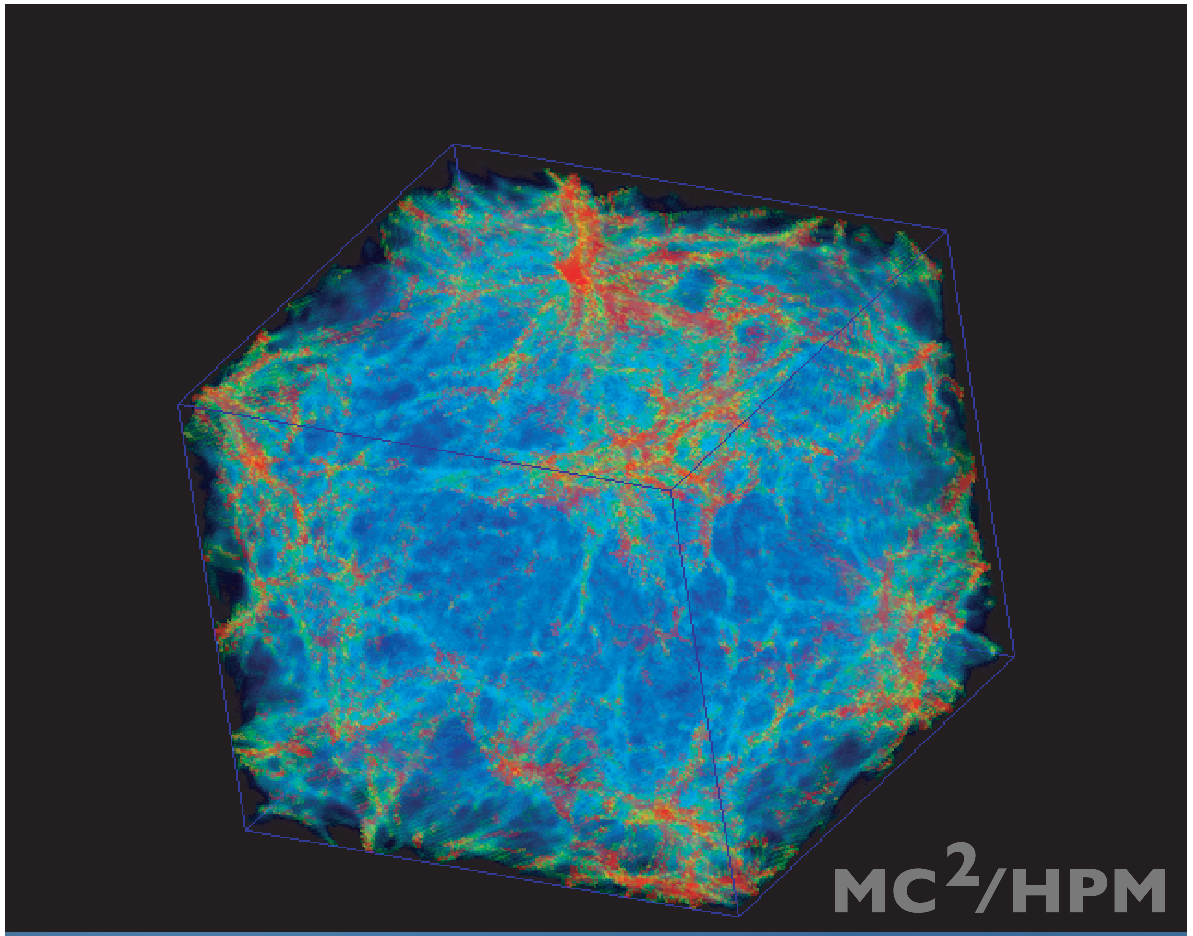
Computing the Universe

Adrian Pope
High Energy Physics Division
Argonne National Laboratory

ANL HEP: S. Habib, K. Heitmann, N. Frontiere (UofC)
ANL LCF/MCS: H. Finkel , V. Morozov, J. Insley, V. Vishwanath, T. Peterka
LANL: D. Daniel, P. Fasel, J. Woodring, J. Ahrens
LBNL/UC: Z. Lukic , M. White, J. Carlson

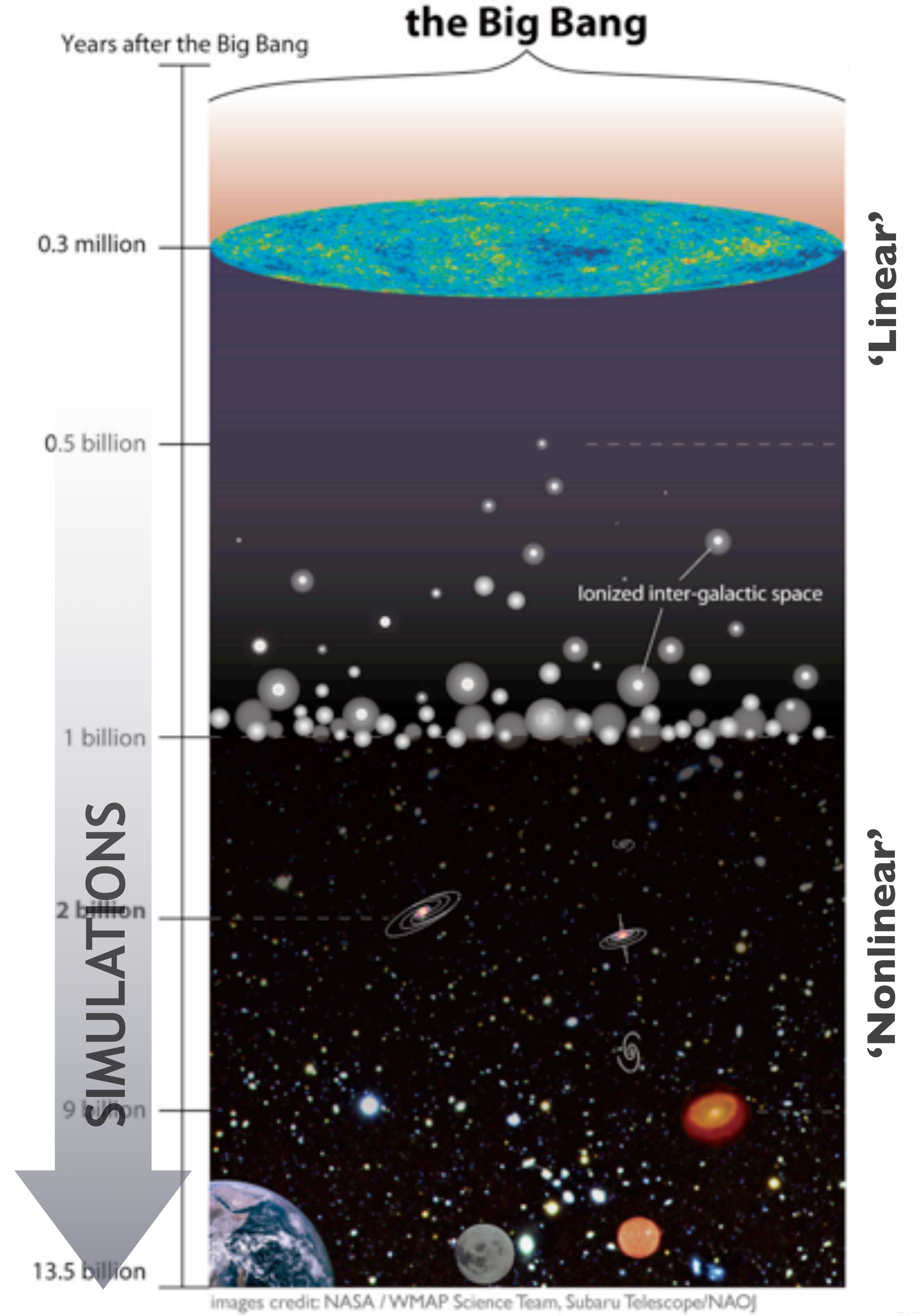
Computational Cosmology: A ‘Particle Physics’ Perspective

- **Primary Research Target:** Cosmological signatures of physics beyond the Standard Model
- **Structure Formation Probes:** Exploit nonlinear regime of structure formation
  - **Discovery Science:** Derive signatures of new physics, search for new cosmological probes
  - **Precision Predictions:** Aim to produce the best predictions and error estimates/distributions for structure formation probes
  - **Design and Analysis:** Advance ‘Science of Surveys’; contribute to major ‘Dark Universe’ missions: BOSS, DES, DESI, LSST...



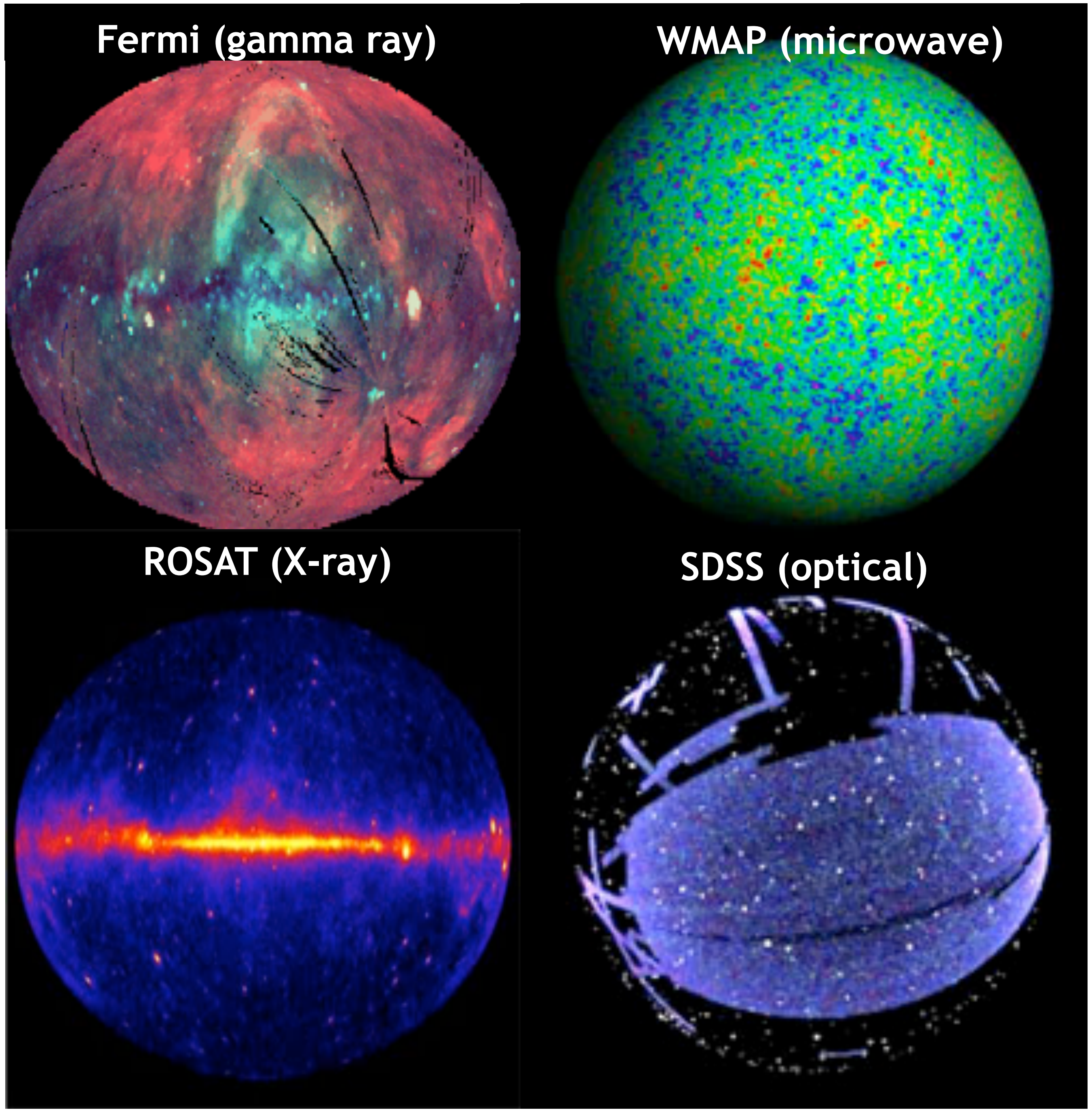
Structure Formation: The Basic Paradigm

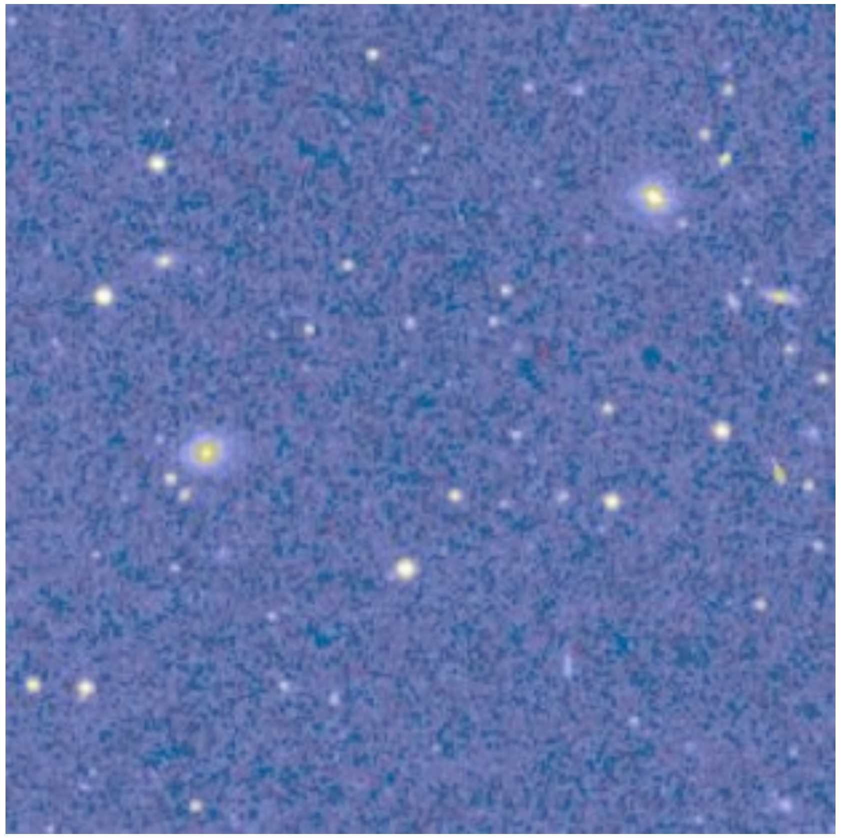
- **Solid understanding of structure formation; success underpins most cosmic discovery**
  - Initial conditions laid down by inflation
  - Initial perturbations amplified by gravitational instability in a dark matter-dominated Universe
  - Relevant theory is gravity, field theory, and atomic physics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



Cosmological Probes of Physics Beyond the Standard Model

- **Dark Energy:**
  - Properties of DE equation of state, modifications of GR, other models?
  - Sky surveys, terrestrial experiments
- **Dark Matter:**
  - Direct/Indirect searches, clustering properties, constraints on model parameters
  - Sky surveys, targeted observations, terrestrial experiments
- **Inflation:**
  - Probing primordial fluctuations, CMB polarization, non-Gaussianity
  - Sky surveys
- **Neutrino Sector:**
  - CMB, linear and nonlinear matter clustering
  - Sky surveys, terrestrial experiments

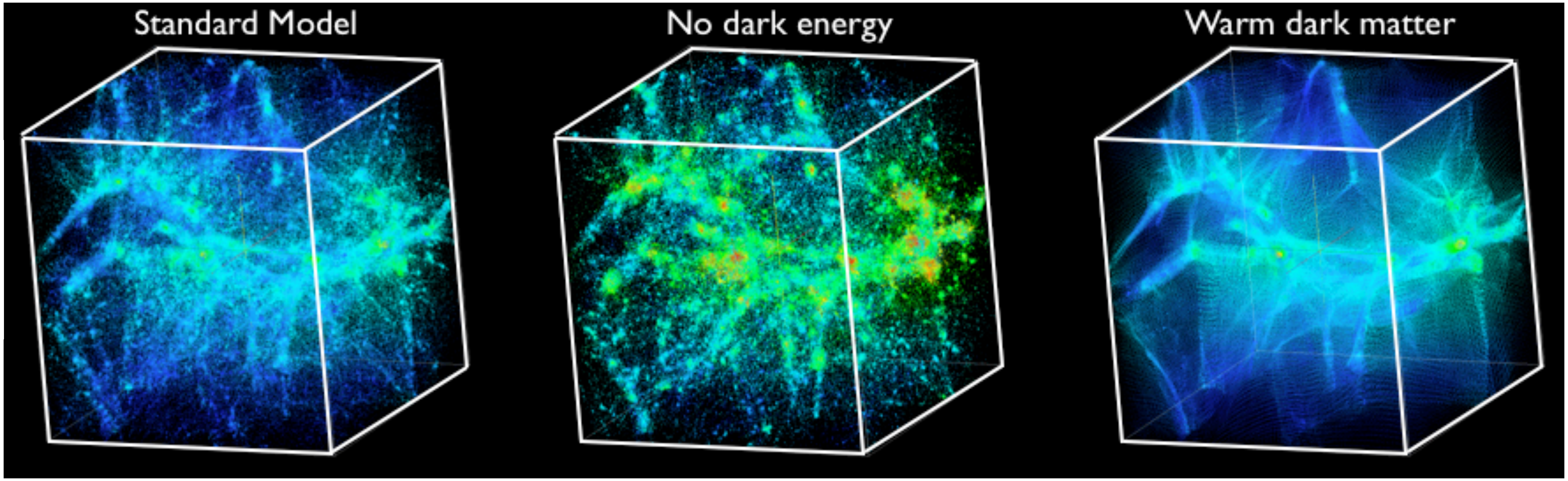




Digitized Sky Survey
1950s-1990s

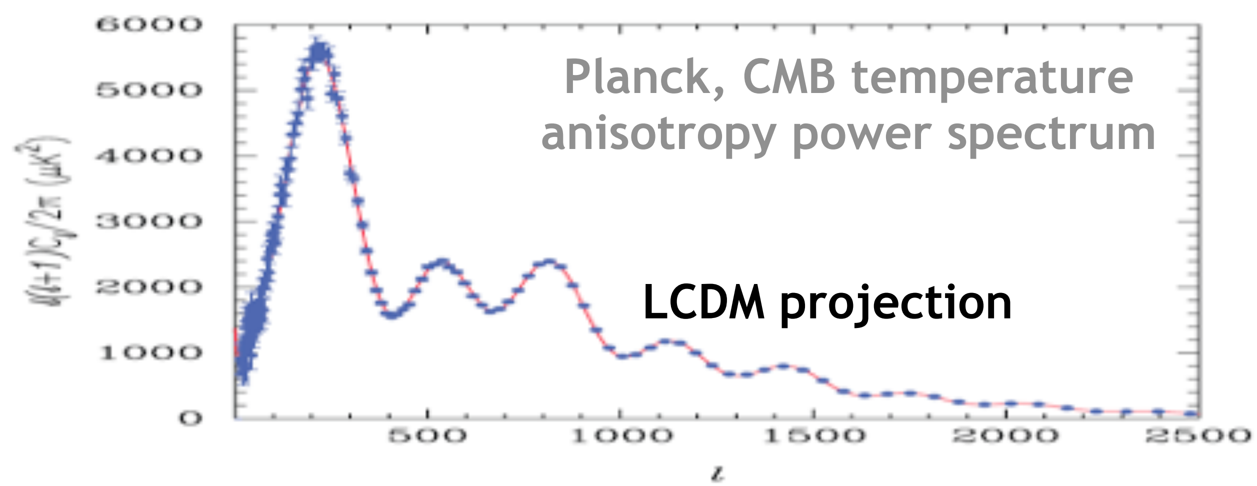
Sloan Digital Sky Survey
2000-2008

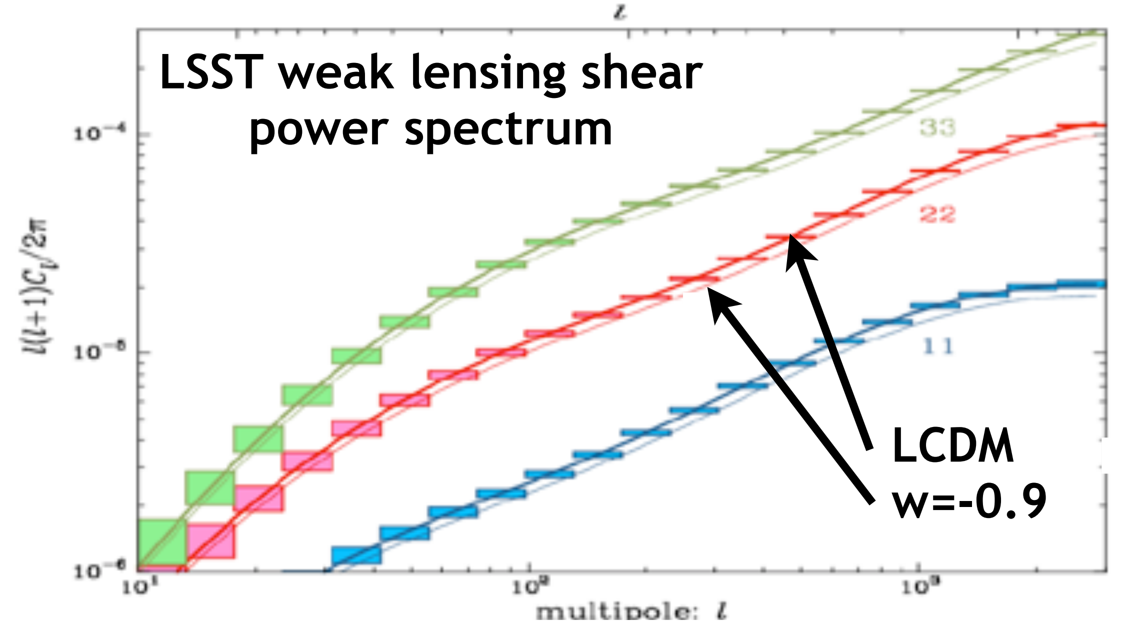
Large Synoptic Survey Telescope
2020-2030
(Deep Lens Survey image)

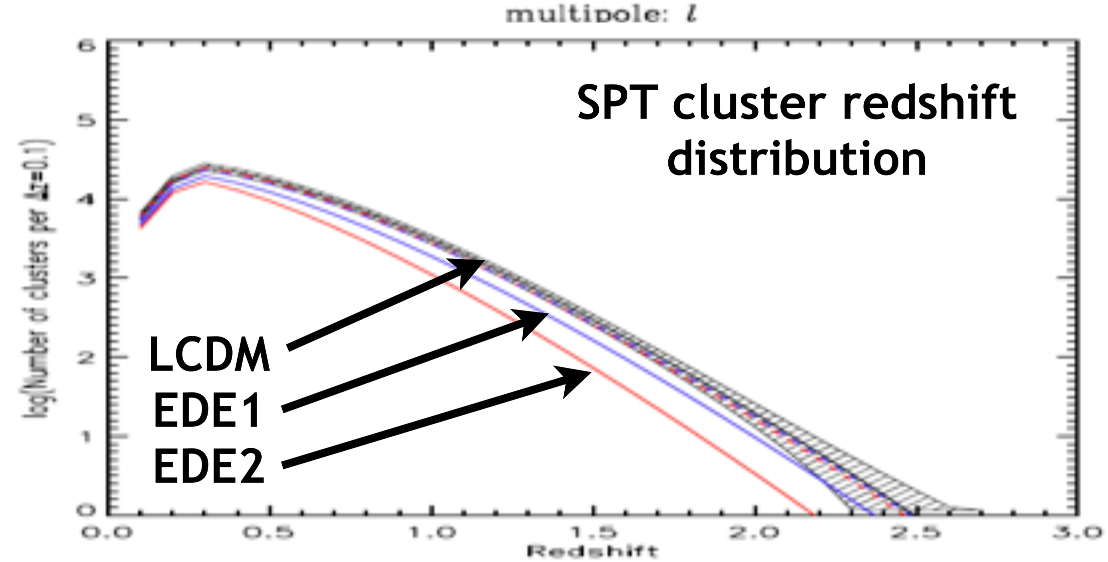


Precision Cosmology: “Inverting” the 3-D Sky

- **Cosmic Inverse Problem**:
  - From sky maps to scientific inference
- **Cosmological Probes**:
  - Measure geometry and presence/growth of structure (linear and nonlinear)
- **Examples**:
  - Baryon Acoustic Oscillations (BAO), cluster counts, CMB, weak lensing, galaxy clustering...
- **Cosmological Standard Model**:
  - Verified at 5-10% with multiple observations
- **Future Targets**:
  - Aim to control survey measurements to ~1%
- **The Challenge**:
  - Theory and simulation must satisfy stringent criteria for inverse problems and precision cosmology not to be theory-limited!



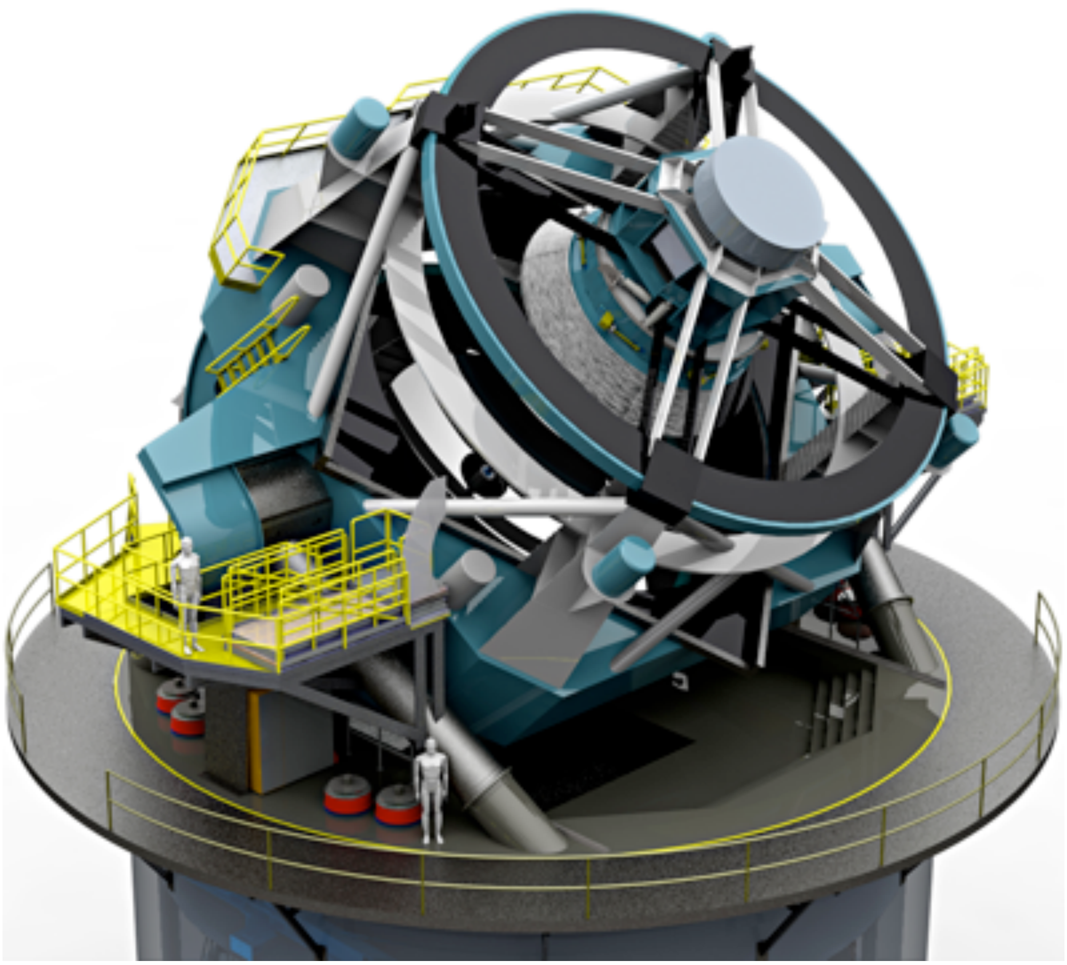




Alam, Lukic, & Bhattacharya 2011

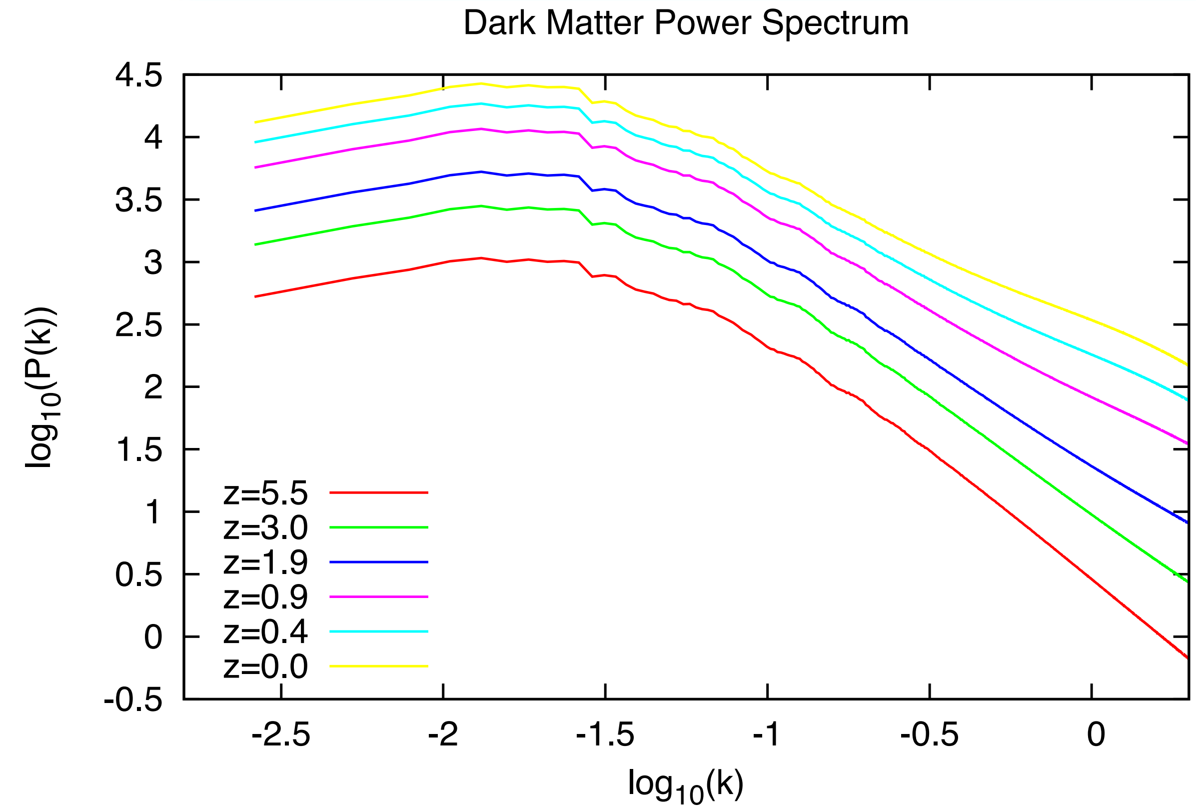
Computing the Universe: Simulations for Surveys

- **Survey Support**: Many uses for simulations
  - Mock catalogs, covariance, emulators, etc.
- **Simulation Volume**: Large (volume, sky-fraction) surveys, weak signals
  - ~ $(3\ \mathrm{Gpc})^3$, memory required ~100 TB -- 1 PB
- **Number of Particles**: Mass resolutions depend on objects to be resolved
  - ~10^8 -- 10^{10} solar masses requires N ~ 10^{11} -- 10^{12}
- **Force Resolution**: ~kpc resolution
  - (Global) spatial dynamic range of 10^6
- **Throughput**:
  - Large numbers of simulations required (100 --1000),
  - Development of analysis suites, and emulators
  - Petascale-exascale computing
- **Computationally very challenging!**

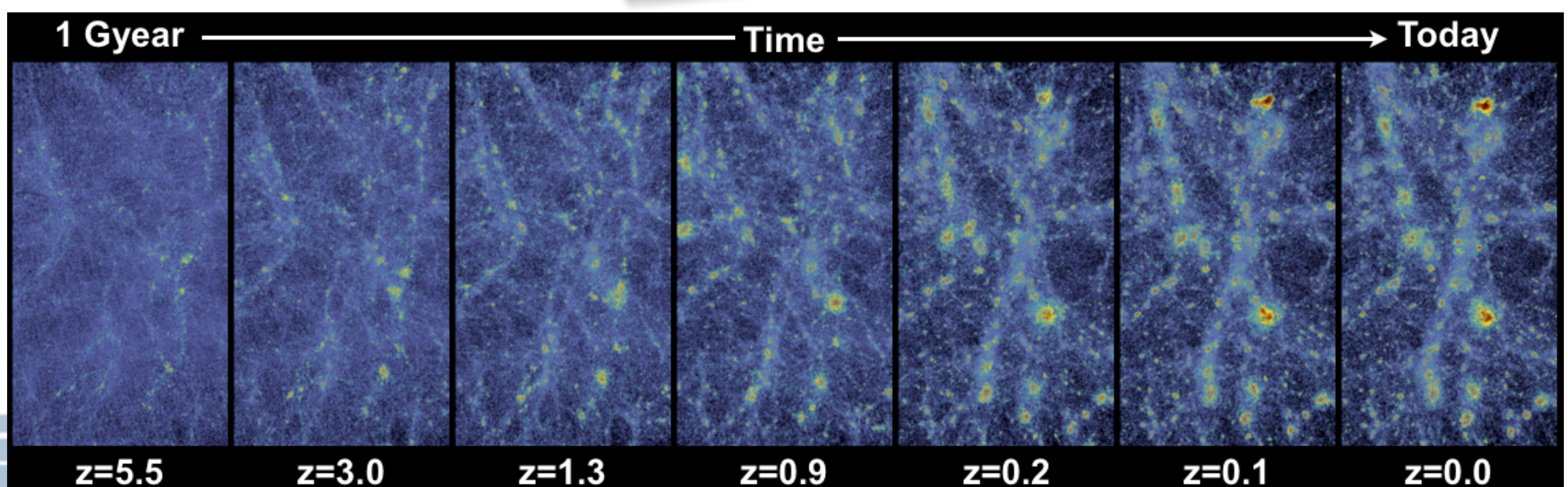


Simulating the Universe

- Gravity dominates at large scales
  - Vlasov-Poisson equation (VPE)
- VPE is 6D, cannot be solved as a PDE
- N-body methods for gravity
  - No shielding
  - Naturally Lagrangian
- Additional small-scale physics
  - Gas, feedback, etc.
  - Sub-grid modeling eventually
  - **HACC** is **gravity only** (for now)



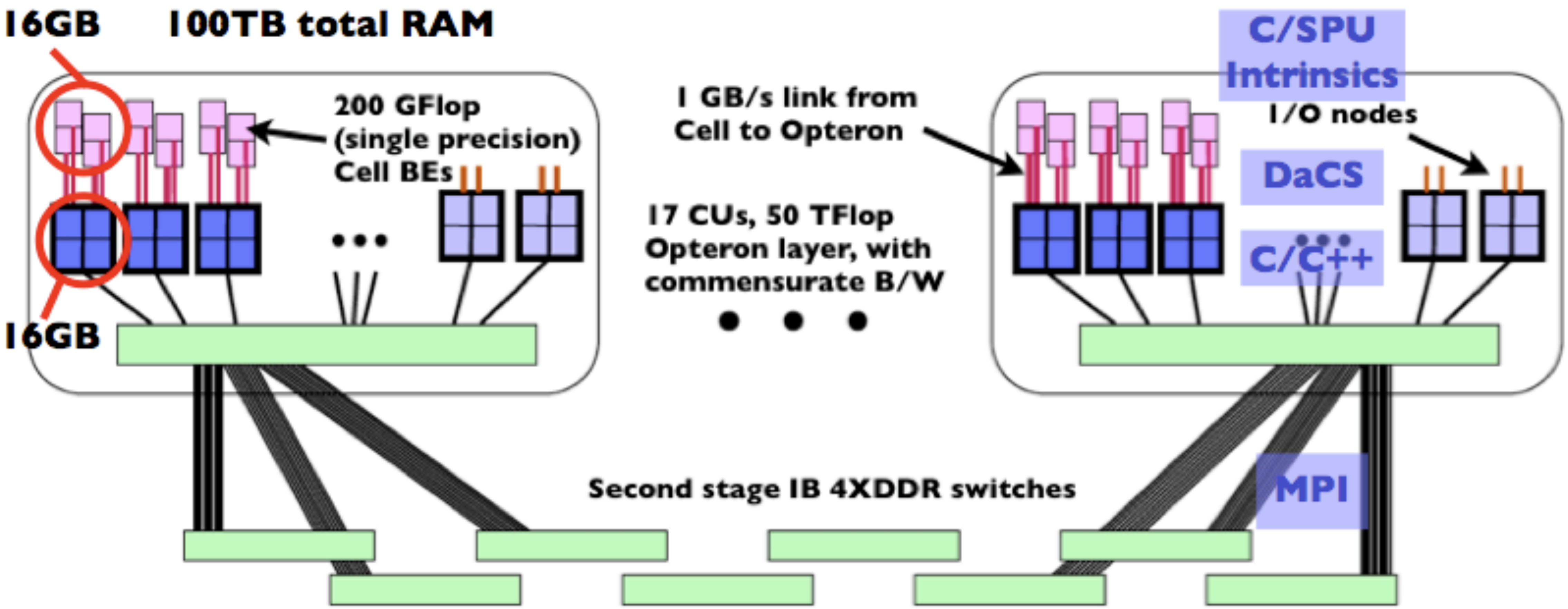
$$
\begin{aligned}
\frac{\partial f_i}{\partial t} + \dot{\mathbf{x}}\frac{\partial f_i}{\partial \mathbf{x}} - \nabla\phi\frac{\partial f_i}{\partial \mathbf{p}} = 0, \qquad \mathbf{p} = a^2\dot{\mathbf{x}},\\
\nabla^2\phi &= 4\pi G a^2(\rho(\mathbf{x},t) - \langle\rho_{\rm dm}(t)\rangle) = 4\pi G a^2\Omega_{\rm dm}\delta_{\rm dm}\rho_{\rm cr},\\
\delta_{\rm dm}(\mathbf{x},t) &= (\rho_{\rm dm} - \langle\rho_{\rm dm}\rangle)/\langle\rho_{\rm dm}\rangle),\\
\rho_{\rm dm}(\mathbf{x},t) &= a^{-3}\sum_i m_i \int d^3\mathbf{p} f_i(\mathbf{x},\dot{\mathbf{x}},t).
\end{aligned}
$$



How It All Started: Roadrunner (LANL)

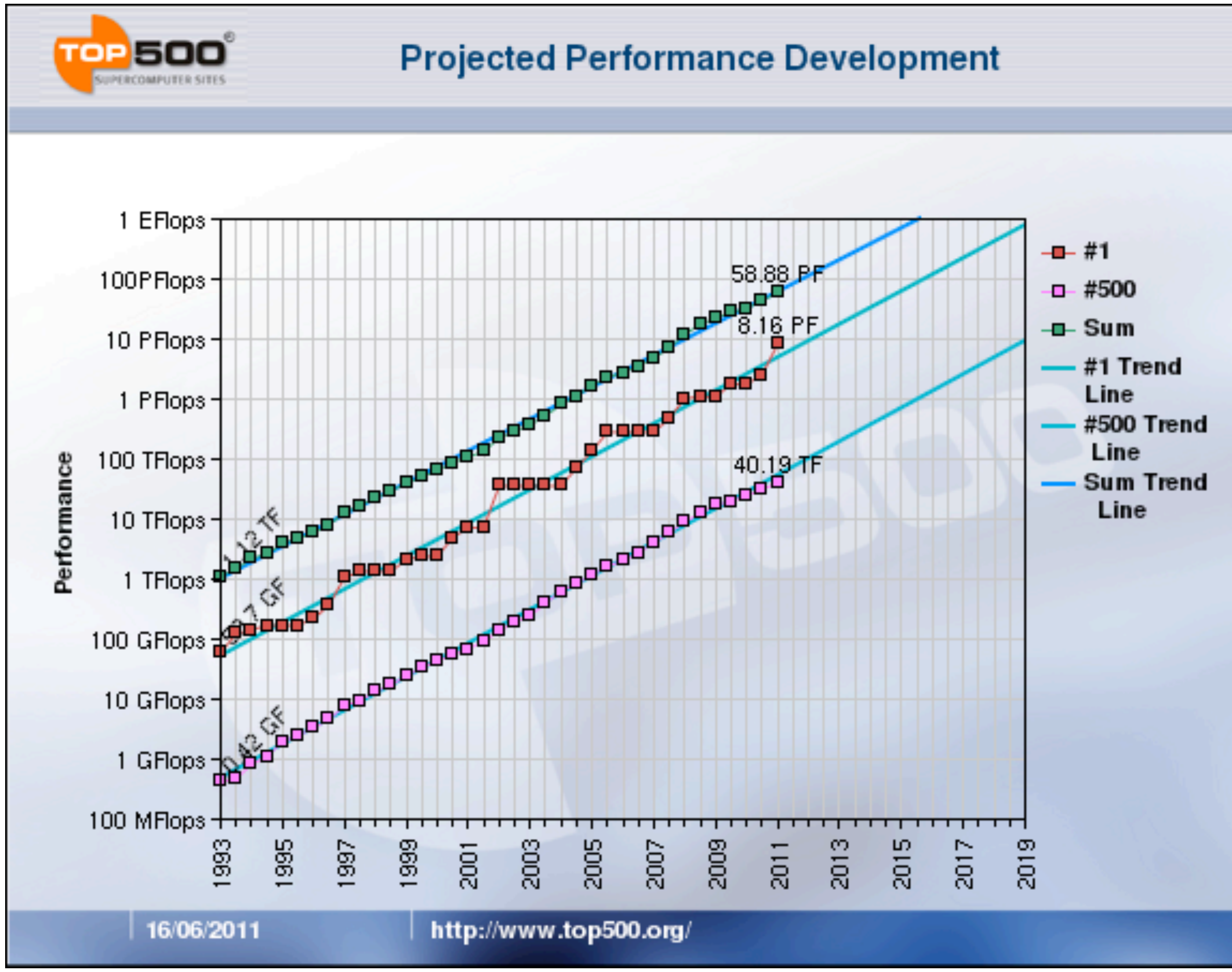
Andrew White Dec 7, 2007 + What if you had a petaflop/s

But what if it looked like this?



High Performance Computing

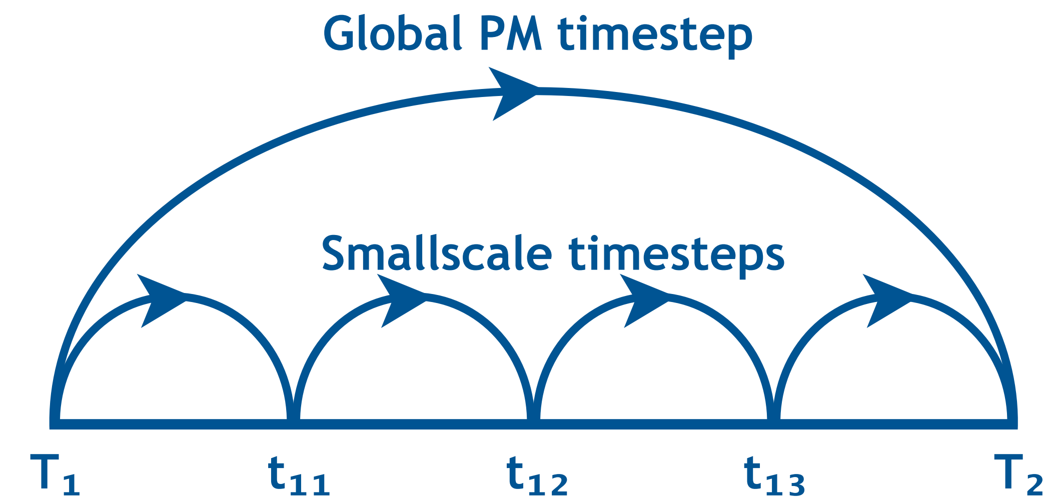
- **Supercomputers**: faster = more "parallel"
  - More nodes
    - Distributed memory parallel (eg. MPI)
    - Network communication, somewhat standard
    - Weak scaling (memory limited)
  - More cores per node
    - Shared memory parallel, "threading" (eg. OpenMP)
    - Many possible models
    - Strong scaling (use local compute)
  - "Memory hierarchy"
    - Balance computational speed, memory movement
- **Architecture**:
  - How to divide real estate (power) on chip
  - Heterogeneity
    - Hybrid chips (complicated)
    - Accelerators (PCI bottleneck)
    - Multiple programming styles



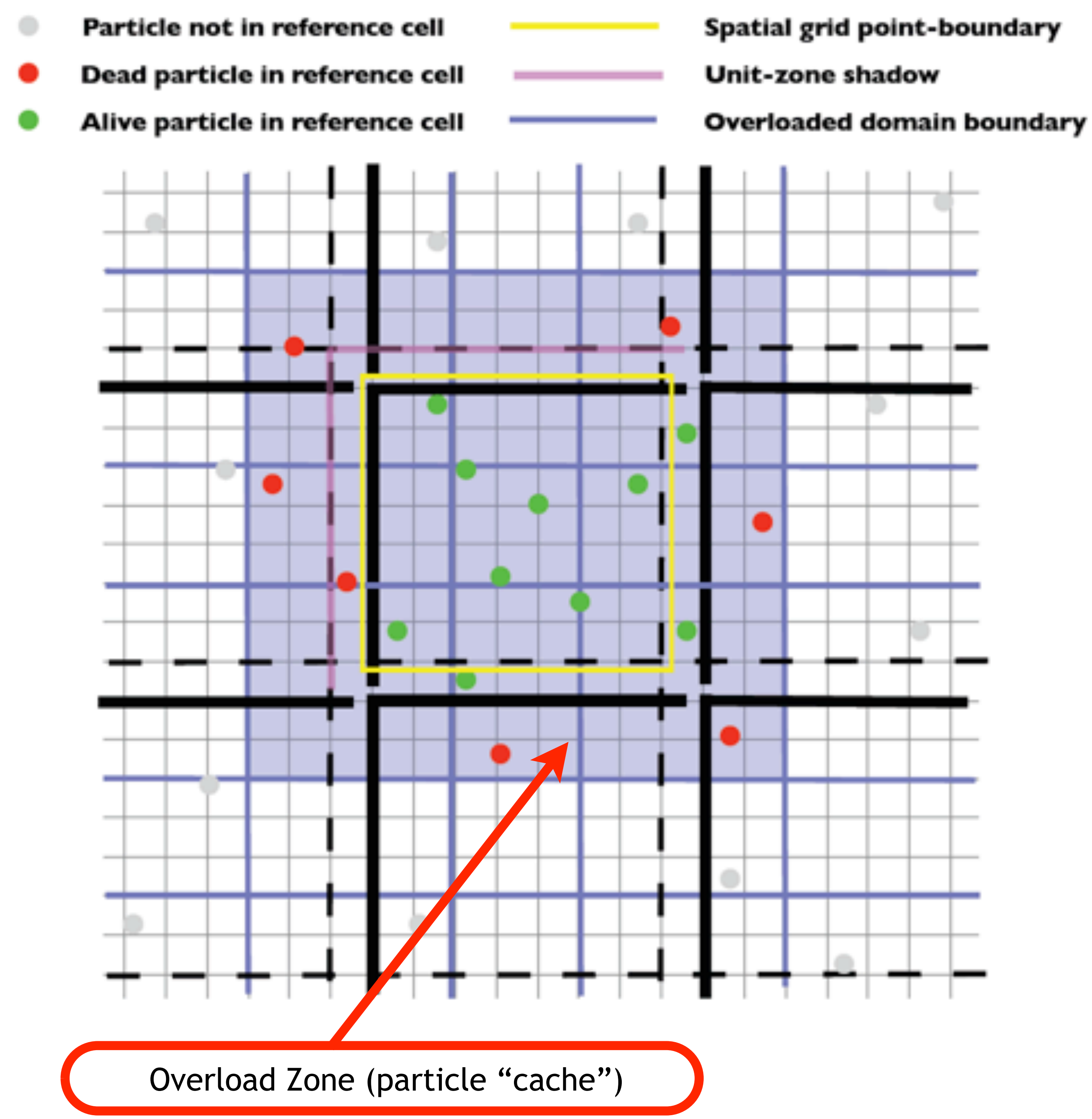
HACC (Hybrid/Hardware Accelerated Cosmology Code)

- **Large volume, high throughput** (weak lensing, large-scale structure, surveys)
  - Dynamic range: volume for long wavelength modes, resolution for halos/galaxy locations
  - Repeat runs: vary initial conditions (realizations), sample parameter space
  - Error control: 1% results
  - Low memory footprint: more particles = better mass resolution
  - Scaling: current and future computers (many MPI ranks, even more cores)
- **Flexibility**
  - Supercomputer architecture (CPU, Cell, GPGPU, Blue Gene)
  - Compute intensive code takes advantage of hardware
  - Bulk of code easily portable (MPI)
- **Development/maintenance**
  - (Relatively) few developer FTEs
  - Simpler code easier to develop, maintain, and port to different architectures
- **On-the-fly analysis, data reduction**
  - Reduce size/number of outputs, ease file system stress

Force Splitting

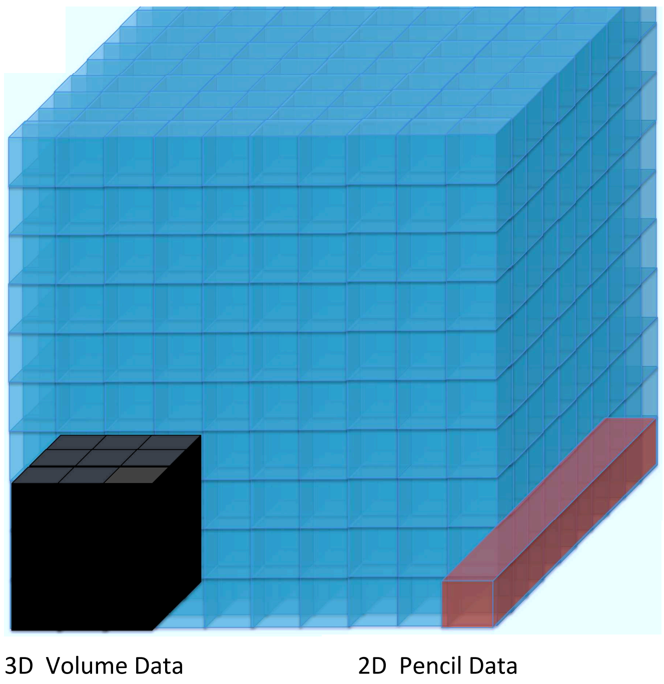


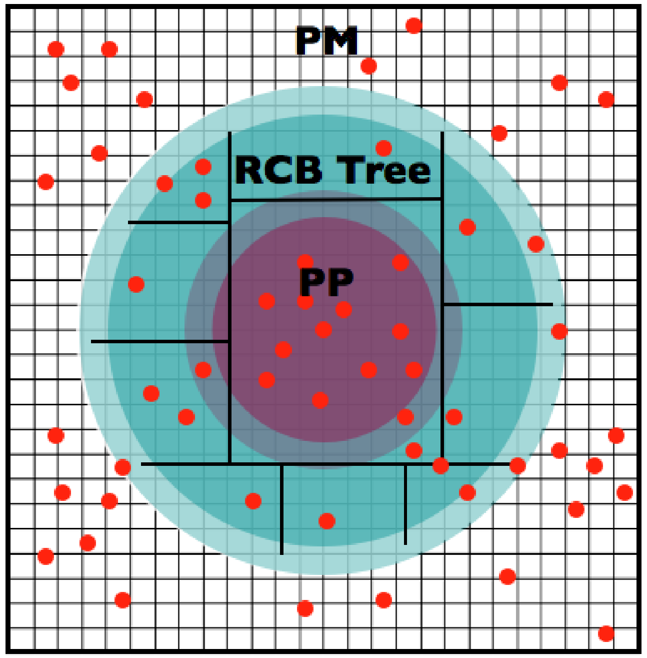
- Gravity is infinite range with no shielding
  - Every particle vs. every other particle
  - Split all-to-all comparison by separation length
- Long-range: Particle-Mesh (PM)
  - Distributed memory, MPI grid/FFT methods
  - ~10^4 dynamic range, slowly varying
  - Portable
- Short-range:
  - Shared memory, particle methods
  - ~10^2 dynamic range, quickly varying
  - Particle "cache" in overload zone
    - No additional MPI code
  - Modular
- Symplectic Integrator:
  - Standard operator splitting
  - "Subcycle" short-range steps

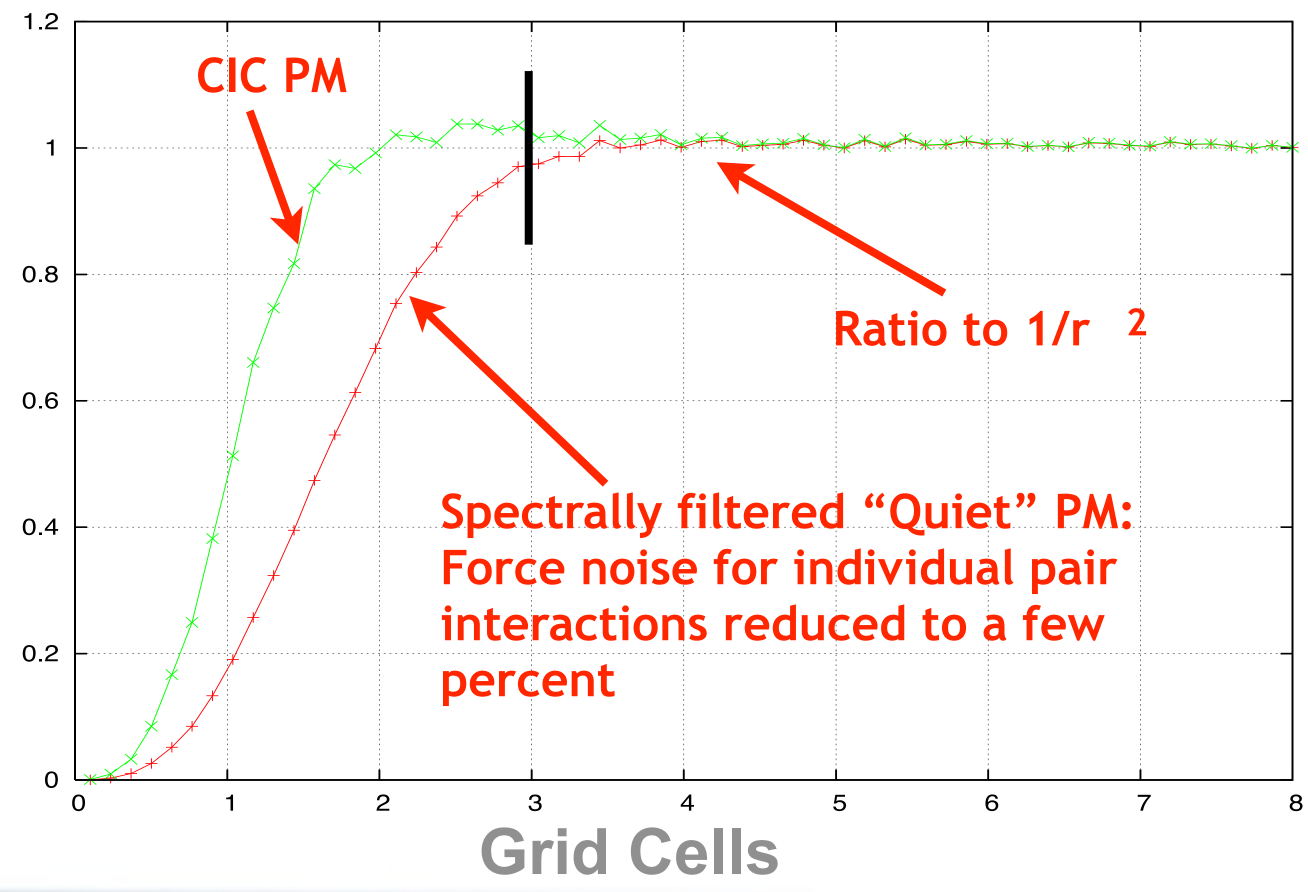


Force Handover

- **Spectral control of force hand-over**
  - Cloud-in-Cell grid deposition
    - Simple, local, noisy, anisotropic
  - Spectral manipulation of grid force
    - "Quiet" PM, cancellation of low-order error terms
  - Empirical fit for real-space short-range force
    - Average Quiet PM over many configurations
- **Modular short-range force solver**
  - **P³M**: direct particle-particle comparisons
    - Only for floating-point intense hardware
    - Small handover scale limits N^2 comparisons
  - **TreePM**: low order multipole approximation
    - More complex data-structures and control flow
    - Tree "local" to MPI rank

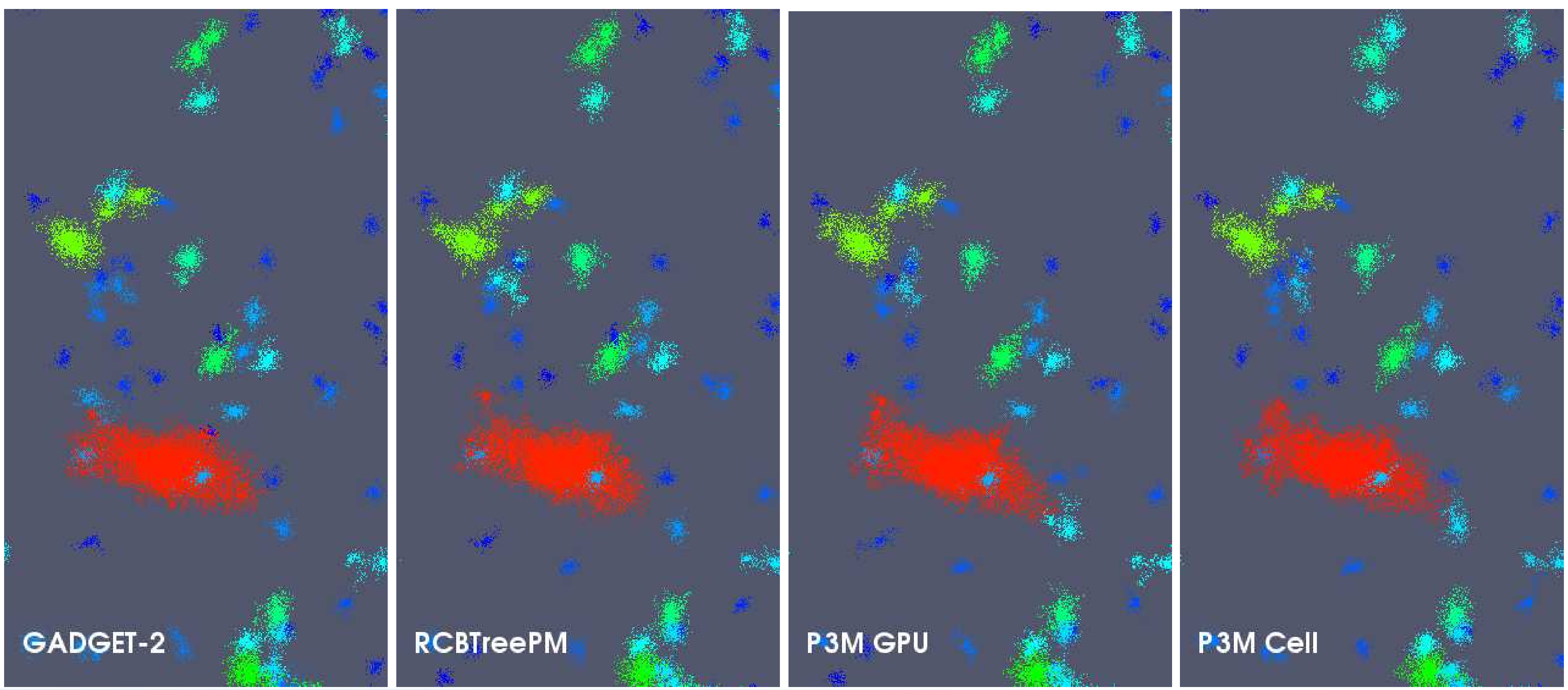






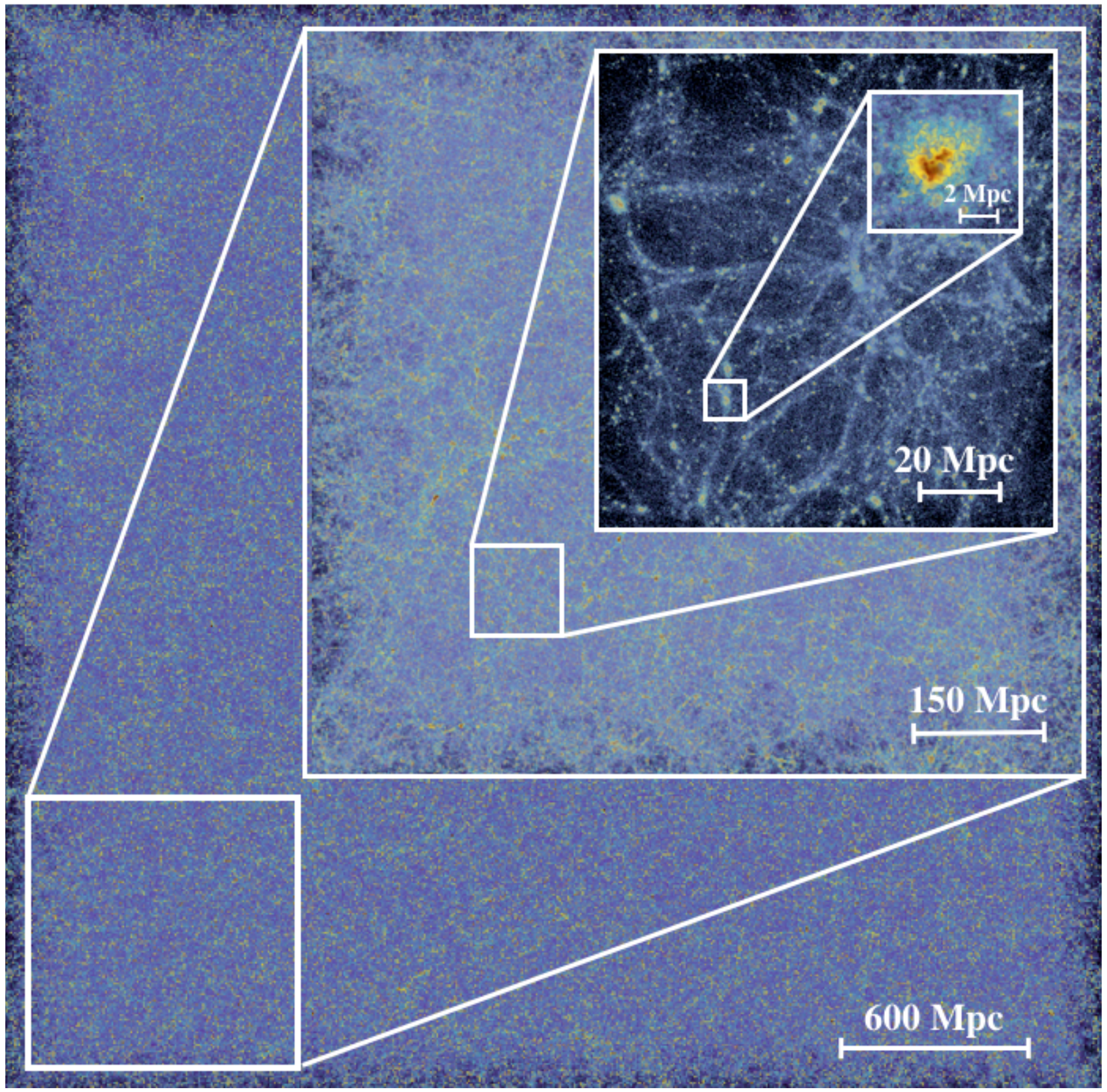
Architectures and Algorithms

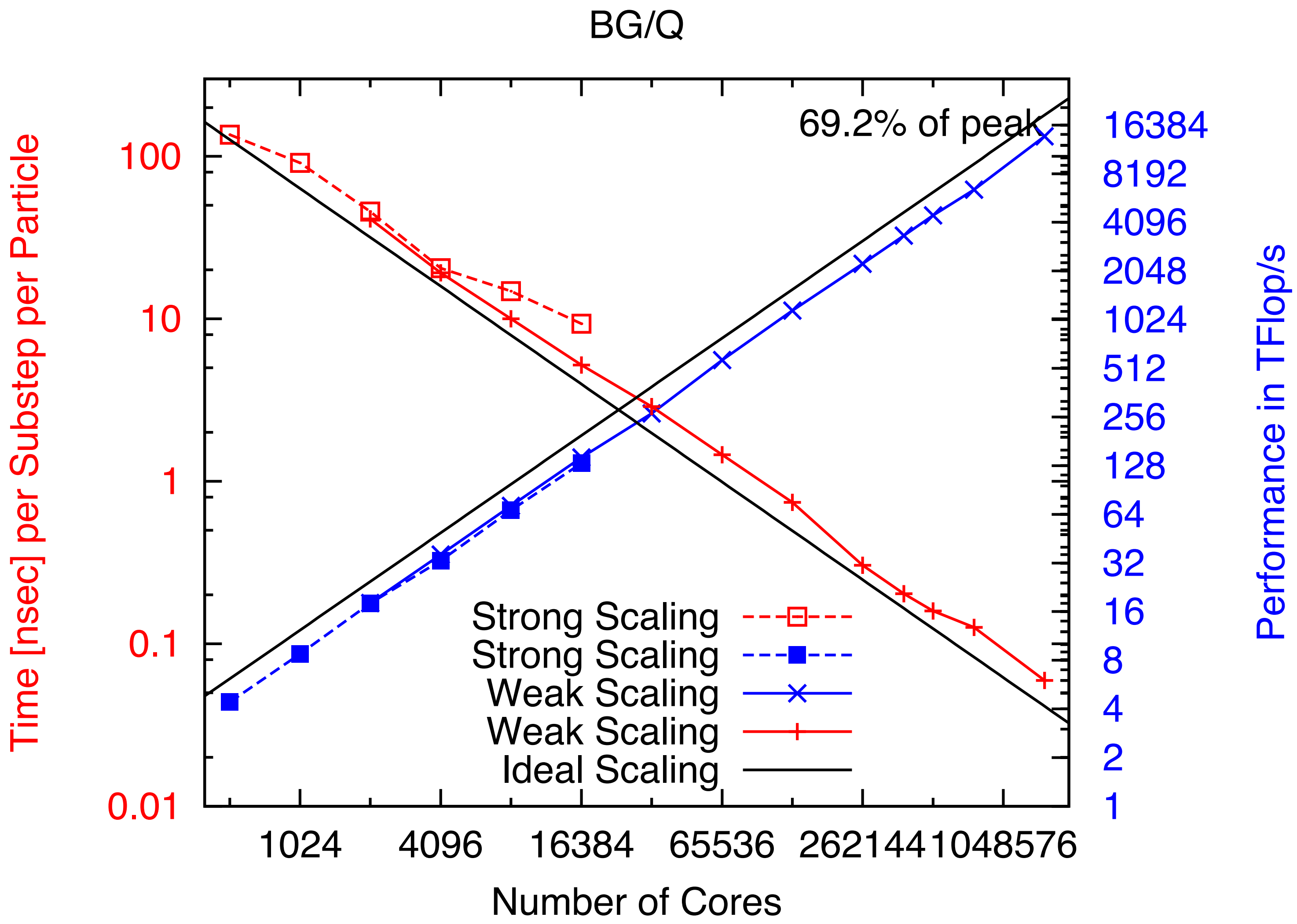
- **IBM Cell Broadband Engine Accelerator:** LANL/Roadrunner (2008)
  - P^3M, MPI + Cell SDK
- **IBM Blue Gene/Q:** ANL/Mira, LLNL/Sequoia (2012)
  - PPTreePM, MPI + OpenMP + IBM QPX (BG/Q intrinsics)
- **GPGPU:** ORNL/Titan (2012)
  - P^3M, MPI + OpenMP + OpenCL



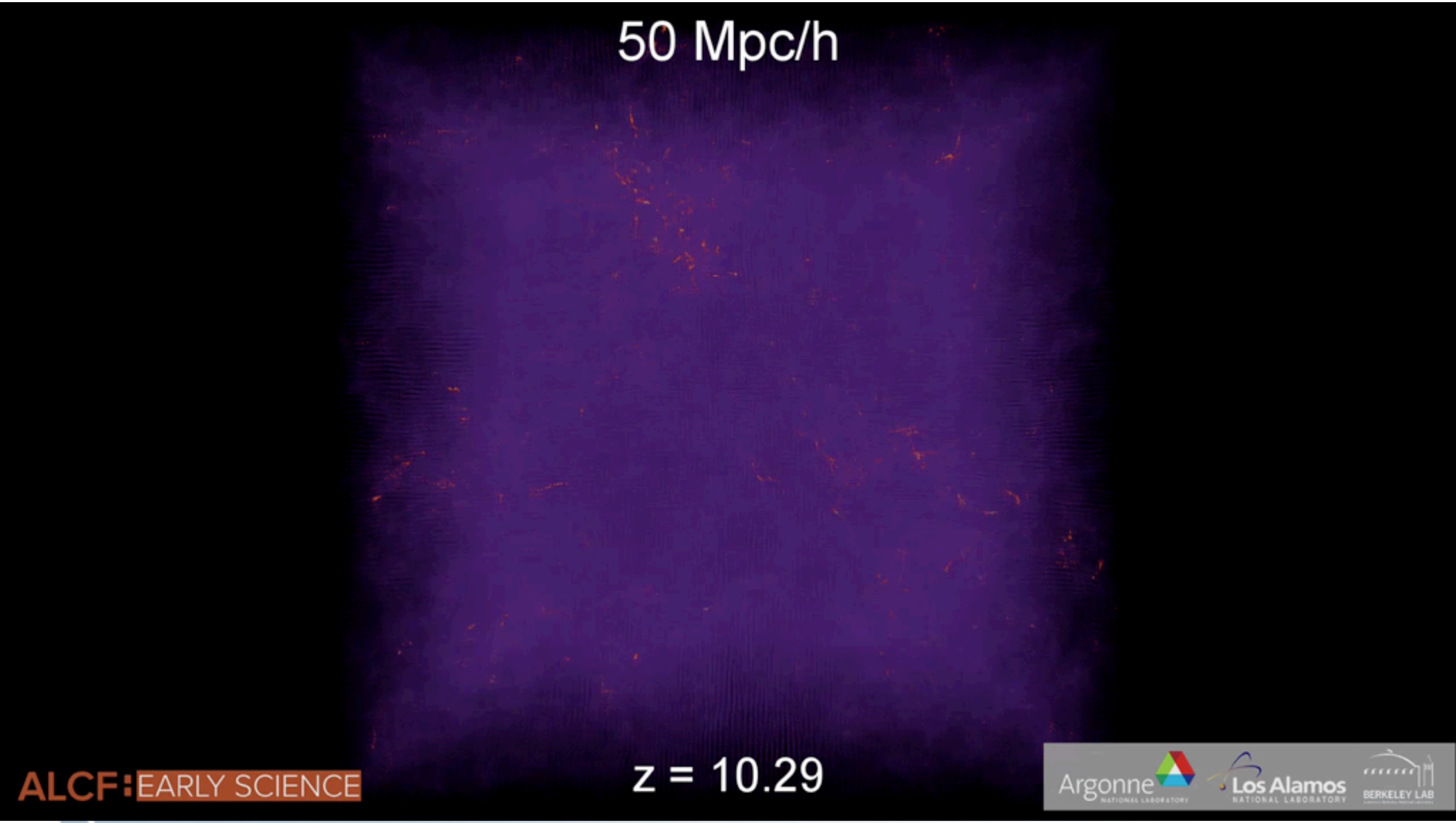
Outer Rim Simulation Run

- ANL/Mira (BG/Q), 3 Gpc/h box, 1.1 trillion particles!

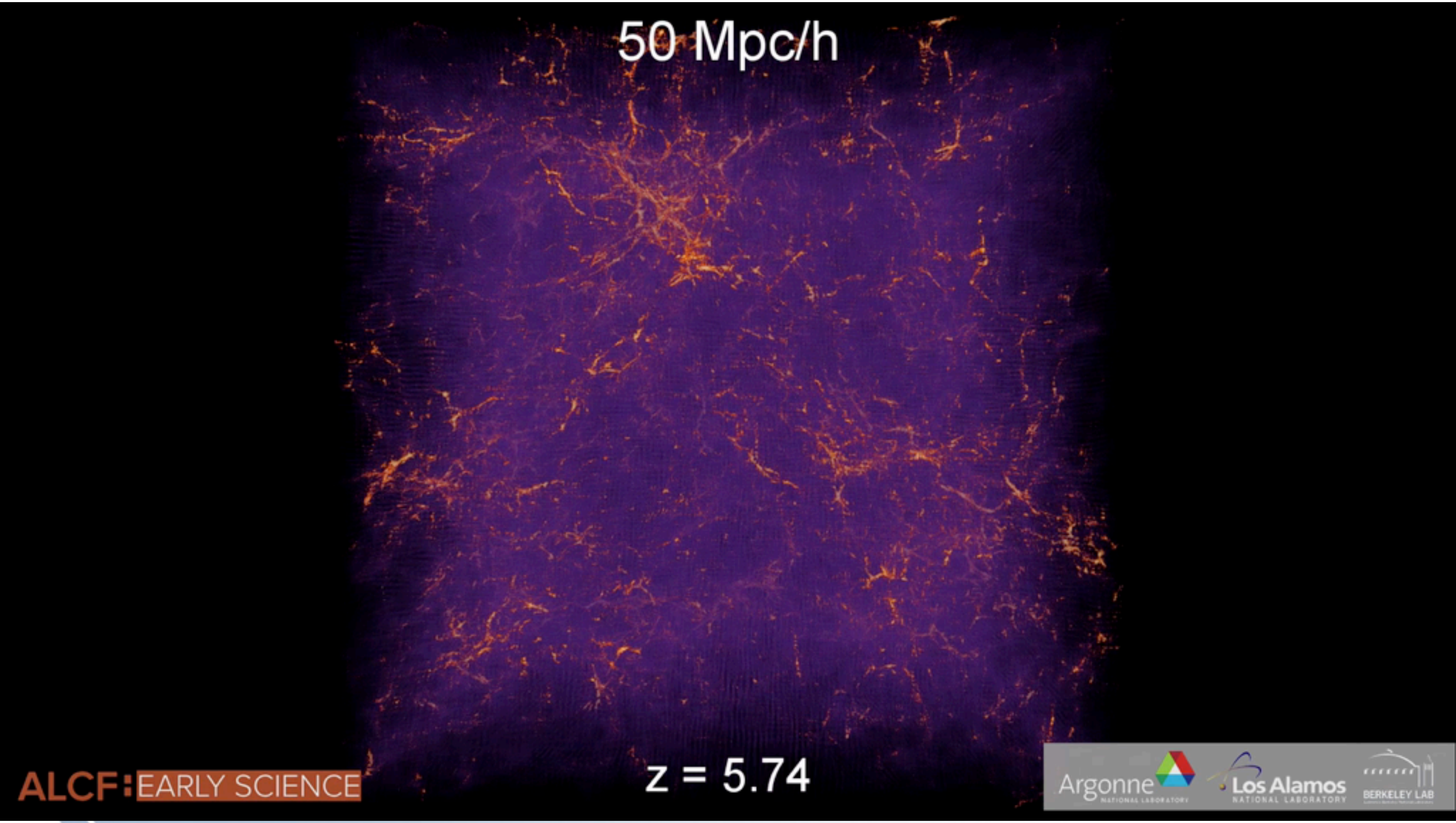




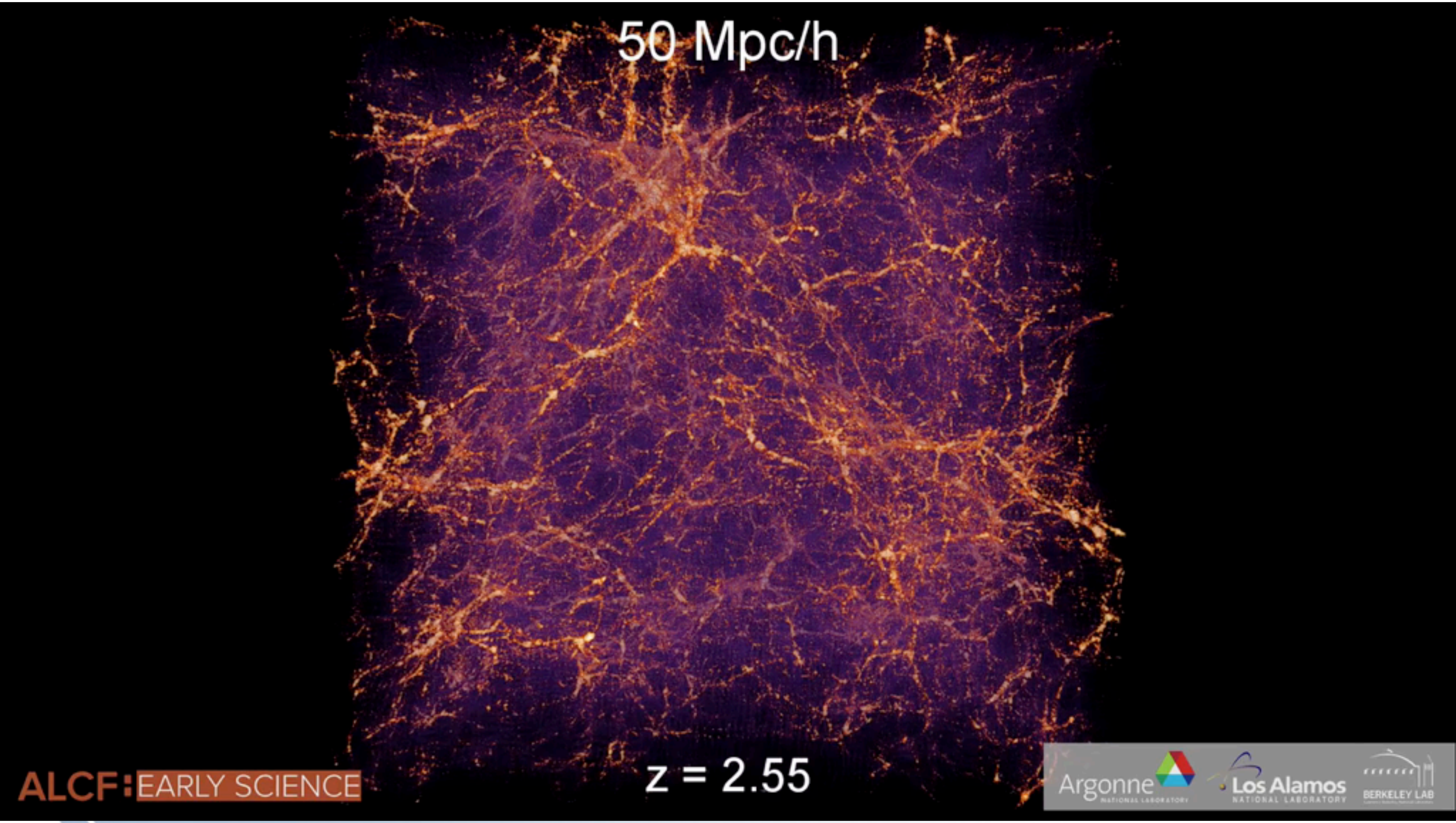
Movie Captures: Growth of Structure



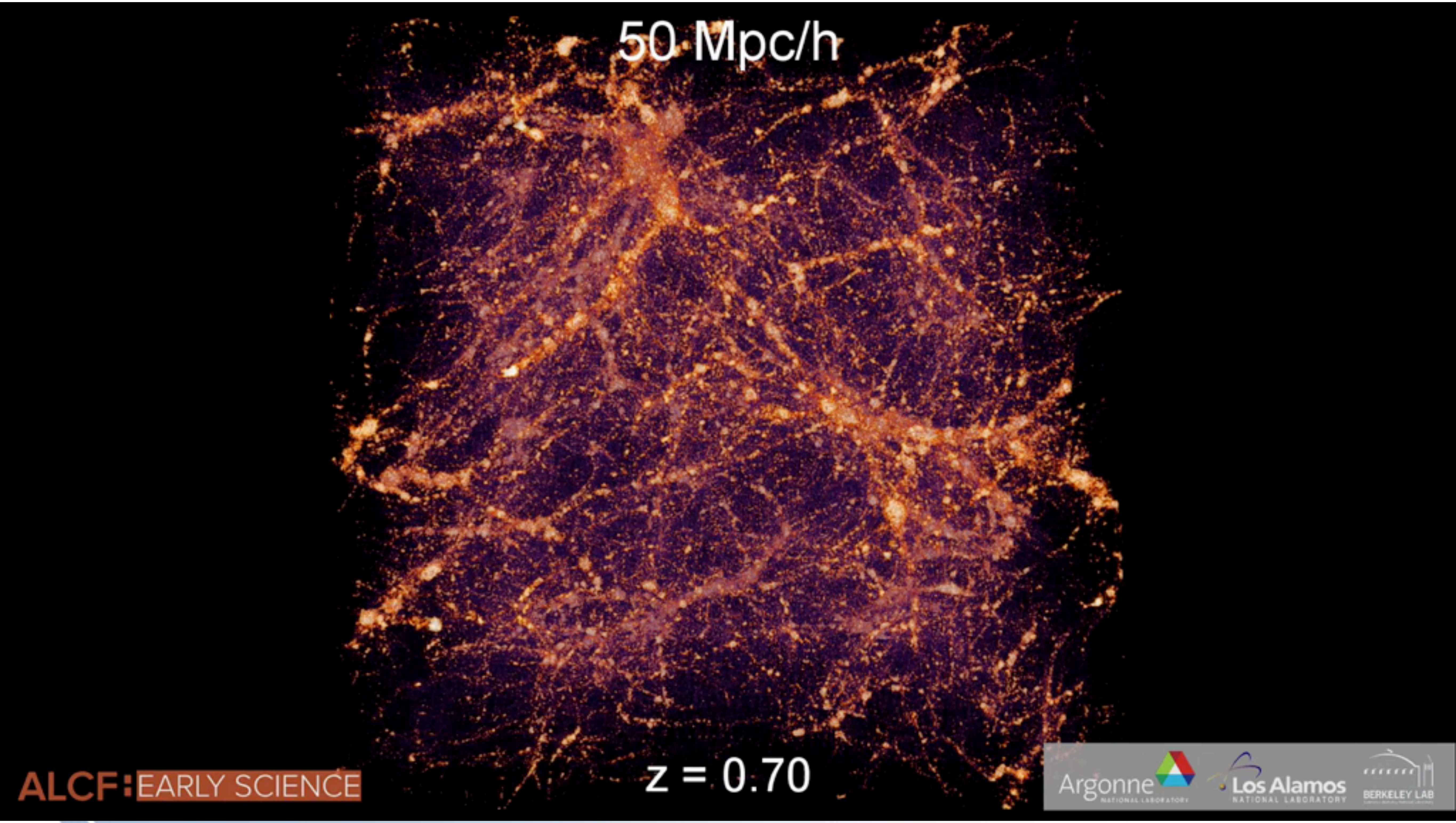
Movie Captures: Growth of Structure



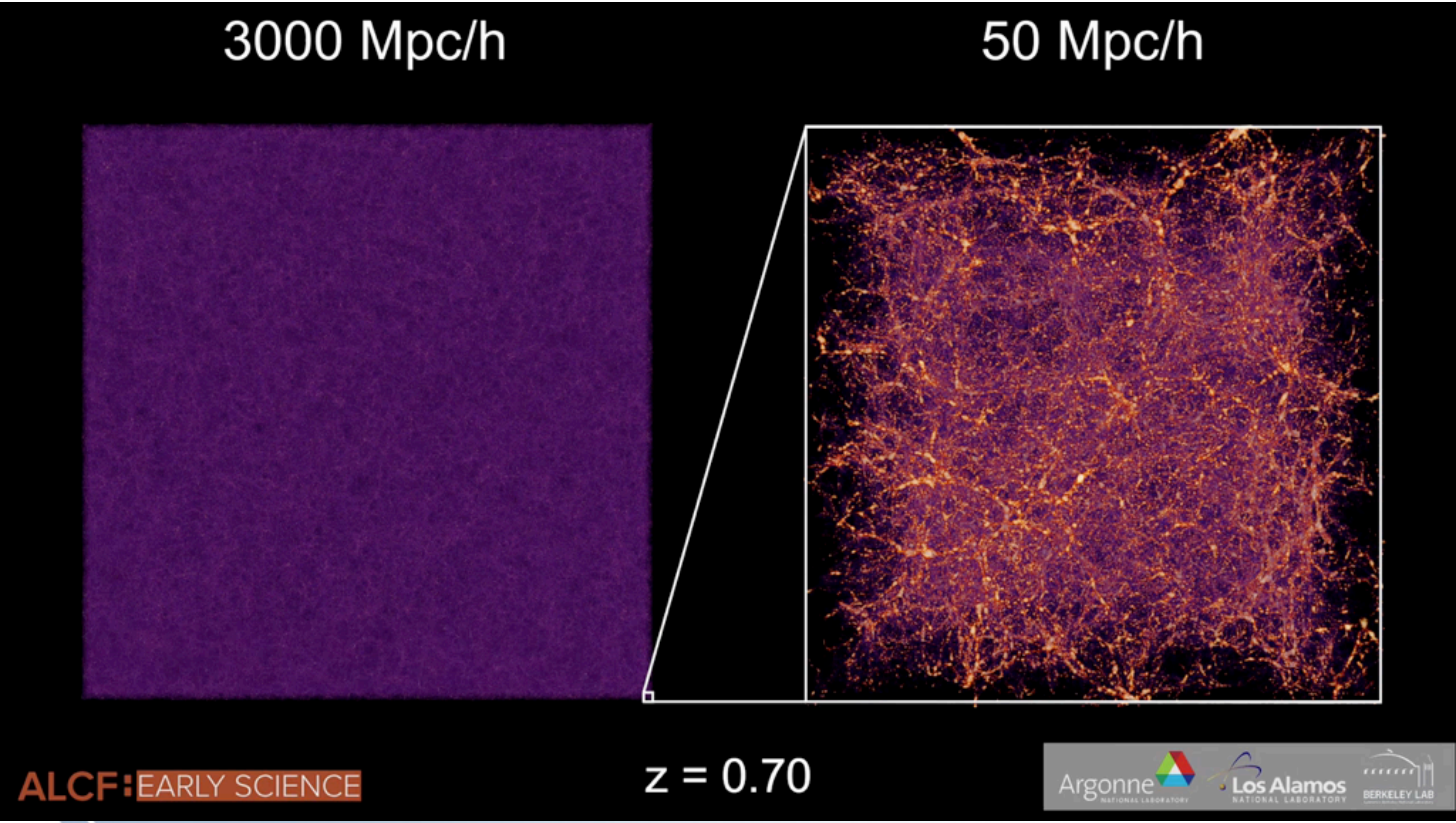
Movie Captures: Growth of Structure



Movie Captures: Growth of Structure



Movie Captures: Spatial Dynamic Range



Movie Captures: Fly-Through