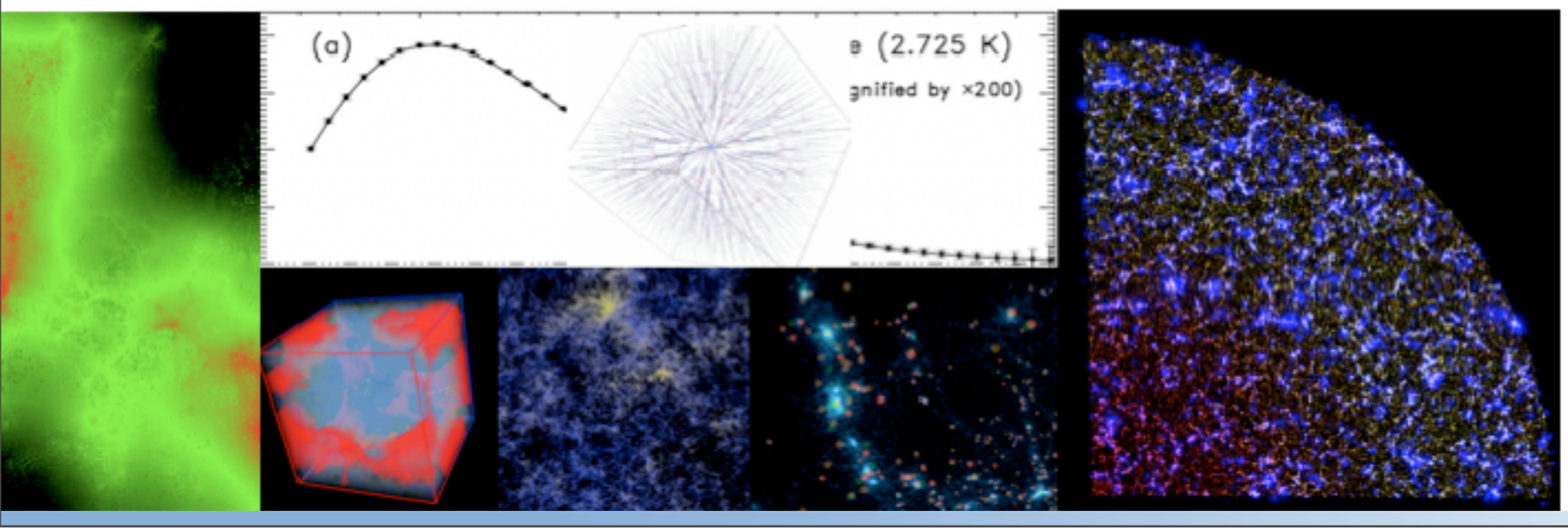
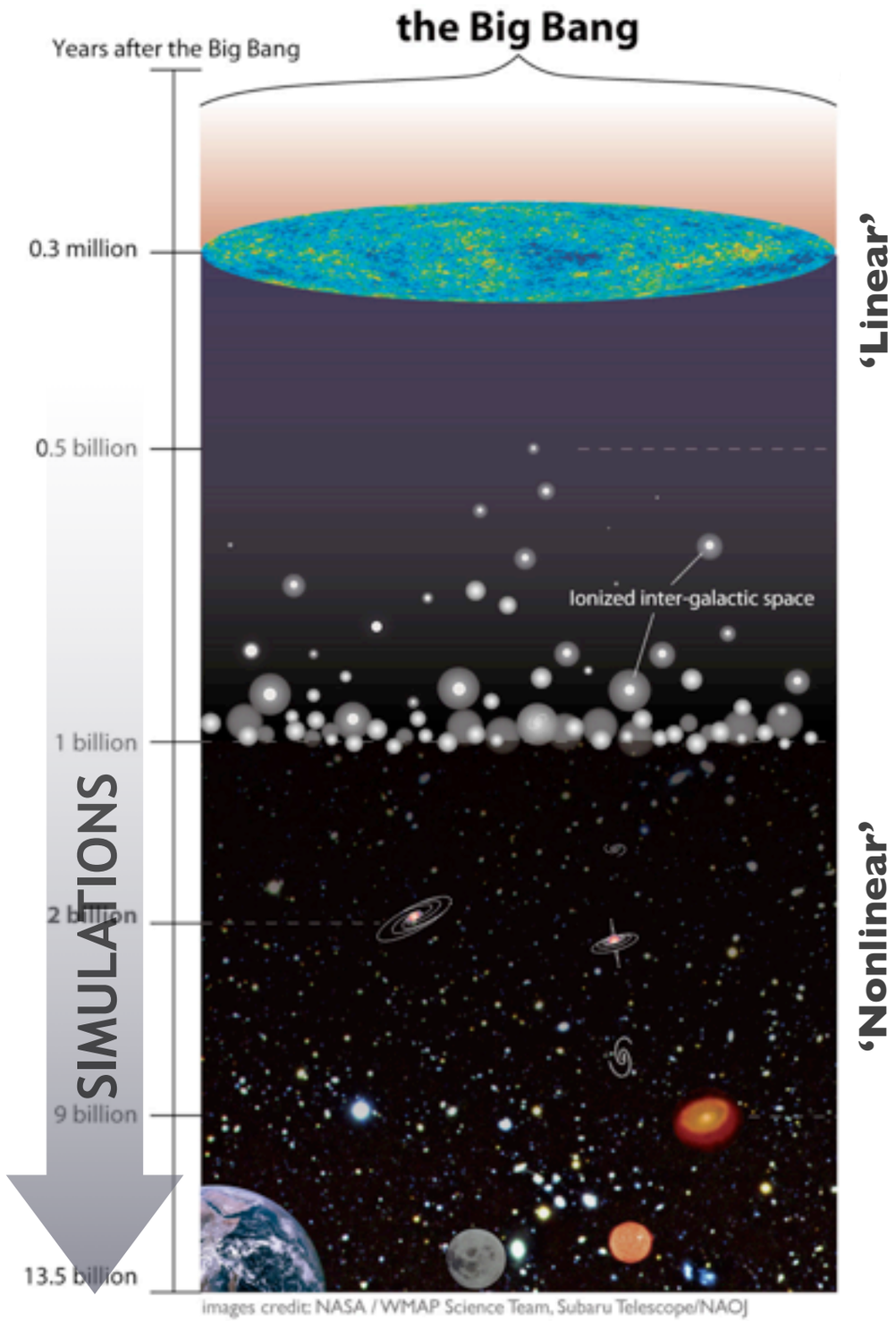
Summary/Survey:Computational Cosmology

Salman Habib



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 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**



Prehistory: An Early Simulation

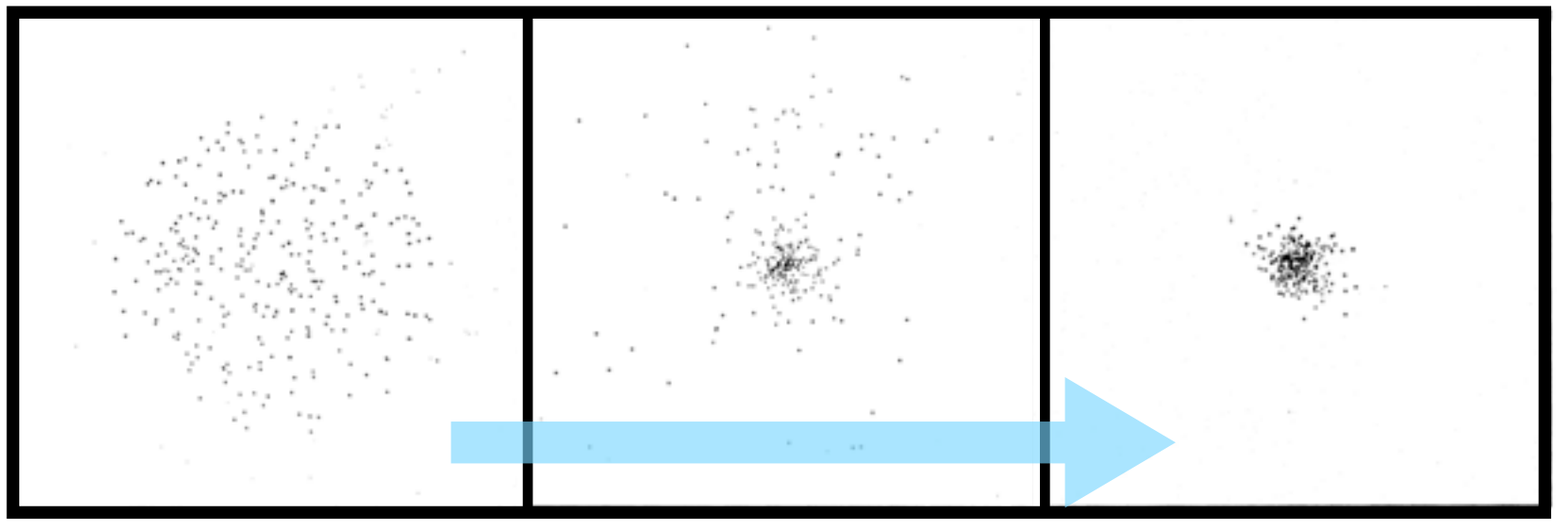
- Suite of 300 (and less) particle simulations
- Run on a CDC 3600, ~1Mflops, 32KB+ at LANL
- Is nine orders of magnitude improvement in both performance and memory good enough for precision cosmology?

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.

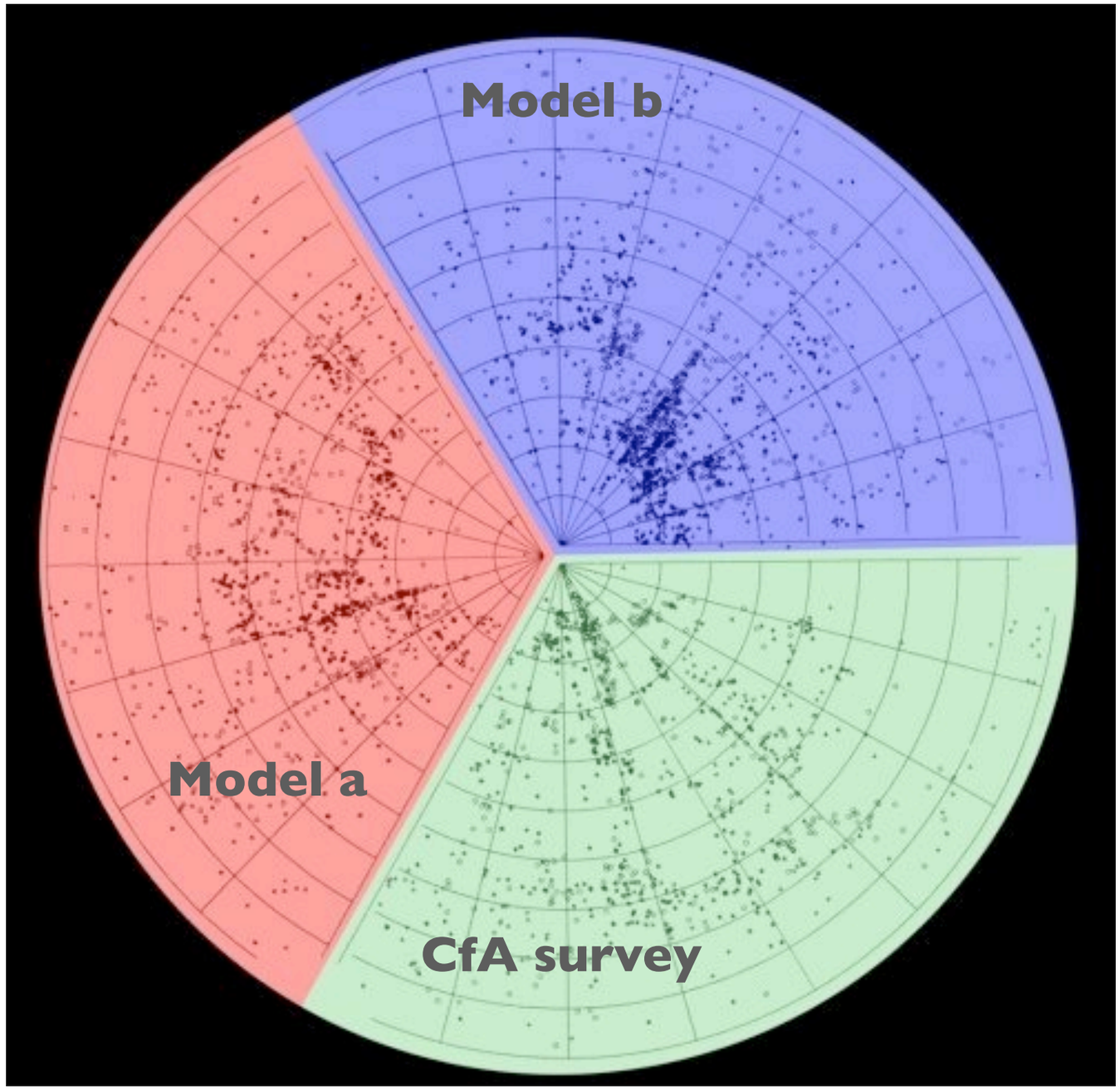


"The Universe is far too complicated a structure to be studied deductively, starting from initial conditions and solving the equations of motion."

Robert Dicke (Jayne Lectures, 1969)

Fast Forward: A Compressed History

- **Beginnings (60s/70s)**
  - Primarily direct particle-particle methods, no theory of initial conditions
- **'Medieval' Period (80s)**
  - Inflation appears; Zeldovich approximation applied to ICs, application of state of the art simulation methods
- **'Modern' Period (90s)**
  - Multi-resolution parallel codes, gasdynamics simulations approach maturity
- **Current Era**
  - Transition from qualitative to quantitative predictions, driven by observations

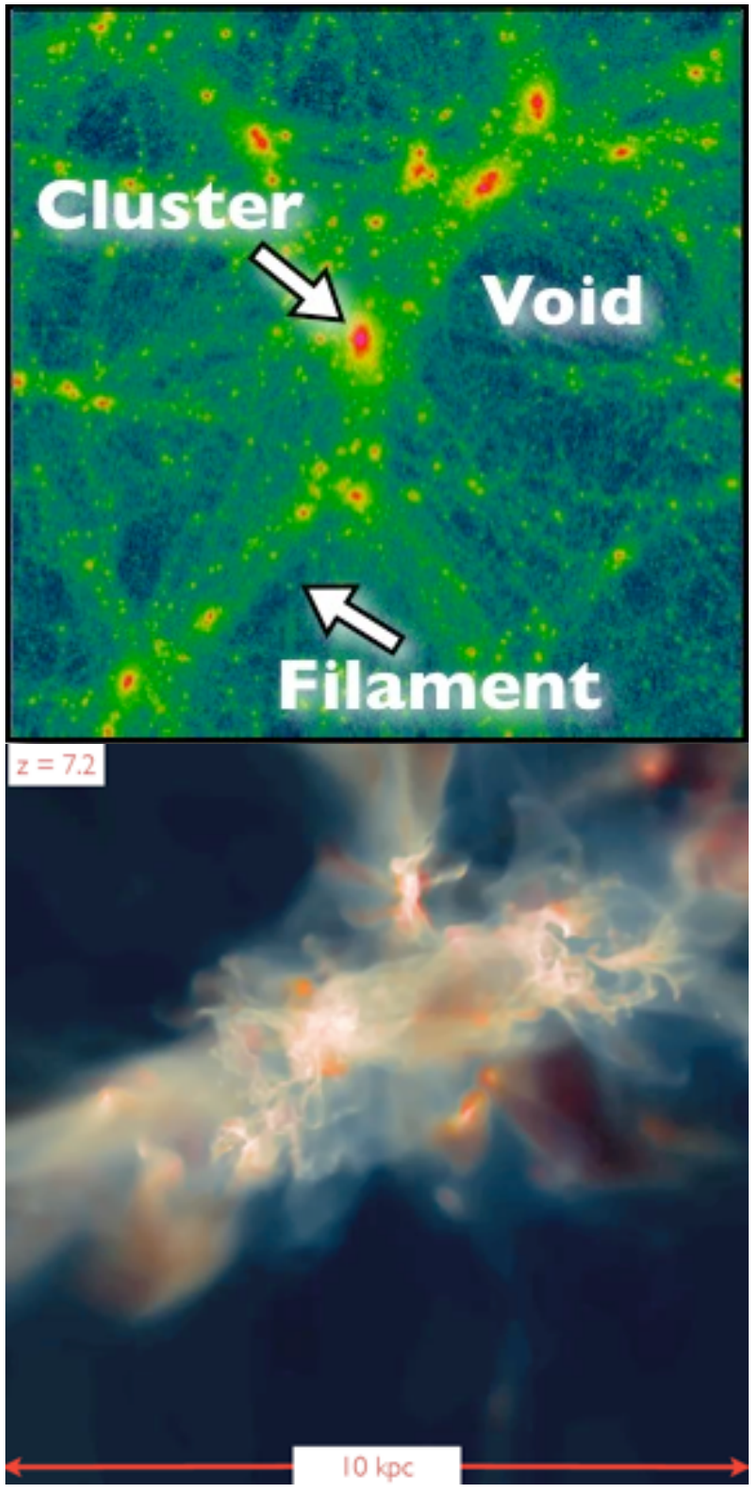


Davis et al 1985

Mock catalogs from 25 years ago for "eyeball" comparisons with the CfA galaxy survey

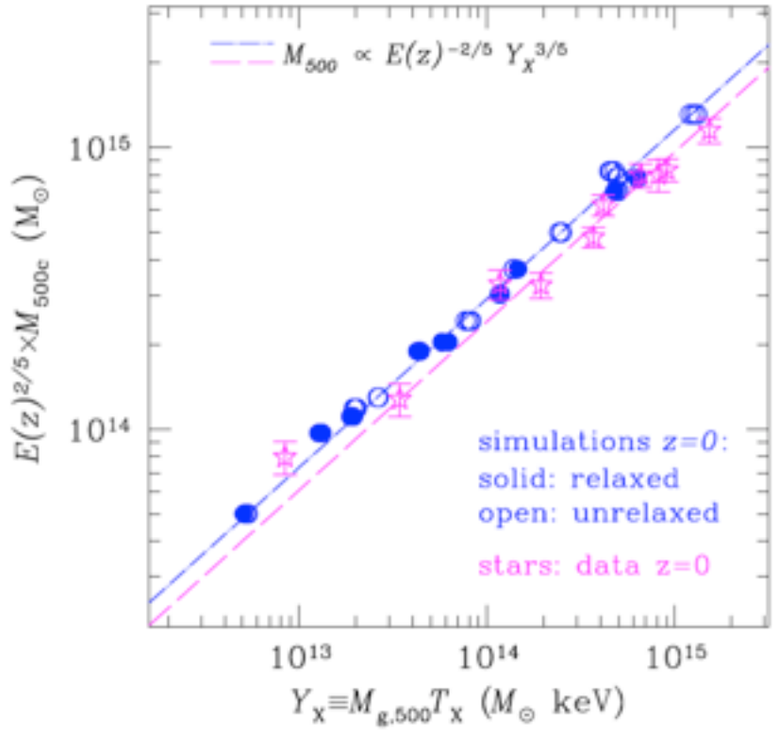
Achievements

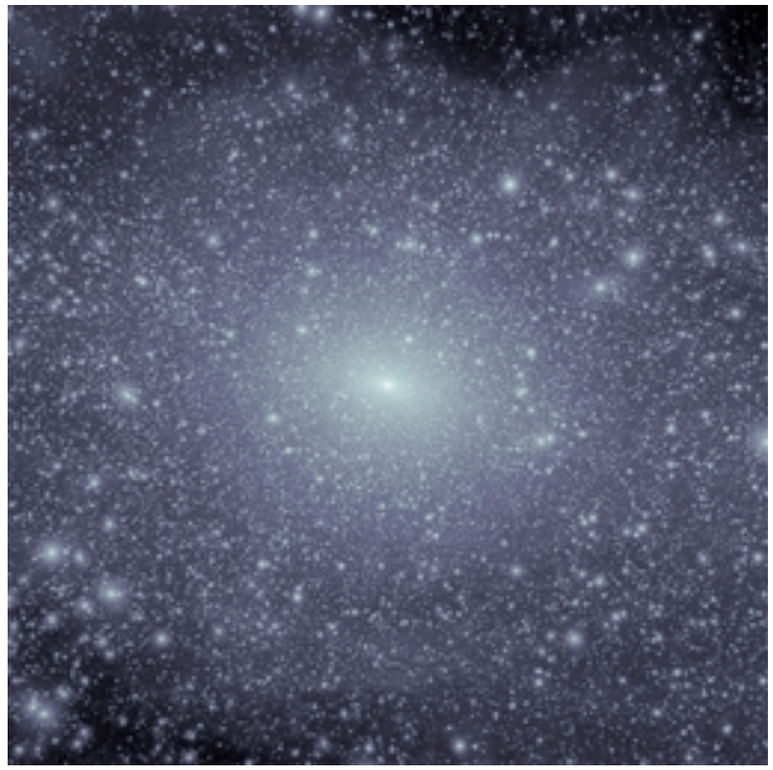
- **Paradigm of Structure Formation**
  - "Cosmic Web" unifies a wide array of observations, from galaxy surveys to the Ly-alpha forest
  - Predictions for halo shapes, substructure, spatial statistics, halo bias (clustering relative to underlying density field), and their evolution
- **Cosmic Probes**
  - Halo abundance and mass distribution
  - Nonlinear domains of clustering probes
- **Physics of the IGM, clusters, reionization --**
  - Multi-resolution, multi-physics simulations essential tools for building more detailed picture, understanding of empirical relations and their scatter



Examples of Current Results

- **Cosmology Simulation Suites**
  - LasDamas -- large suite for a single cosmology, many realizations and box sizes (SDSS mocks)
  - Coyote Universe -- large simulation for multiple cosmologies, very fast ('instantaneous') emulation for observables
- **Galaxy Modeling for Surveys**
  - High-resolution, 'small' volume simulation to follow details of structure formation (Bolshoi, Millennium II, MW TreePM, --)
  - Baryonic effects on weak lensing
- **Cluster Physics**
  - Realistic cluster simulations for understanding scatter of mass-observable relations
- **Ultra-high Resolution 'Milky Way'**
  - Aquarius, Via Lactea, Silver River, --

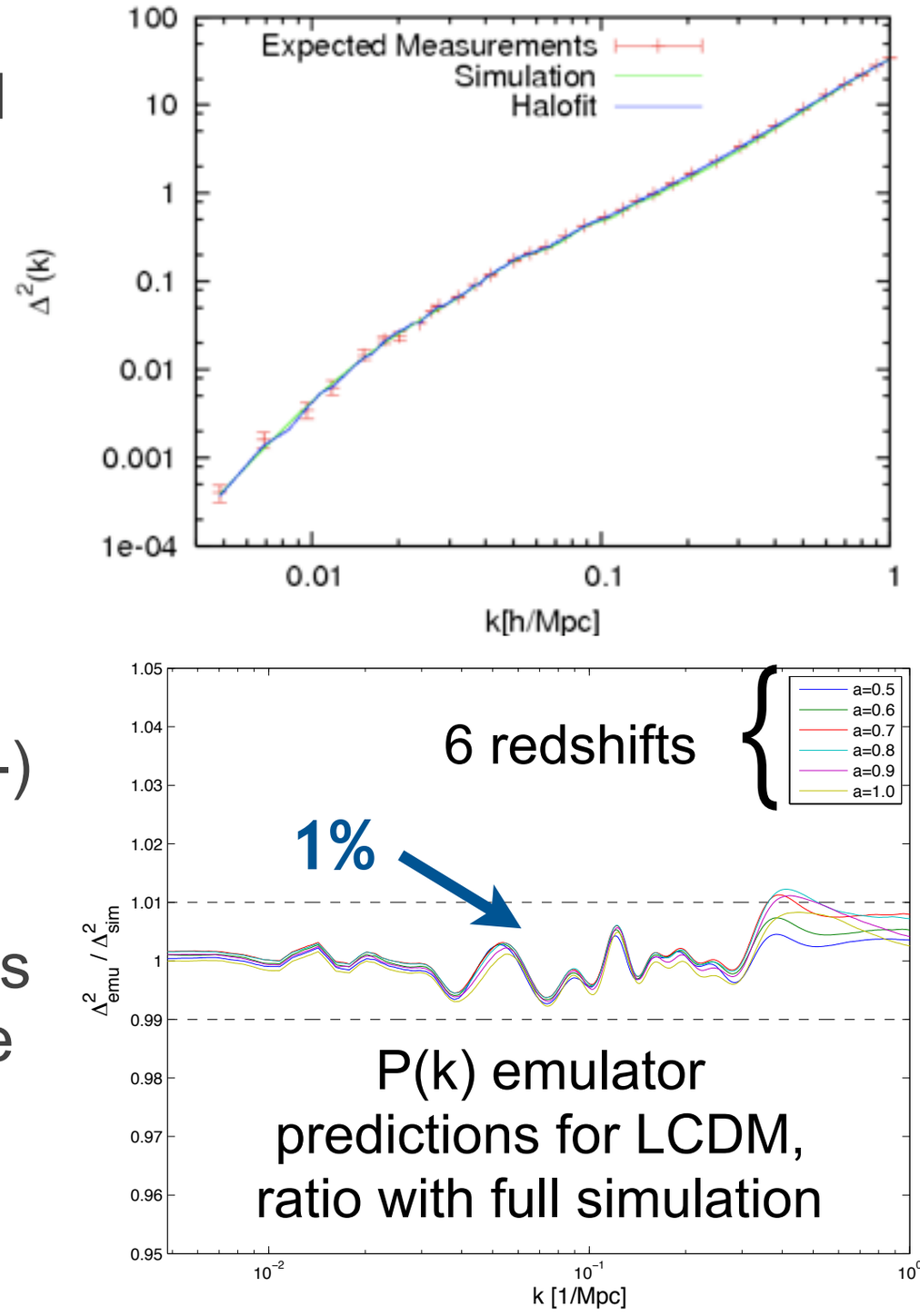




U.S. DEPARTMENT OF ENERGY | Office of Science

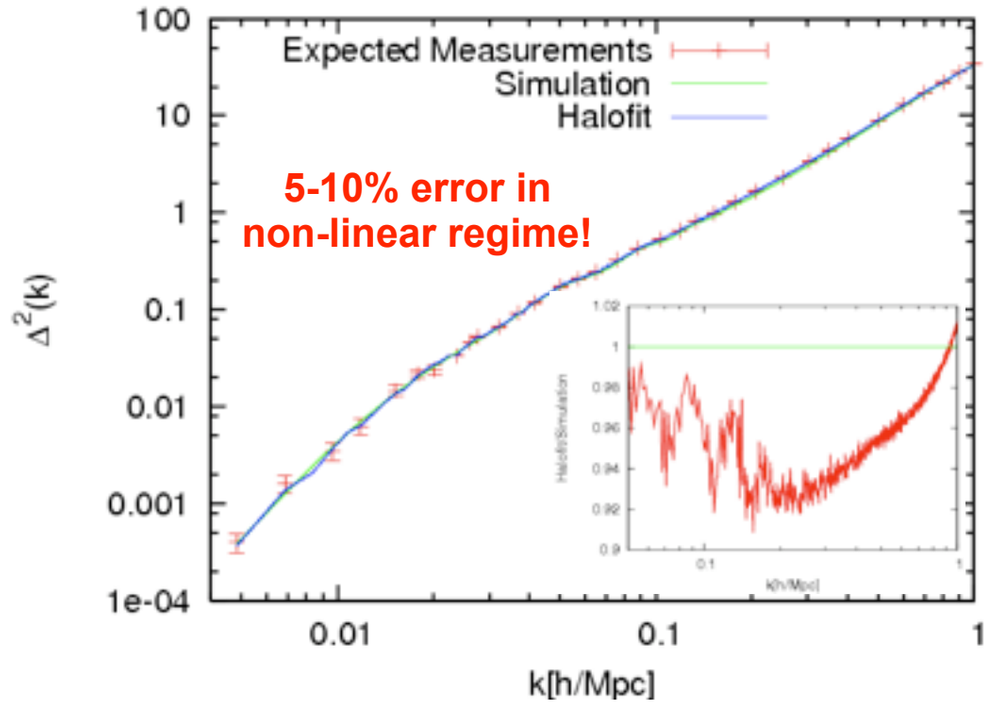
Precision Cosmology: Modeling and Observations

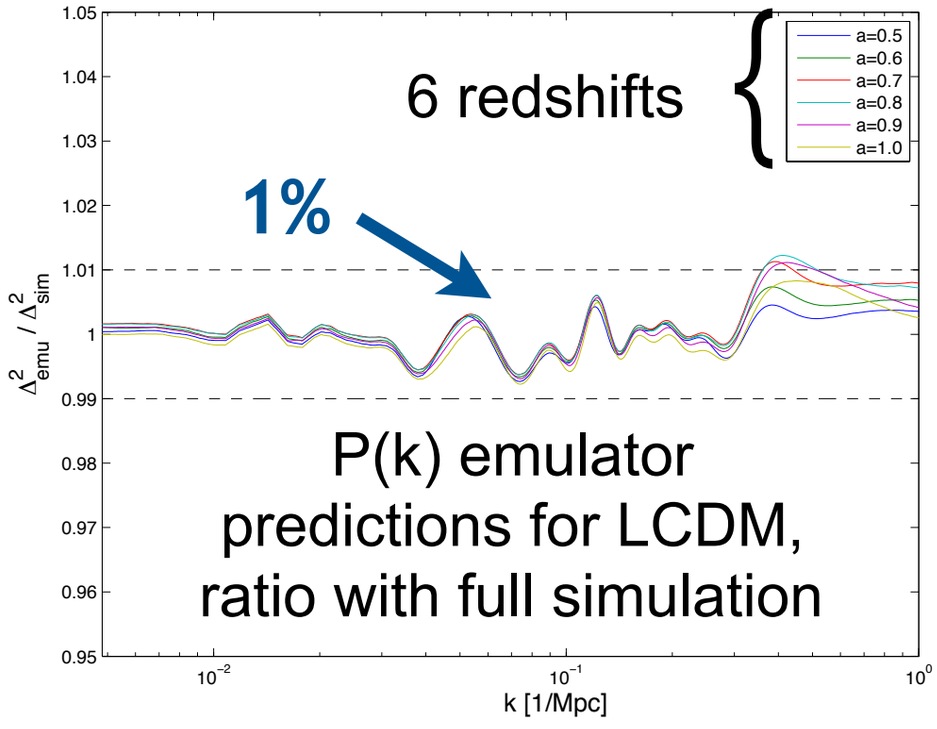
- **Early Period**
  - Access to large data choked (technology did not exist), insignificant computing
  - Characterized by small datasets, ‘eyeball’ comparisons, simple statistics
- **Intermediate Phenomenology (~10%)**
  - Use simulation to build intermediate, simplified theoretical model; use this to interact with observations (HOD models, HaloFit, scaling relations from simulations, --)
- **‘Direct’ Numerical Phenomenology (~1%)**
  - ‘Theory’ = direct interaction with observations by sophisticated simulations (or intermediate numerical products, ‘emulators’); complex space -- understand systematic errors (missing/wrong physics), bias



Precision Cosmology: Modeling and Observations

- **Early Period**
  - Access to large data choked (technology did not exist), insignificant computing
  - Characterized by small datasets, ‘eyeball’ comparisons, simple statistics
- **Intermediate Phenomenology (~10%)**
  - Use simulation to build intermediate, simplified theoretical model; use this to interact with observations (HOD models, HaloFit, scaling relations from simulations, --)
- **‘Direct’ Numerical Phenomenology (~1%)**
  - ‘Theory’ = direct interaction with observations by sophisticated simulations (or intermediate numerical products, ‘emulators’); complex space -- understand systematic errors (missing/wrong physics), bias

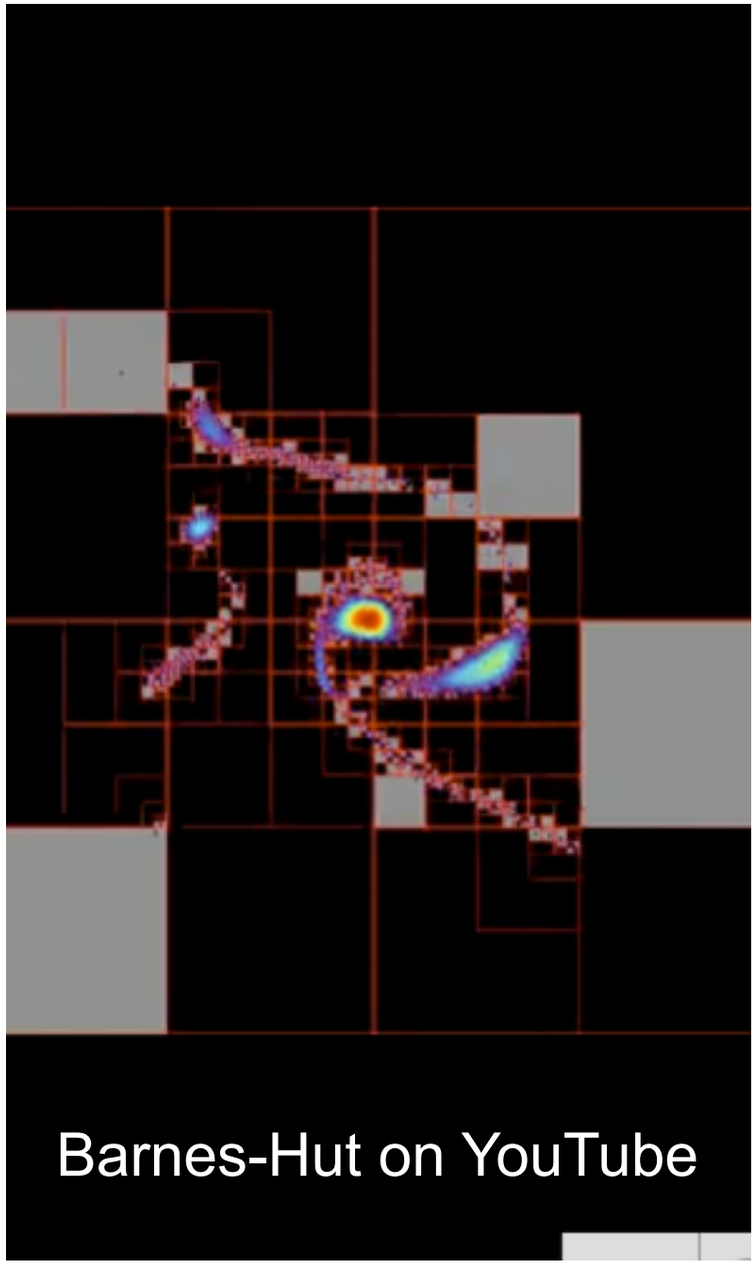




U.S. DEPARTMENT OF ENERGY | Office of Science

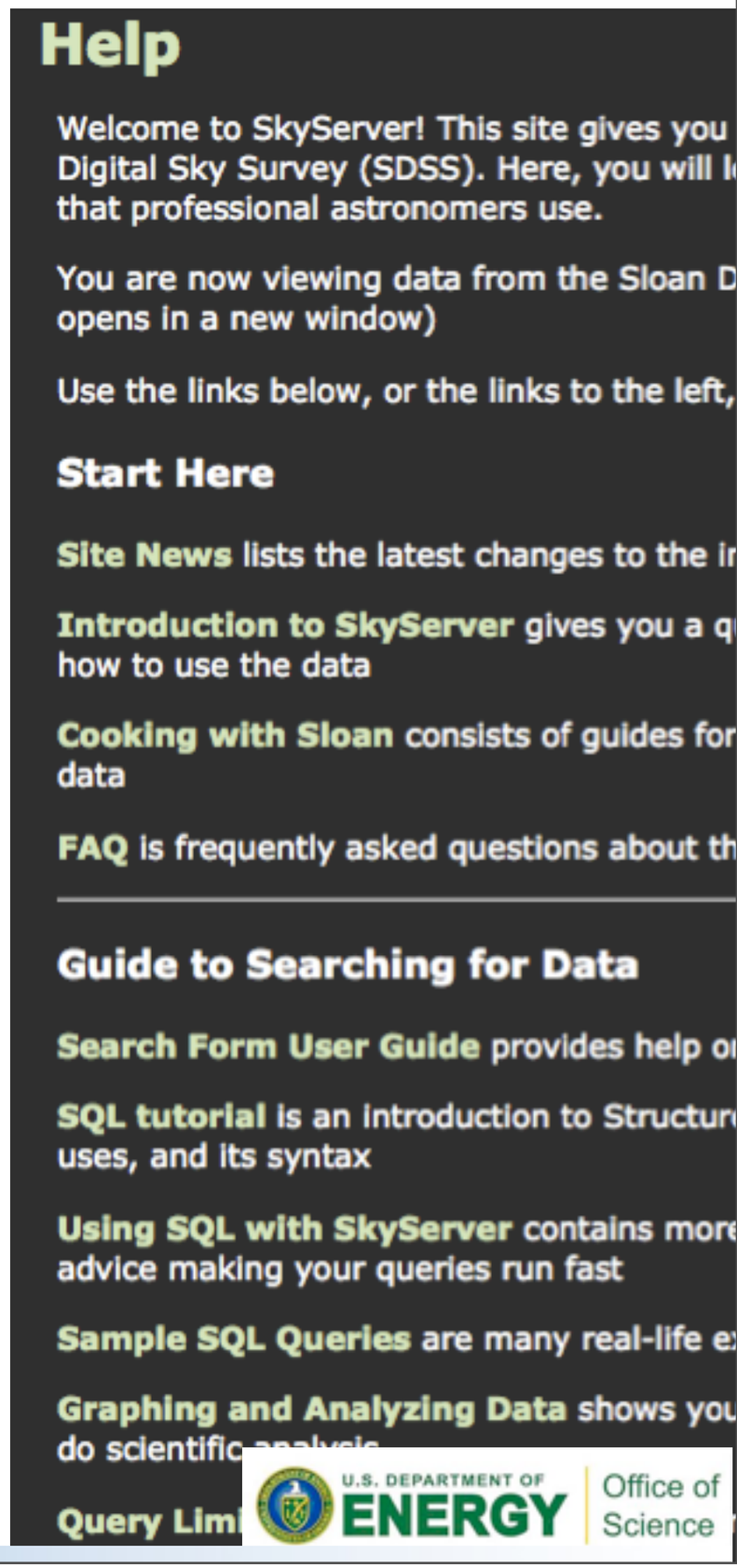
Computational Cosmology and HPC

- **Early Era**
  - Early use of vector systems, methods borrowed from plasma physics (PM, adaptive P3M, AMR) plus tree algorithms
- **Parallel Simulations**
  - Early parallelization of all codes, efficient utilization of supercomputing resources (several Gordon Bell prizes, including custom hardware)
- **More Physics: Multi-Physics/Multi-Resolution**
  - Several cosmology codes have complexity similar to combustion codes, others are aimed at optimizing performance and memory (survey scale simulations)
- **Next-Generation Architectures**
  - Increased concurrency with communication bottlenecks and costs -- codes must evolve to deal with these issues



Computational Cosmology and Large Data

- **Digital Data**
  - Plate digitization
  - CCD sky surveys, SDSS (10s of TB), many ‘use’ modes -- very rich scientifically
- **Next-Generation Surveys**
  - DES starting (~PB total)
  - LSST ‘threatening’ (~100 PB total)
  - Planck ongoing (analysis at NERSC)
- **Analysis modes**
  - Many ‘single’ applications (each complex)
  - Very many cross-correlations to constrain systematics, tease out new signals
- **Problems of Scale**
  - Simulation data and observational data sizes enormous, need to implement new algorithms and methodologies



U.S. DEPARTMENT OF ENERGY | Office of Science

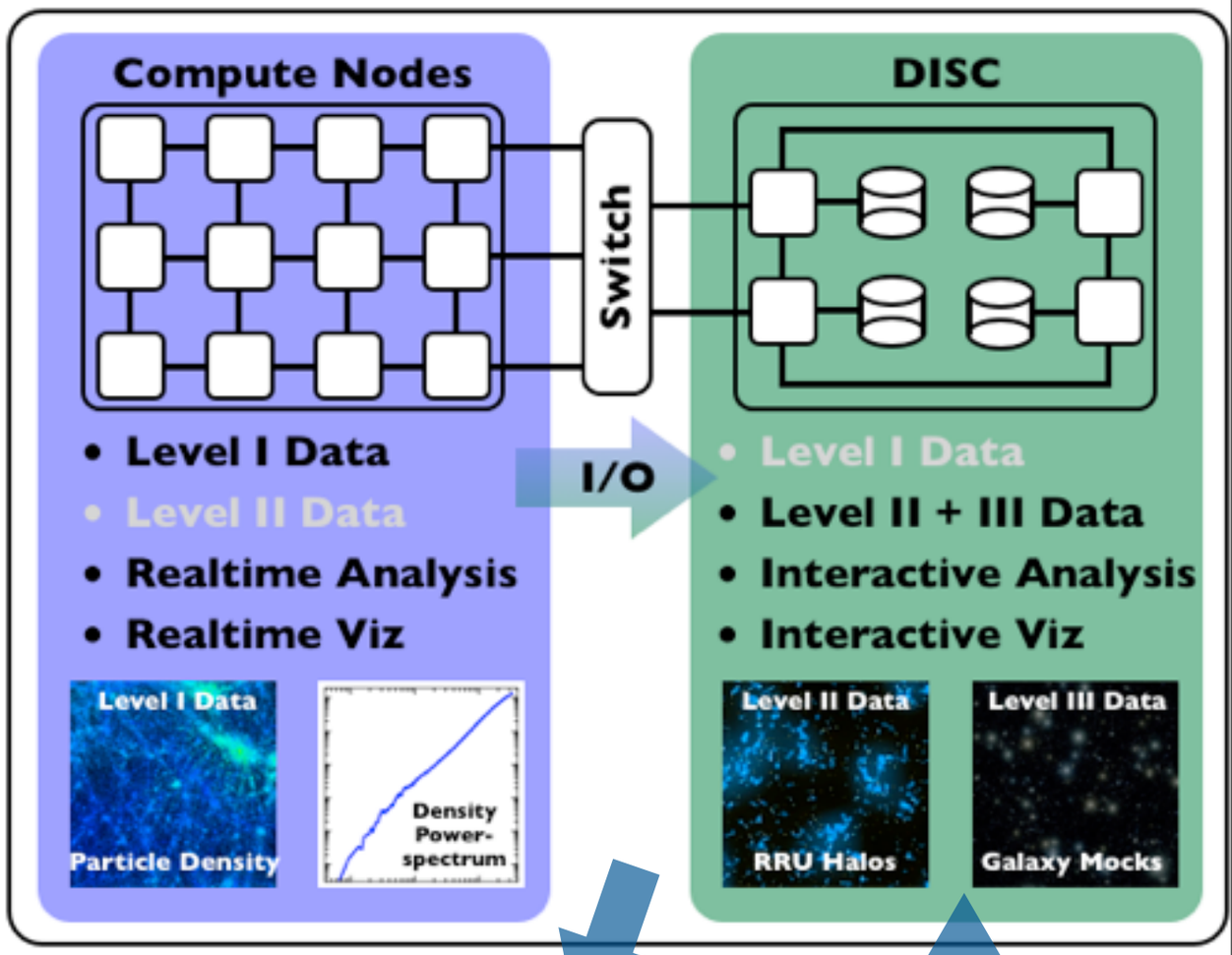
‘Ultimate Vision’

- **Clean up ‘Discovery Space’:** Robust theory with subtle signals
- **Precision Cosmic Calibration at Scale:** ‘All Sky’ solution of the cosmic inverse problems in the nonlinear regime
- **Cosmology Simulations at the Exascale:** Next-generation computing and beyond as essential theoretical and analysis tools

18 INTERNATIONAL EXASCALE SOFTWARE PROJECT

SciDAC Scientific Discovery through Advanced Computing

- **Large Data:** Simulation & observational datastreams: Archiving, serving, quality assurance, (joint) analyses
- **Simulation/Data/Analysis:** Cross-platform, multi-source, analysis and interrogation frameworks



U.S. DEPARTMENT OF ENERGY | Office of Science