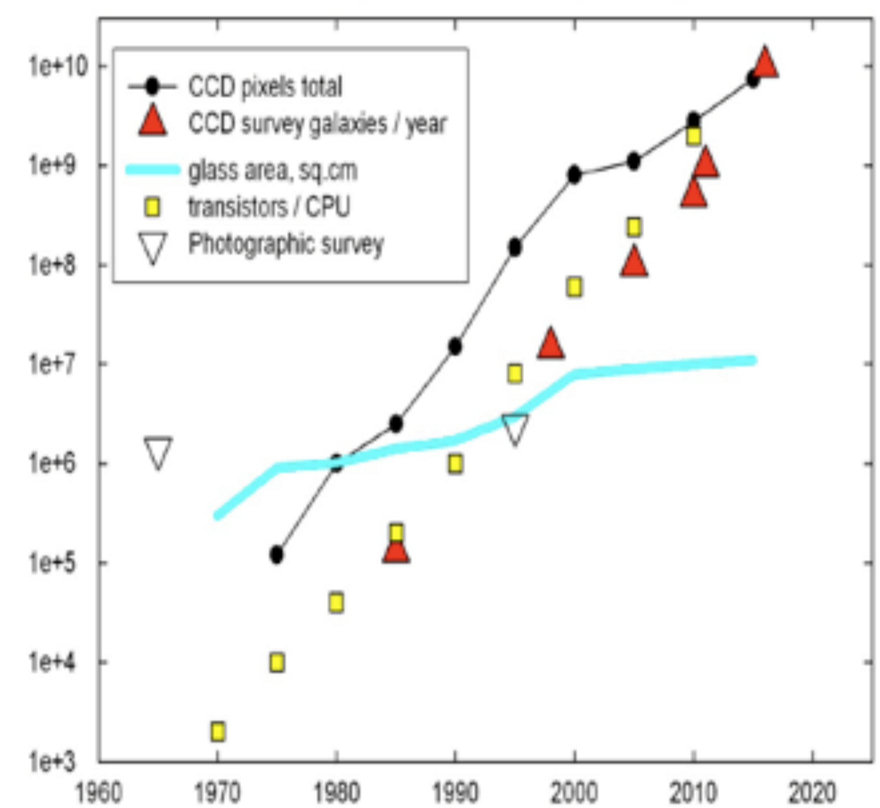
Precision Cosmology: “Inverting” the 3-D Sky

- **Cosmological Probes:** Measure geometry and presence/growth of structure
- **Examples:** Baryon acoustic oscillations (BAO), cluster counts, CMB, weak lensing, galaxy clustering, --
- **Standard Model:** Verified at the 5-10% level across multiple observations
- **Future Targets:** Aim to control survey measurements to the ~1% level, can theory and simulation keep up?





Cosmic content pie charts



Optical survey ‘Moore’s Law’

Simulating the Universe

- Gravity dominates at large scales, key task: solve the Vlasov-Poisson equation (VPE)
- VPE is 6-D and cannot be solved as a PDE
- N-body methods; gravity has (i) no shielding but is (ii) naturally Lagrangian
- Are errors controllable?
- At smaller scales add gas physics, feedback, etc. (subgrid modeling inevitable)
- Calibrate simulations against observations

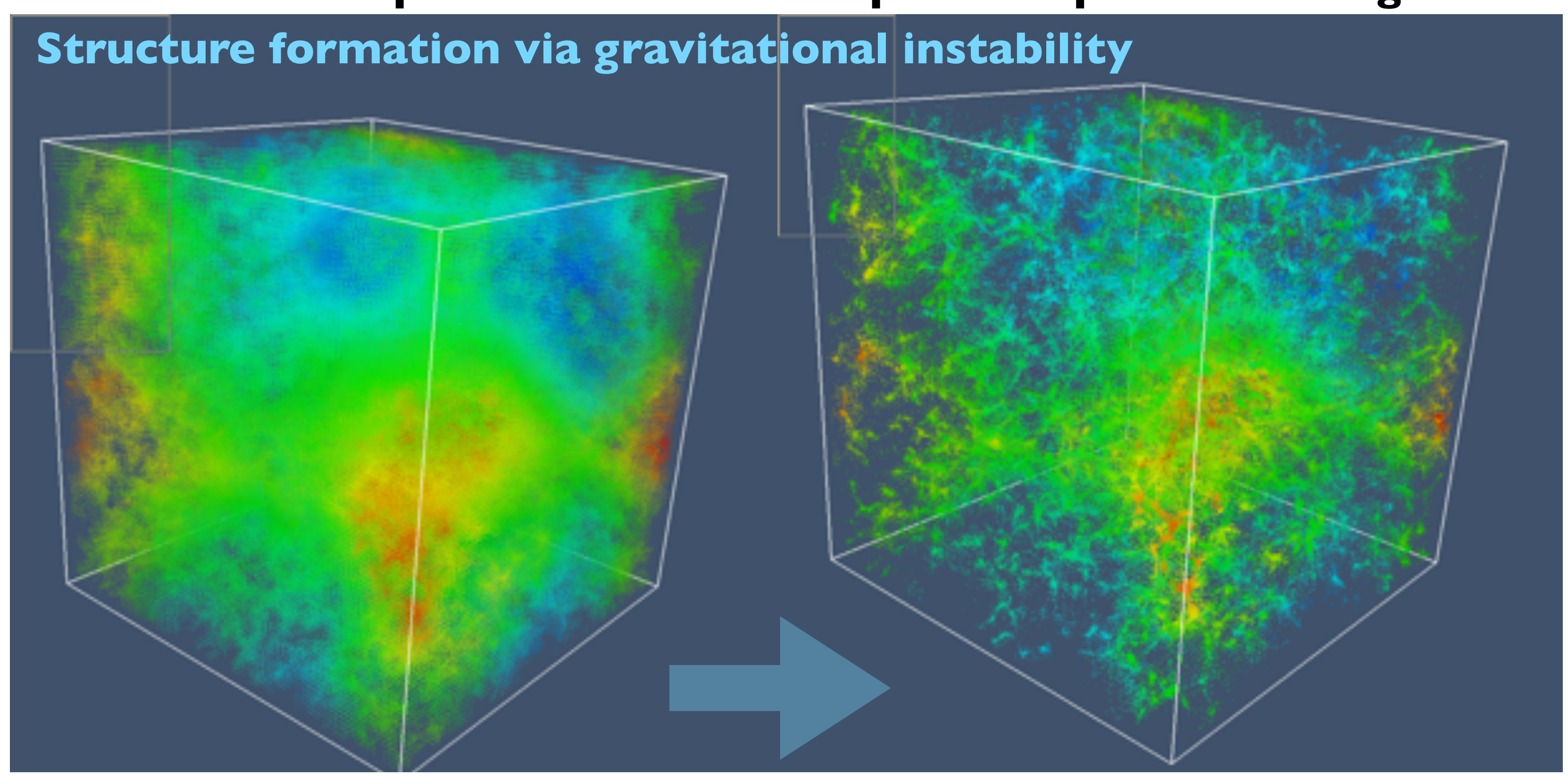
$$\frac{\partial f_i}{\partial t} + \dot{\mathbf{x}}\frac{\partial f_i}{\partial \mathbf{x}} - \nabla\phi\frac{\partial f_i}{\partial \mathbf{p}} = 0, \qquad \mathbf{p} = a^2\dot{\mathbf{x}},$$

$$\nabla^2\phi = 4\pi G a^2(\rho(\mathbf{x},t) - \langle\rho_{\rm dm}(t)\rangle) = 4\pi G a^2\Omega_{\rm dm}\delta_{\rm dm}\rho_{\rm cr},$$

$$\delta_{\rm dm}(\mathbf{x},t) = (\rho_{\rm dm} - \langle\rho_{\rm dm}\rangle)/\langle\rho_{\rm dm}\rangle),$$

$$\rho_{\rm dm}(\mathbf{x},t) = a^{-3}\sum_i m_i \int d^3\mathbf{p} f_i(\mathbf{x},\dot{\mathbf{x}},t).$$

Cosmological Vlasov-Poisson Equation: A 'wrong-sign' electrostatic plasma with time-dependent particle 'charge'



An Early Simulation

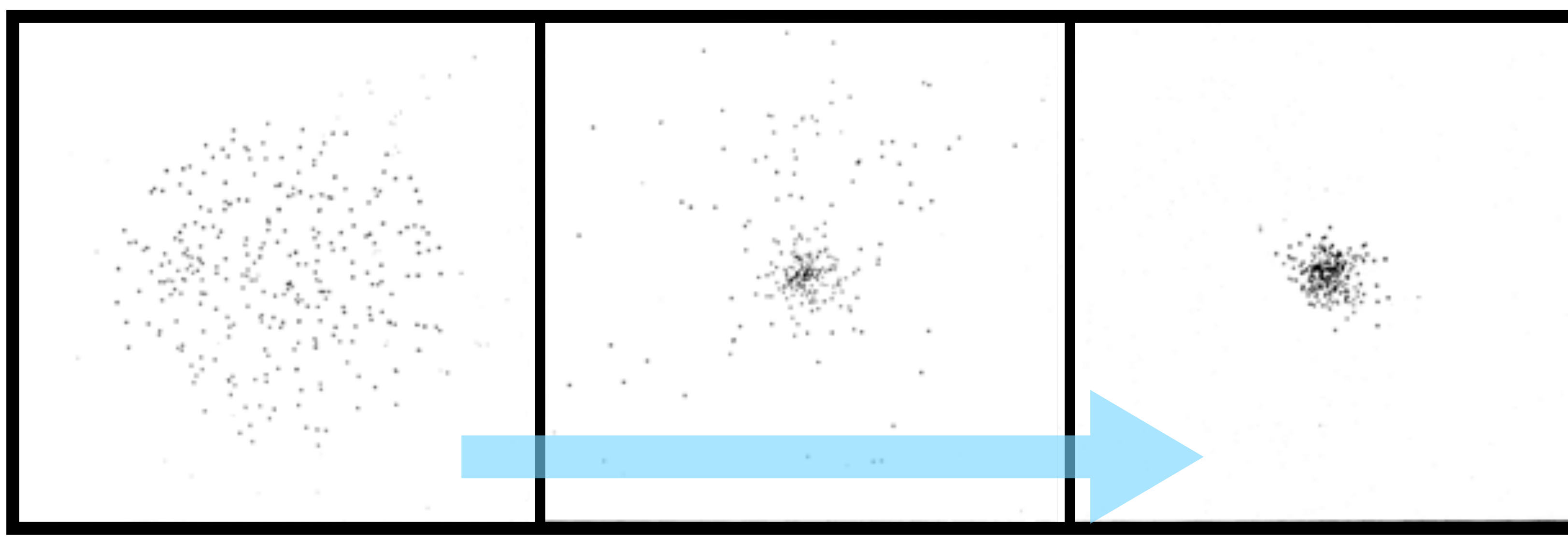
- Suite of 300 (and less) particle simulations
- Run on a CDC 3600, ~1Mflops, 32KB+ at LANL
- Is nine orders of magnitude improvement in both performance and memory good enough for precision cosmology?

THE ASTRONOMICAL JOURNAL VOLUME 75, NUMBER 1 FEBRUARY

Structure of the Coma Cluster of Galaxies*

P. J. E. Peebles†
Palmer Physical Laboratory, Princeton University, Princeton, New Jersey
(Received 7 October 1969)

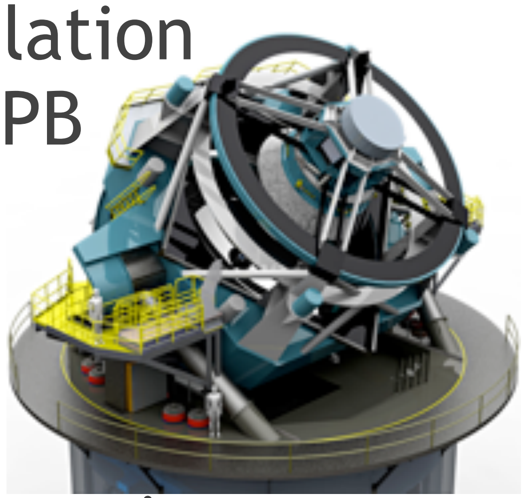
In some cosmologies, a cluster of galaxies is imagined to be a gravitationally bound system which, in analogy with the formation of the Galaxy, originated as a collapsing protocluster. It is shown that a numerical model based on this picture is consistent with the observed features of the Coma Cluster of galaxies. The cluster mass derived from this model agrees with previous values; however, an analysis of the observational uncertainty within the framework of the model shows that the derived mass could be consistent with the estimated total mass provided by the galaxies in the cluster.



"The Universe is far too complicated a structure to be studied deductively, starting from initial conditions and solving the equations of motion."

Robert Dicke (Jayne Lectures, 1969)

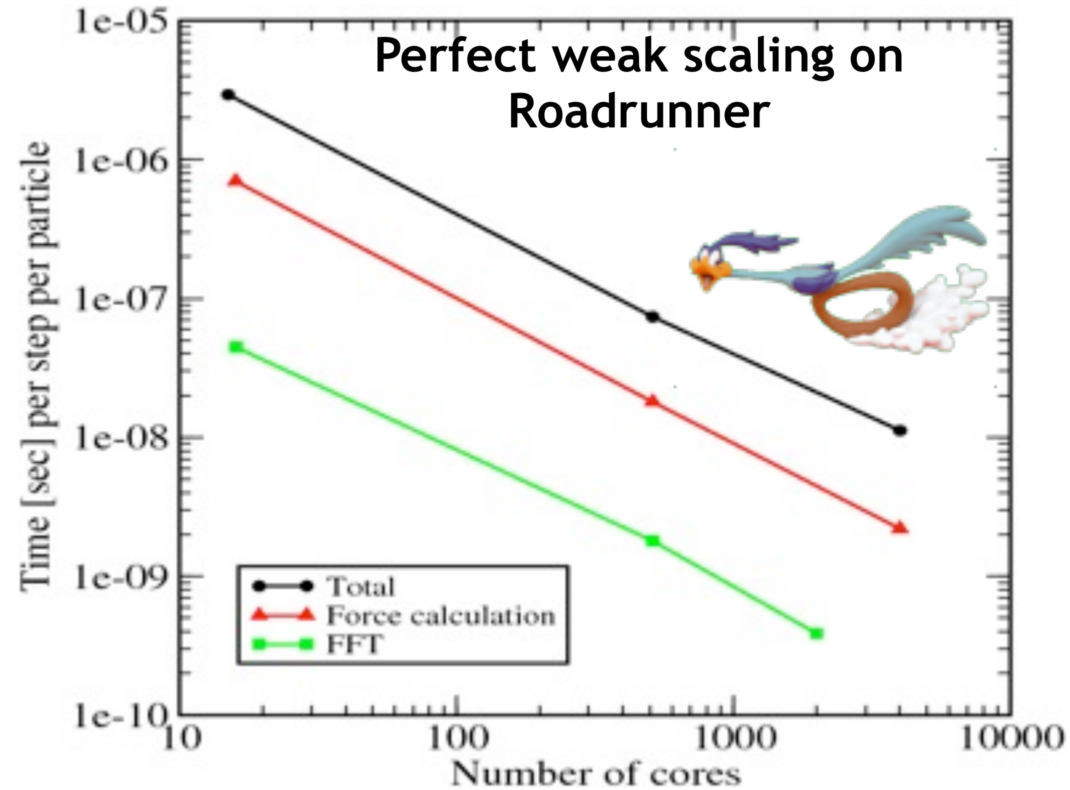
Computing the Universe: Simulating Surveys

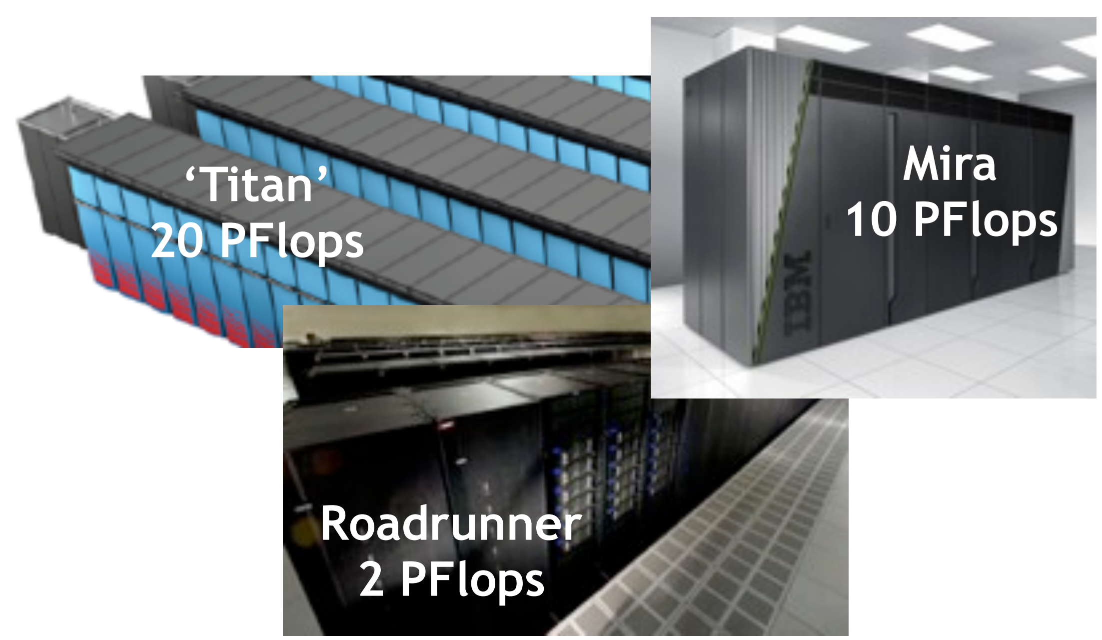


- **Simulation Volume:** Large survey sizes impose simulation volumes ~ (**3** Gpc)^3, memory required ~**100** TB **-- 1** PB
- **Number of Particles:** Mass resolutions depend on ultimate object to be resolved, ~10^8 --10^{10} solar masses, N~10^{11}--10^{12}
- **Force Resolution:** ~kpc, yields a (global) spatial dynamic range of 10^6
- **Hydrodynamics/Sub-Grid Models:** Phenomenological treatment of gas physics and feedback greatly adds to computational cost
- **Throughput:** Large numbers of simulations required (**100**'s **--1000**'s), development of analysis suites, and emulators; peta-exascale computing exploits
- **Data-Intensive-SuperComputing:** End-to-End simulations and observations must be brought together in a DISC environment (theory-observation feedback)

Hardware-Accelerated Cosmology Code (HACC) Framework

- **Architecture Challenge:** HPC is rapidly evolving (clusters/BG/CPU+GPU/MIC --)
- **Code for the Future:** Melds optimized performance, low memory footprint, embedded analysis, and scalability
- **Implementation:** Long/short-range force matching with spectral force-shaping (long-range=PM, short-range=PP, Tree)
- **Key Features:** Hybrid particle/grid design, particle overloading, spectral operators, mixed-precision, node-level ‘plug-ins’, ~50% of peak Flops
- **Cross-Platform:** Designed for all current and future supercomputing platforms
- **Embedded Analysis:** High performance with low I/O and storage requirement





Habib et al. 2009, Pope et al. 2010

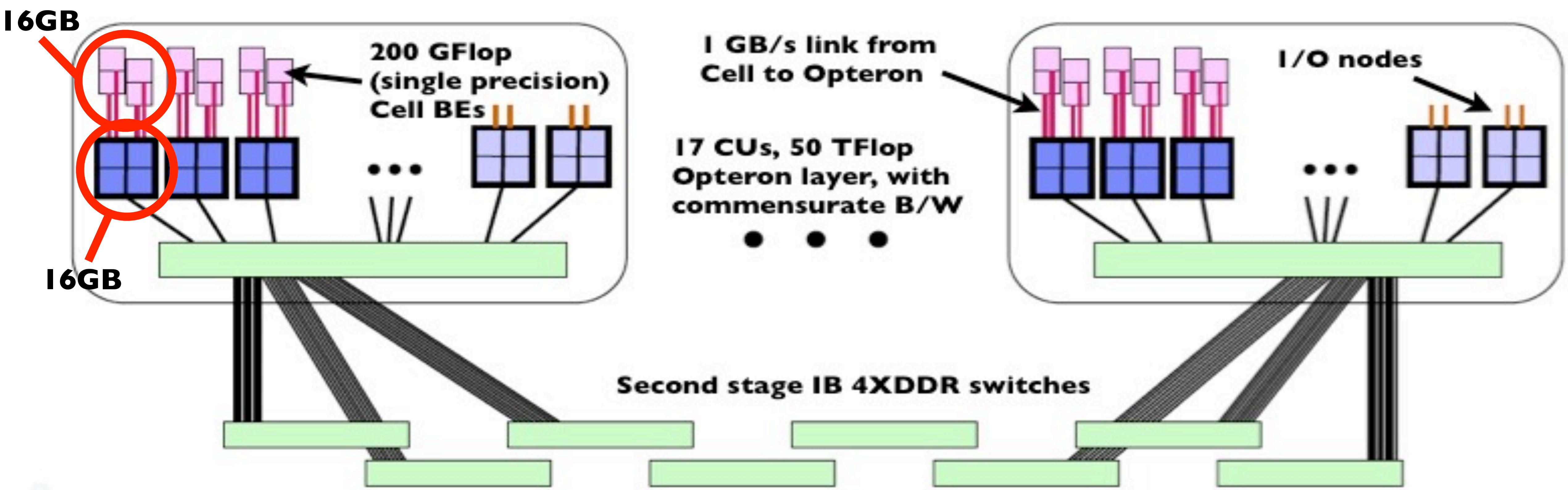
HACC Design Features

- **New Framework:** Not a port of an older code (too difficult)
- **Two-Layer Design:** Anticipates communication bottleneck between CPU and accelerator layers
- **Compute Sharing:** Compute complexity shifted to CPU+MPI layer (new algorithms), simple brute force computations assigned to accelerators, use mixed precision (CPU, double; accelerator, single)
- **Memory Trade-Off:** Small memory overhead used to reduce inter-layer communication and improve modularity
- **Cross-Platform:** Aimed at current and future supercomputing platforms using ‘plug-in’ short-range force modules optimized for a given nodal architecture (and using different algorithms)
- **In Situ Analysis:** Significant attention paid to ‘on the fly’ analysis methods to reduce I/O and storage; code design allows for essentially ‘serial’ methods to be trvially parallelized
- **Simplicity:** Relatively straightforward approach

HACC Beginnings: Roadrunner Universe Project

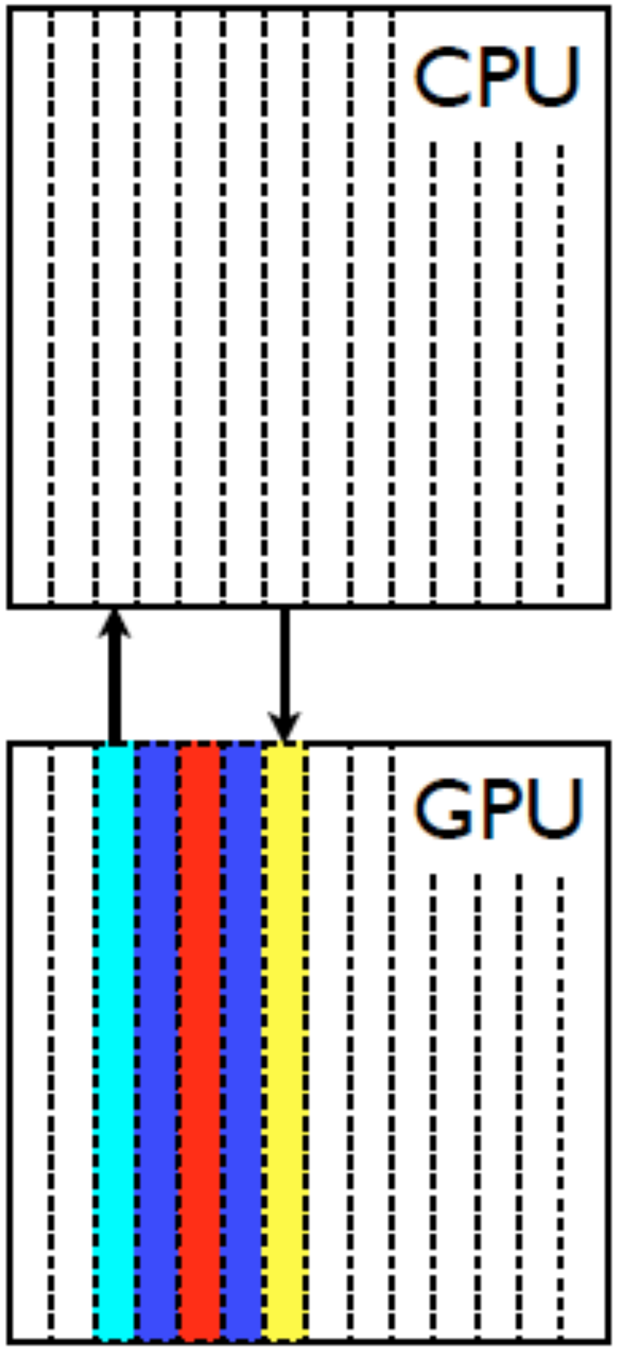
Andrew White Dec 7, 2007 + What if you had a petaflop/s

- Hybrid machine architecture, out of balance communication (50-100) and performance (20)
- Balanced memory (CPU=Cell)
- Multi-level programming paradigm
- Prototype for exascale code design problems
- Scalable approach extensible to all next-generation architectures (BG/Q, CPU/GPU, --)



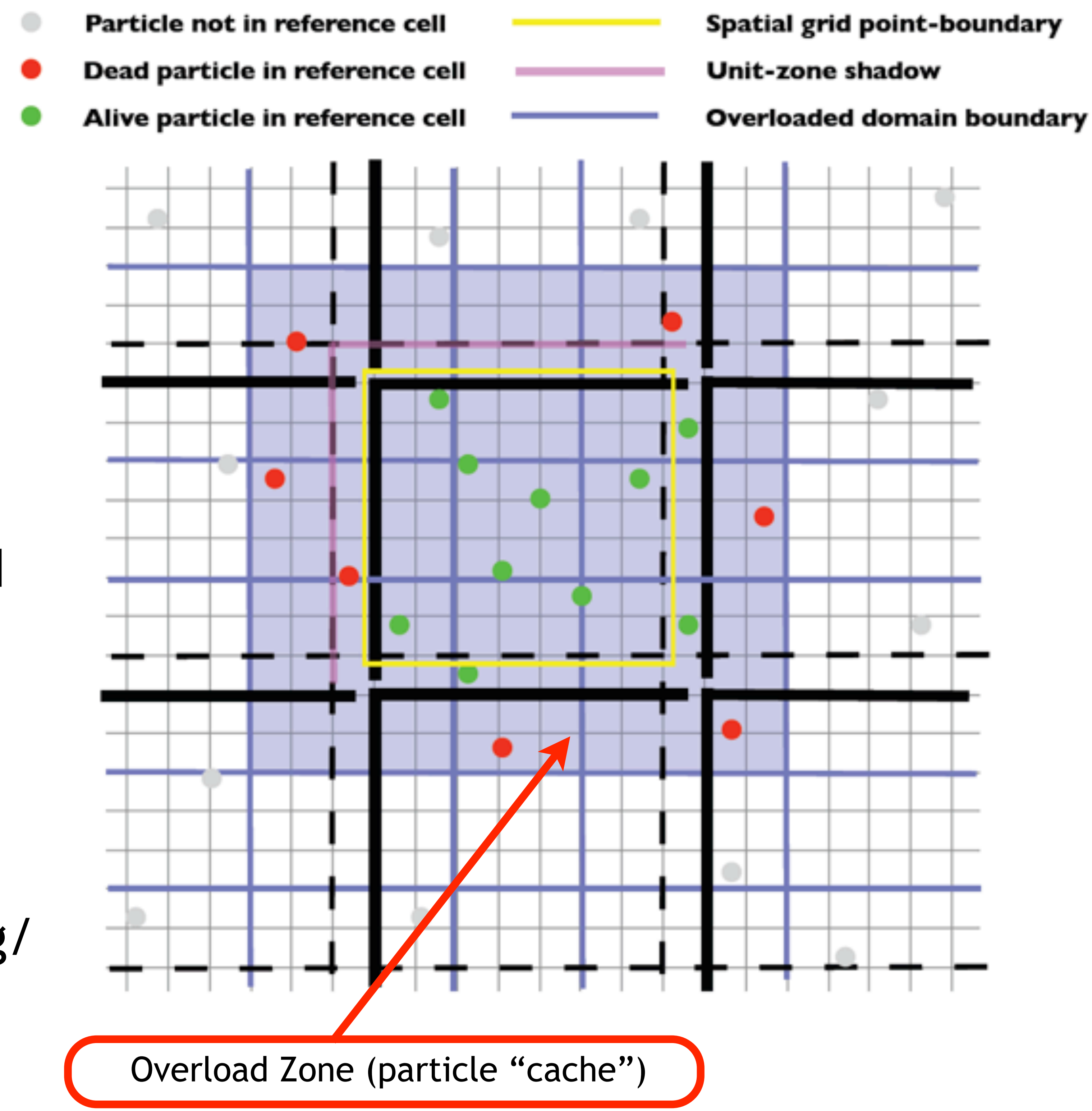
HACC Example 2: CPU+GPU

- CPU/GPU performance and communication out of balance, unbalanced memory (CPU/main memory dominates)
- Multi-level programming (mitigate with OpenCL)
- Particles in CPU main memory, CPU does low flop/byte operations
- Stream slabs through GPU memory (pre-fetches, asynchronous result updates)
- Data-parallel kernel execution
- Many independent work units per slab -- many threads, efficient scheduling, good performance achieved (improves on Cell)
- Scalability of HACC is the same across all ‘nodal’ variants



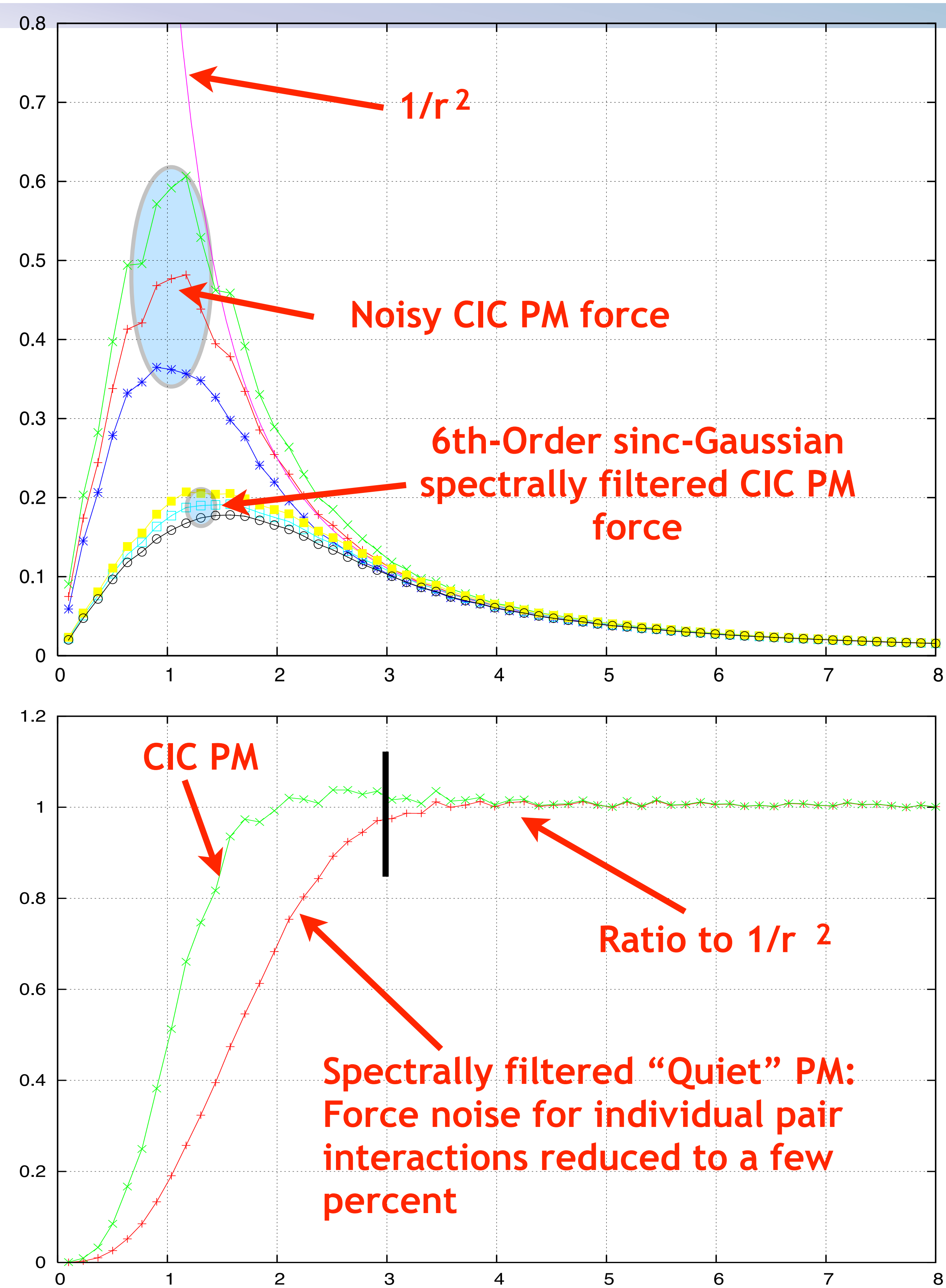
HACC Algorithmic Details 1

- Solve compute imbalance: Split problem into long-range and short-range force updates
- Long-range handled by a grid-based Poisson solver
- Direct particle-particle short-range interactions
- Simplify and speed-up Cell computational tasks
- Reduce CPU/Cell traffic to avoid PCIE bottleneck: use simple CIC to couple particles to the grid, followed by spectral filtering on the grid
- Reduce inter-node particle communication: particle caching/replication (ghost zone analog)
- 'On the fly' analysis and visualization to reduce I/O

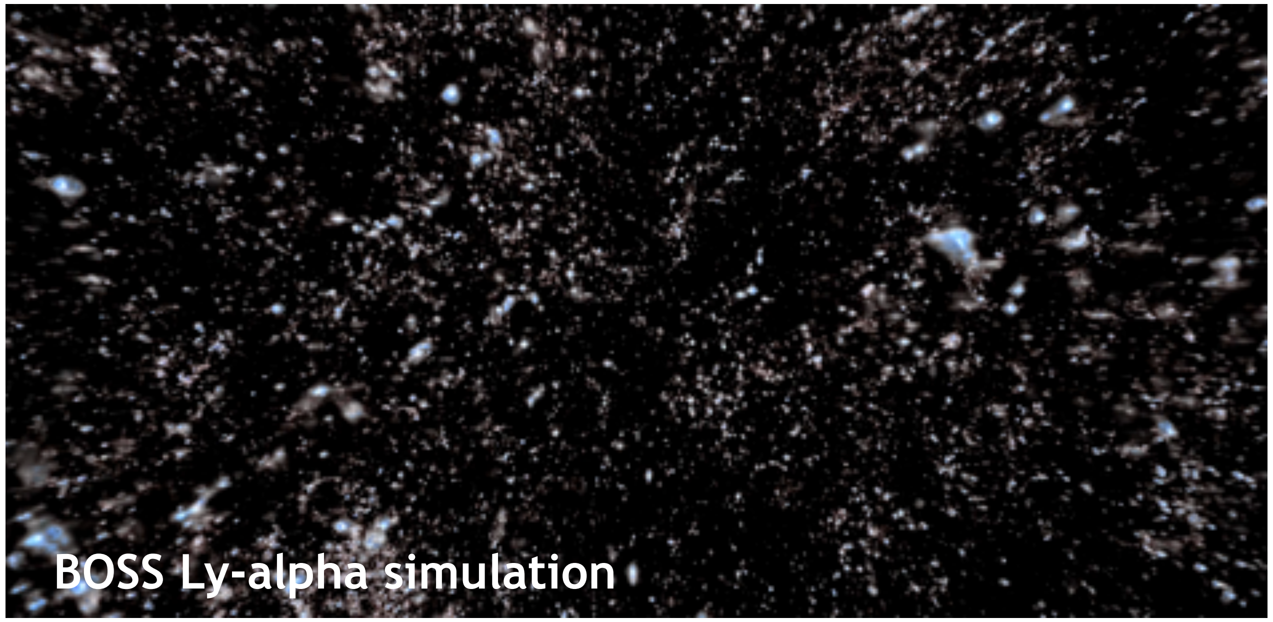


HACC Algorithmic Details 2

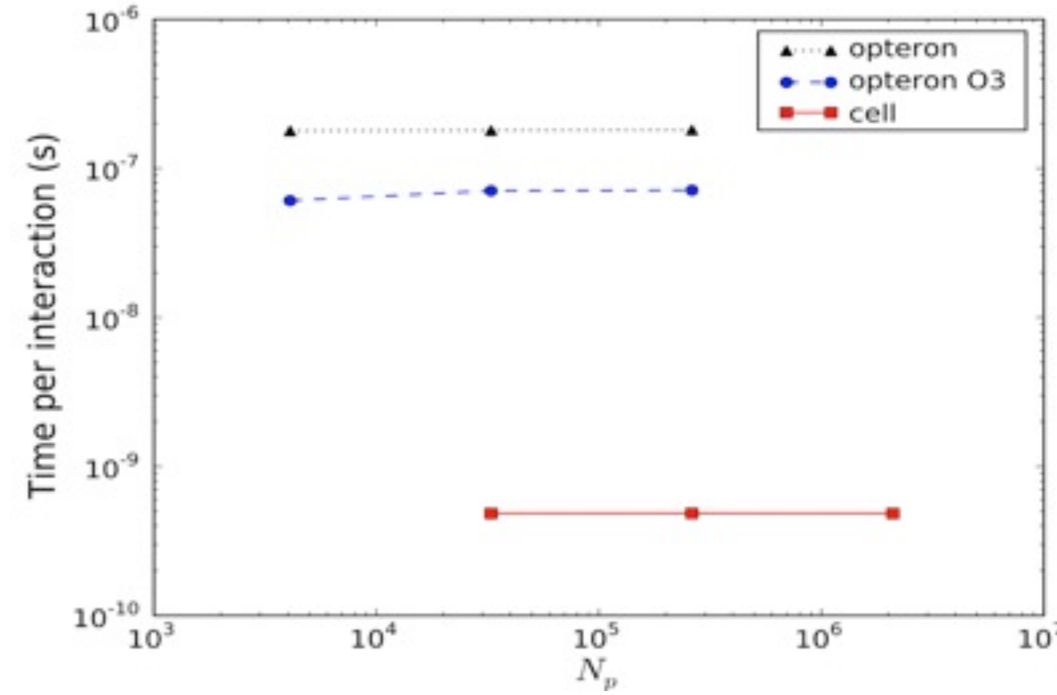
- Spectral smoothing of the CIC density field allows 6-th order Green function and 4th order super-Lanczos gradients for high-accuracy Poisson-solves
- Short-range force is fit to the numerical difference between Newtonian and long-range force (not conventional P^3M)
- Short-range force time-steps are sub-cycled within long-range force kicks via symplectic algorithm
- Short-range computations isolated as essentially 'on-node', replace or re-design for different architectures (e.g., BG/Q or GPU)



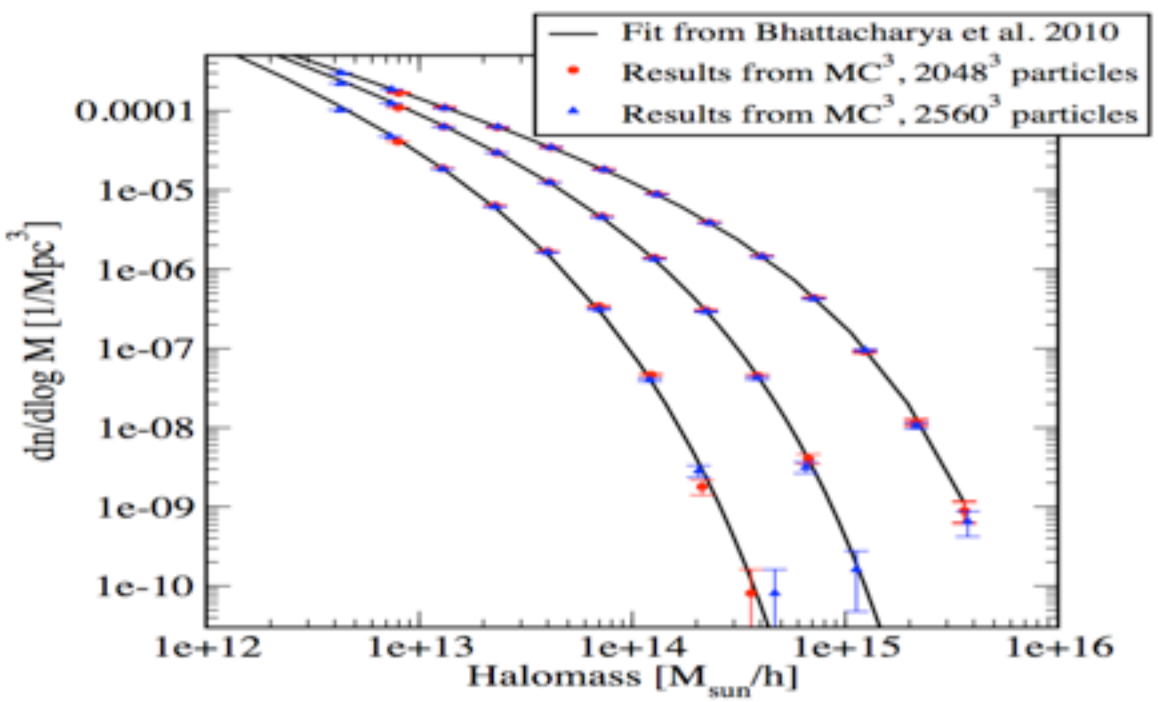
Some Results --



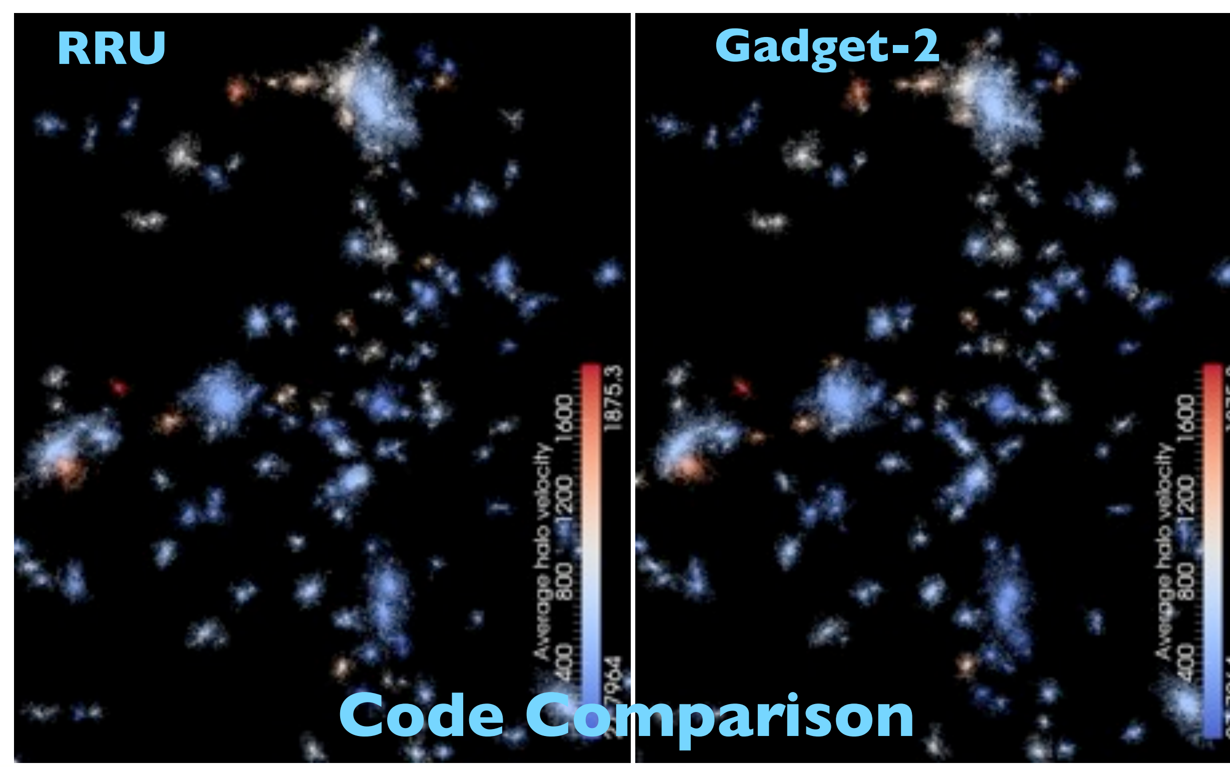
Roadrunner view (halos) of the Universe at z=2 from a 64 billion particle run (9 runs on one weekend)



Particle interactions: Cell computation gave improvement of two orders of magnitude over the Opterons for the short-range force

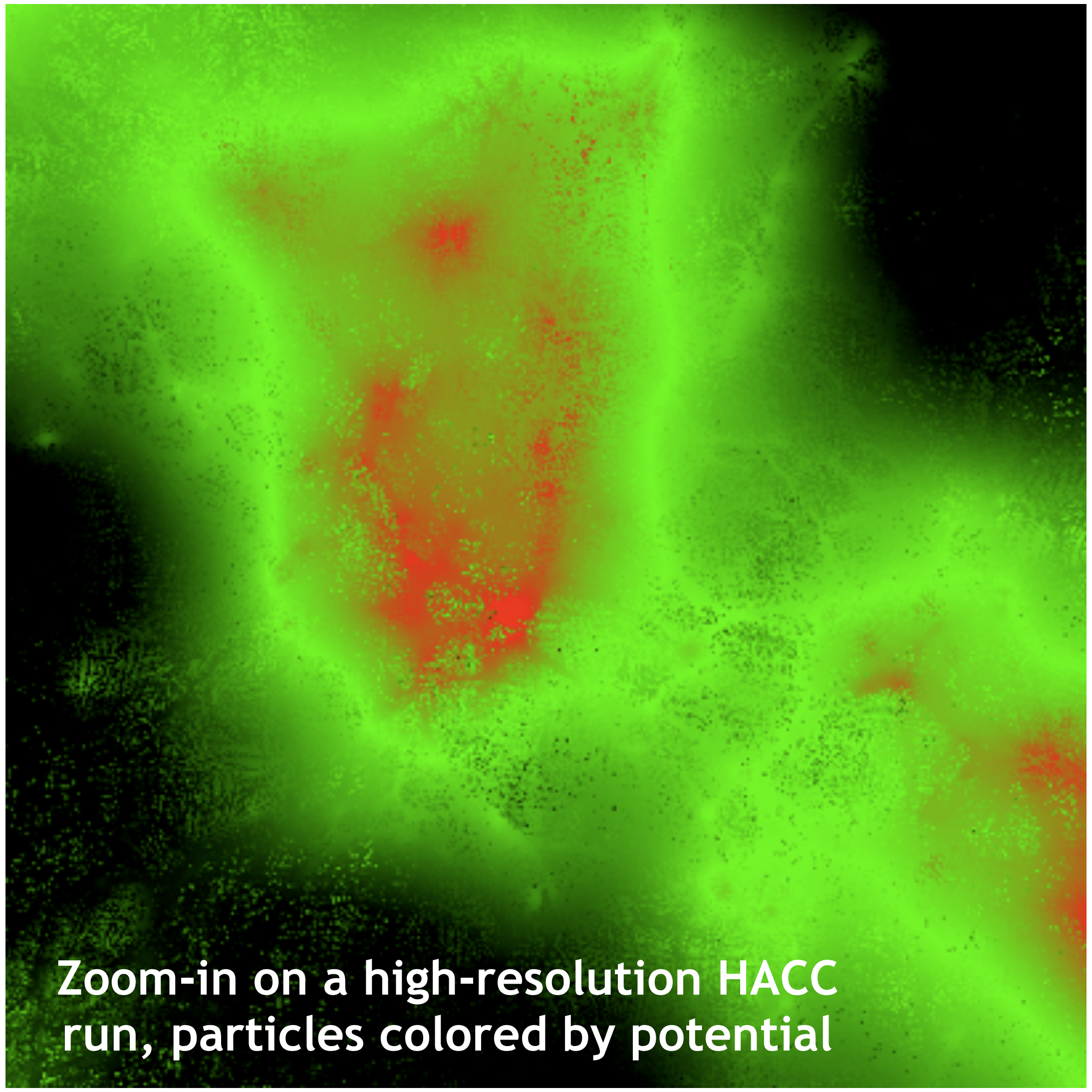


Halo mass function at z=0,1,2 and the corresponding phenomenological fit



HACC on Titan?

- **Simulation Requirements:** Need sets of very large runs (~trillion particle) for solving cosmological inverse problems; fast turn-around important
- **Titan hardware:** Need to understand host CPU/GPU better, but appears that overall ratios (memory, communication) are not too far from our previous HACC implementations (RR and CPU/GPU)
- **Software:** Unlikely to use directive-based approaches, need to decide on CUDA/OpenCL trade-offs
- **Prospects:** With some preliminary work could be ready to hit the ground running for Phase 4-6 (if I understand them correctly) --



Zoom-in on a high-resolution HACC run, particles colored by potential