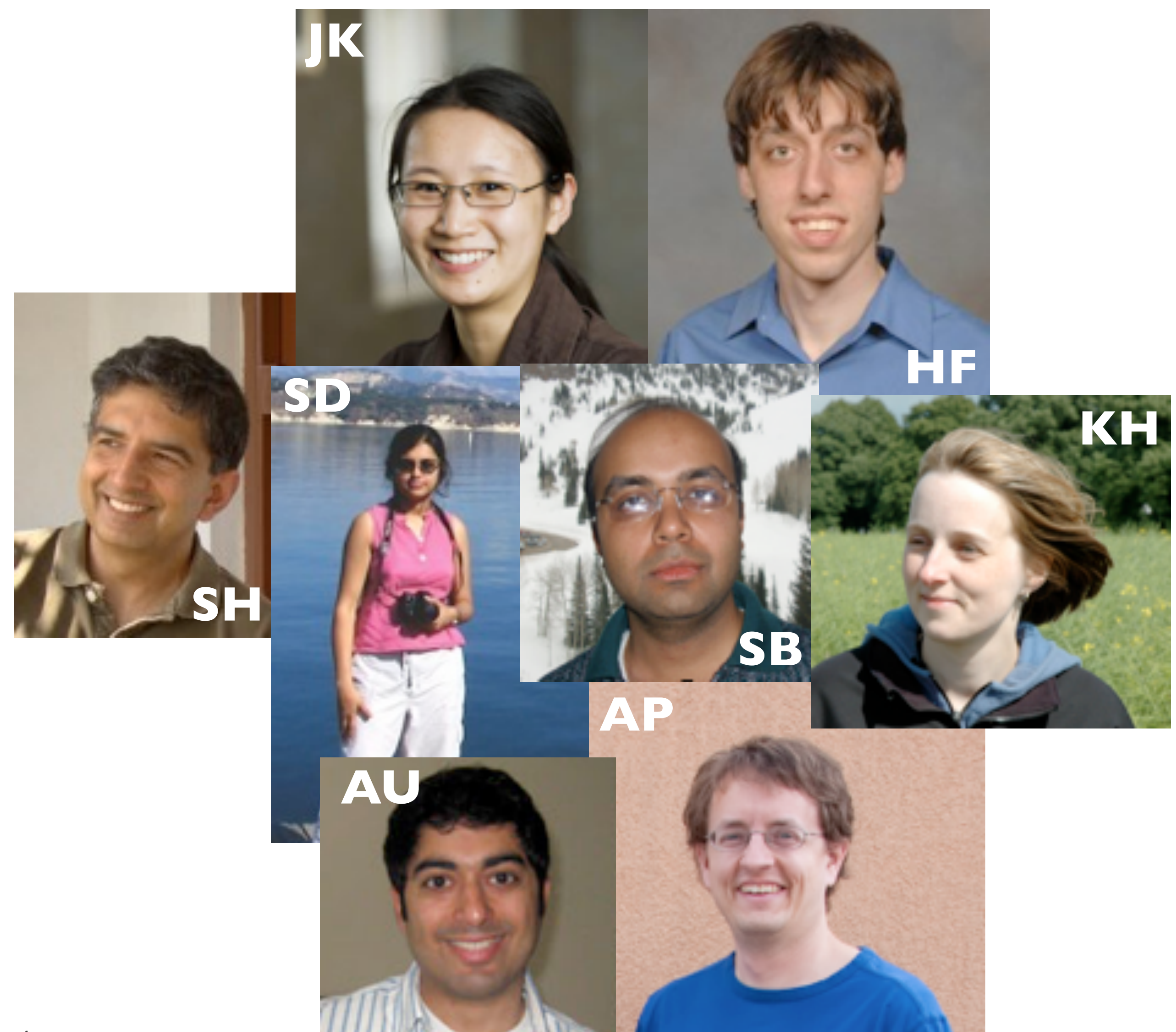
Computational Cosmology Initiative at Argonne National Laboratory

HEP staff:
Salman Habib, Katrin Heitmann

HEP post-docs:
Sanghamitra Deb, Juliana Kwan*, Adrian Pope, Amol Upadhye

Argonne Leadership Computing Facility post-doc:
Hal Finkel

HEP/UChicago post-doc:
Suman Bhattacharya*



*joining Fall 2011

Collaborations and Projects

Kev Abazajian, Jim Ahrens, Ujjaini Alam, Debbie Bard, Andrew Benson, Gary Bernstein, Joe Bernstein, Derek Bingham, Rahul Biswas, Karen Byrum, Anna Cabre, Jordan Carlson, Joanne Cohn, Andrew Connolly, David Daniel, Nehal Desai, Scott Dodelson, Tim Eifler, Gus Evrard, Patricia Fasel, Hume Feldman, Wu Feng, Josh Frieman, NIck Frontiere, David Higdon, Tracy Holsclaw, Chung-Hsing Hsu, Bhuvnesh Jain, Gerard Jungman, Steve Kahn, Lloyd Knox, Savvas Koushiappas, Steve Kuhlmann, Earl Lawrence, Herbie Lee, Adam Lidz, Zarija Lukic, Chris Miller, Charles Nakhleh, Mike Norman, Alex Pang, Uliana Popov, Darren Reed, Paul Ricker, Robert Ryne, Bruno Sanso, Paul Sathre, Michael Schneider, Sergei Shandarin, Alex Szalay, Tony Tyson, Alberto Vallinotto, Licia Verde, Alexei Vikhlinin, Christian Wagner, Mike Warren, Martin White, Brian Williams, Tim Williams, Jon Woodring, ---

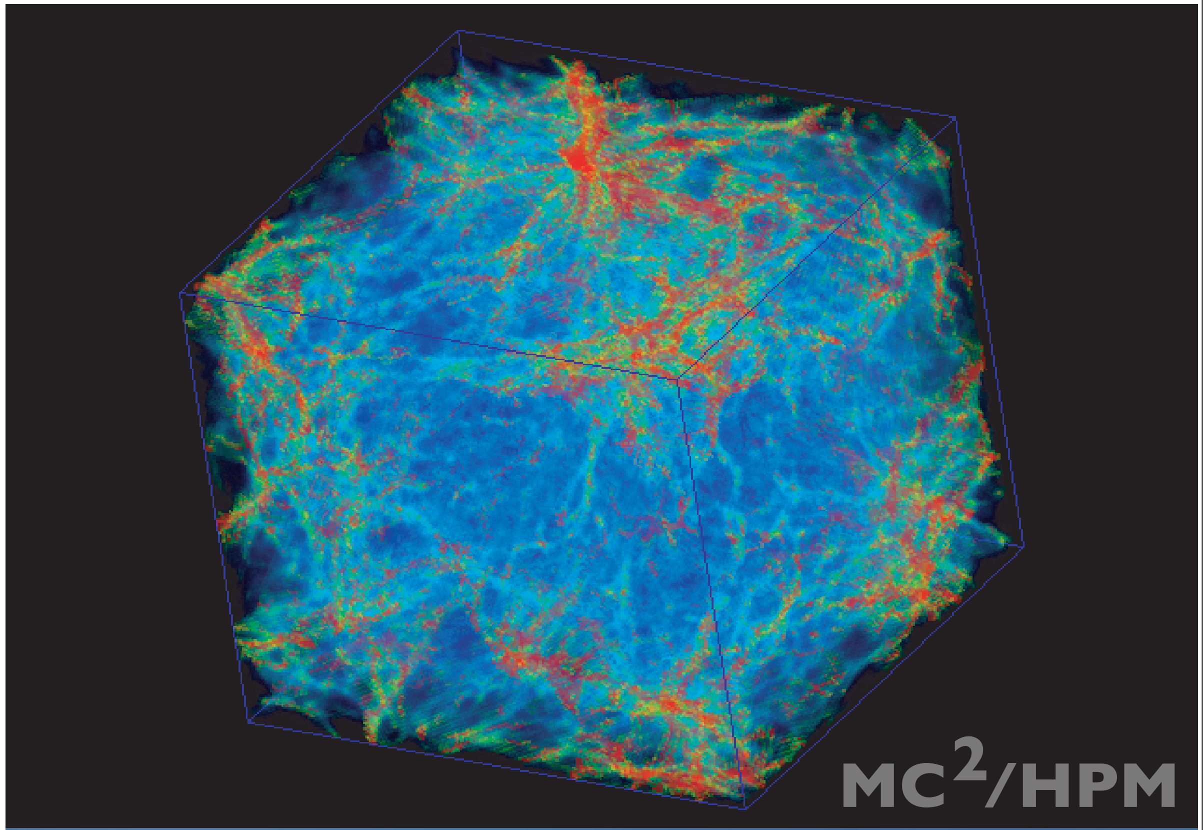
Collaborations and Projects

Kev Abazajian, Jim Ahrens, Ujjaini Alam, Debbie Bard, Andrew Benson, Gary Bernstein, Joe Bernstein, Derek Bingham, Rahul Biswas, Karen Byrum, Anna Cabre, Jordan Carlson, Joanne Cohn, Andrew Connolly, David Daniel, Nehal Desai, Scott Dodelson, Tim Eifler, Gus Evrard, Patricia Fasel, Hume Feldman, Wu Feng, Josh Frieman, NIck Frontiere, David Higdon, Tracy Holsclaw, Chung-Hsing Hsu, Bhuvnesh Jain, Gerard Jungman, Steve Kahn, Lloyd Knox, Savvas Koushiappas, Steve Kuhlmann, Earl Lawrence, Herbie Lee, Adam Lidz, Zarija Lukic, Chris Miller, Charles Nakhleh, Mike Norman, Alex Pang, Uliana Popov, Darren Reed, Paul Ricker, Robert Ryne, Bruno Sanso, Paul Sathre, Michael Schneider, Sergei Shandarin, Alex Szalay, Tony Tyson, Alberto Vallinotto, Licia Verde, Alexei Vikhlinin, Christian Wagner, Mike Warren, Martin White, Brian Williams, Tim Williams, Jon Woodring, ---

Computational Cosmology Collaboration

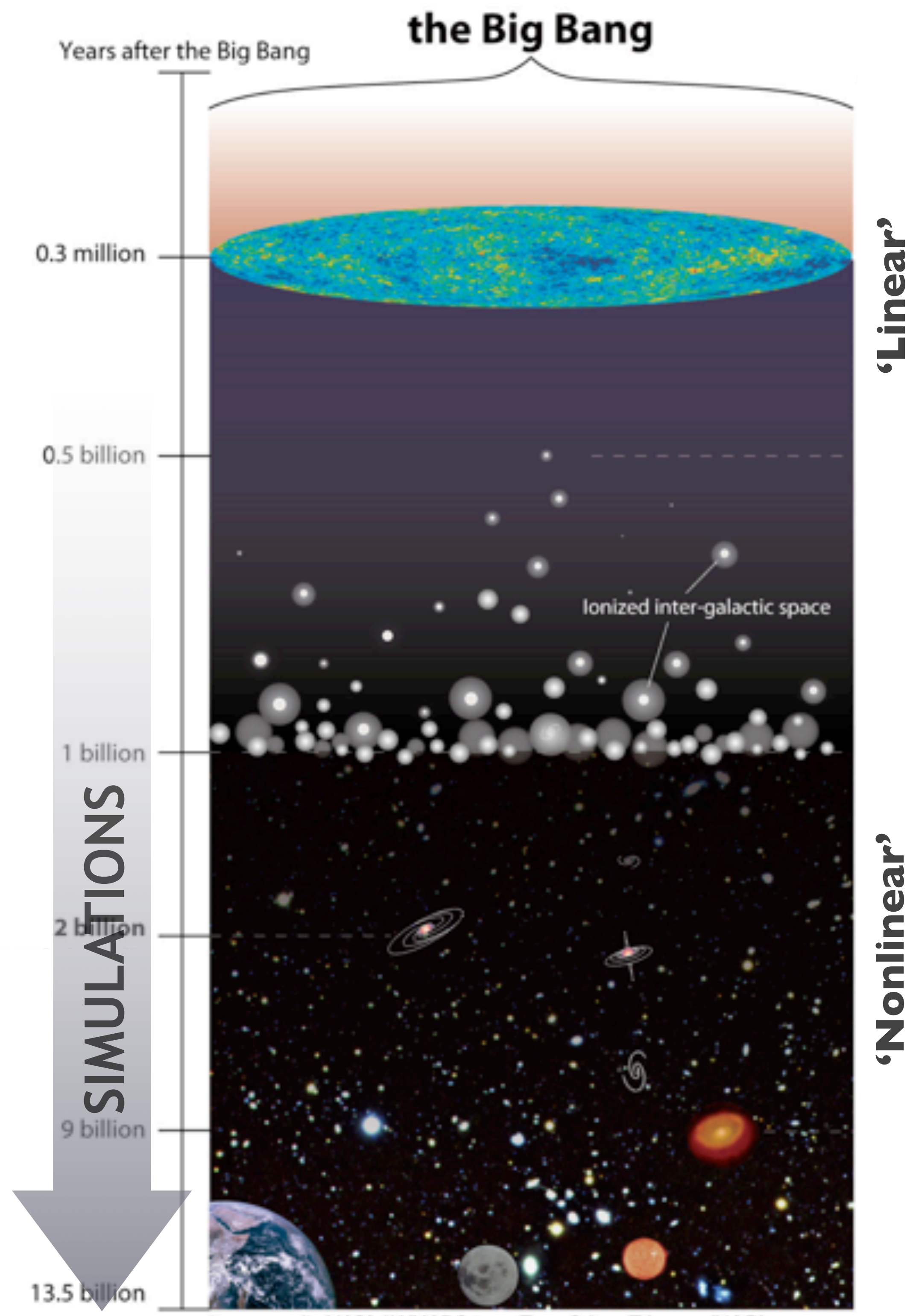
Overview

- **Primary Target of Group Research:** 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 (rough analogy with lattice QCD)
  - **Design and Analysis:** Advance 'Science of Surveys'; contribute to major 'Dark Universe' missions: BOSS, DES, LSST, BigBOSS, DESpec --



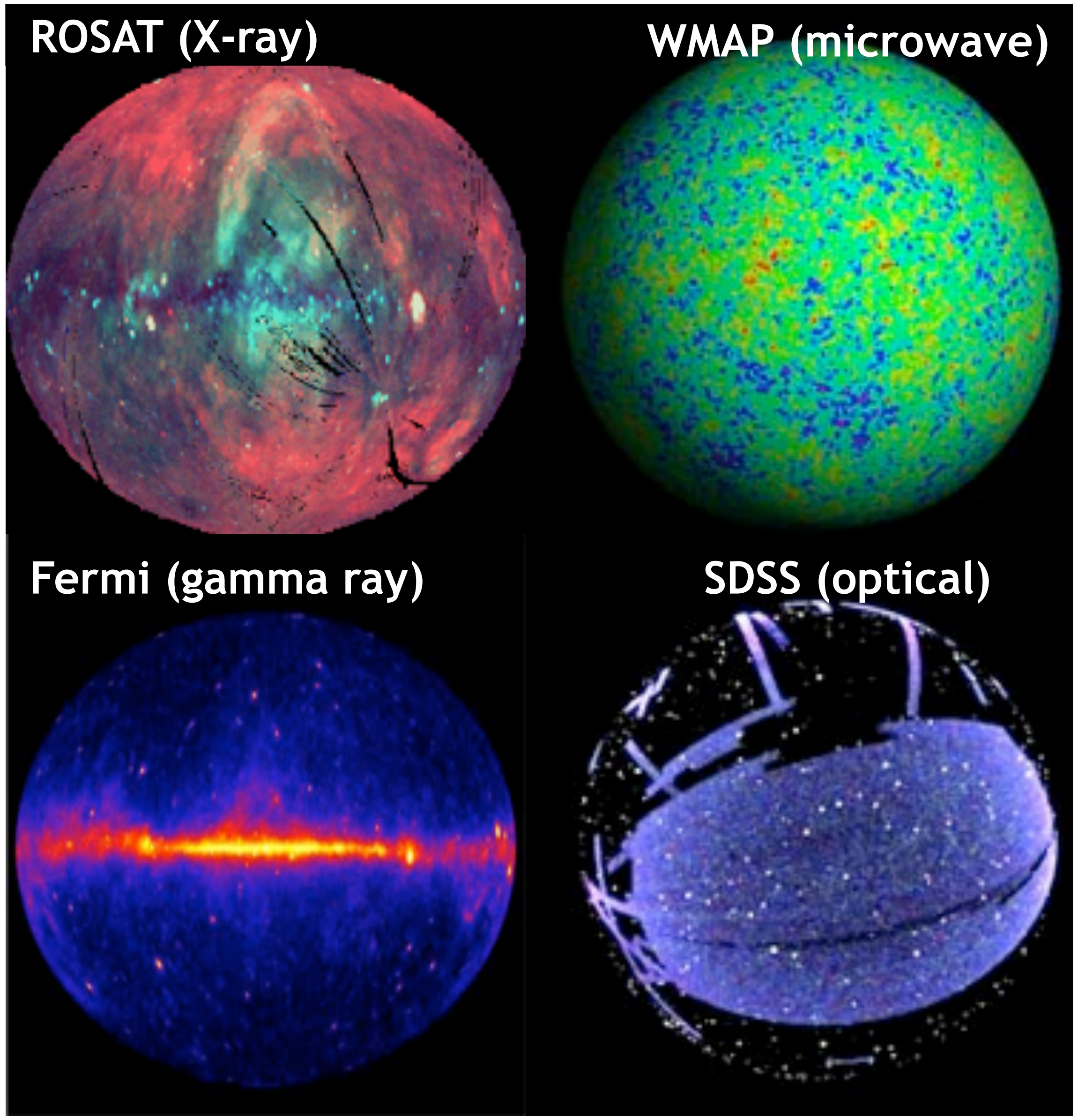
Structure Formation: The Basic Paradigm

- **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



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

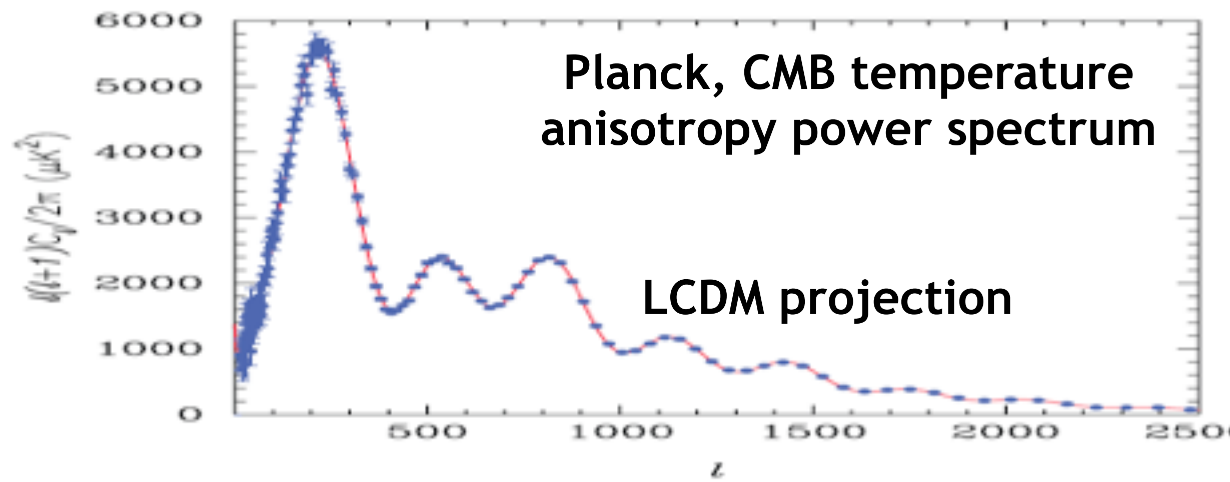


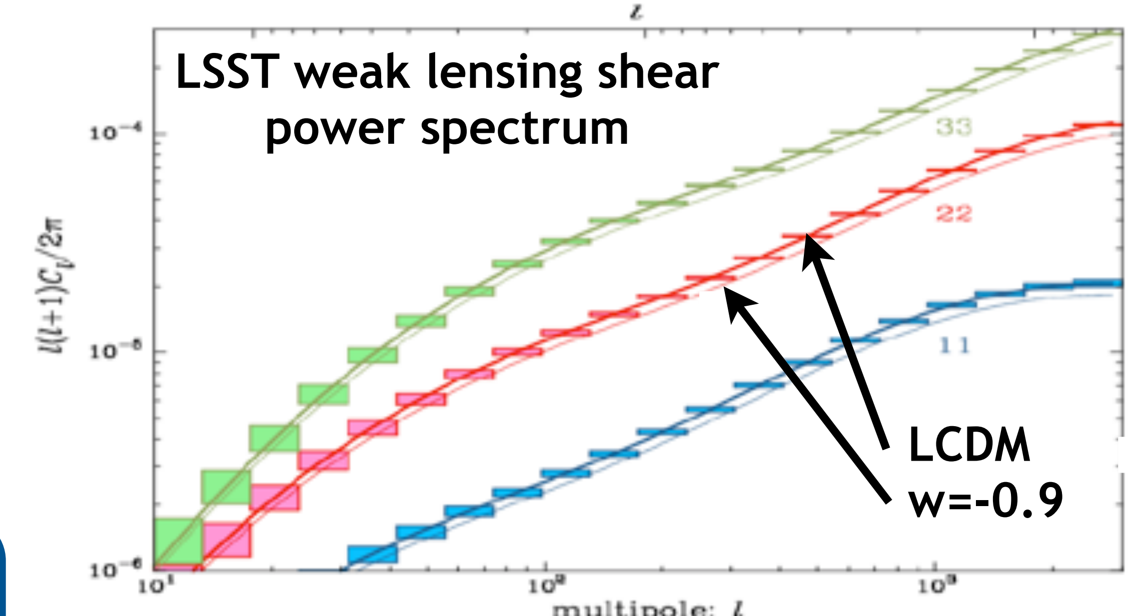
Explosion of information from sky maps: Precision Cosmology

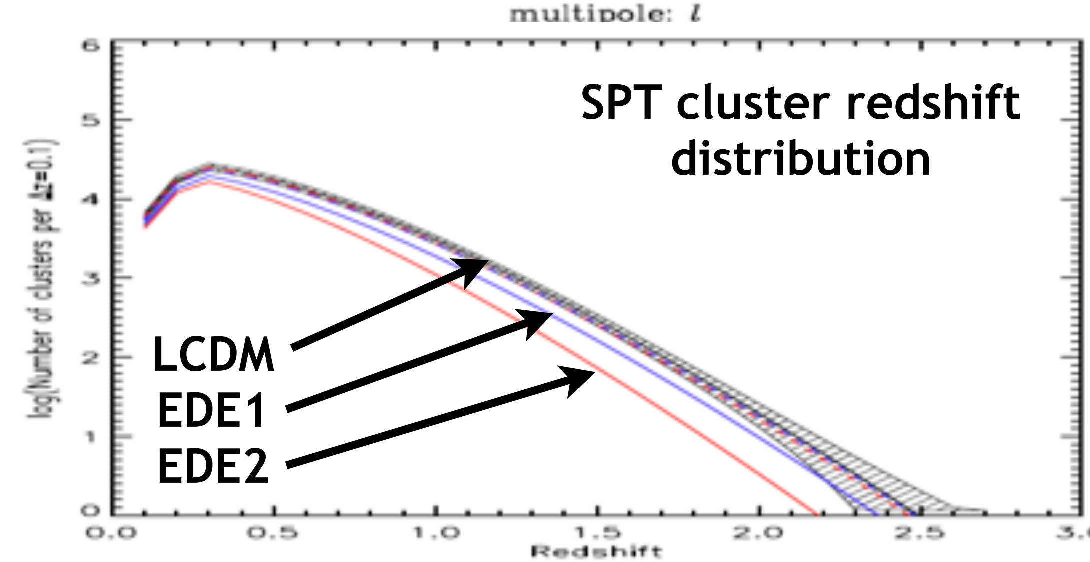
Precision Cosmology: “Inverting” the 3-D Sky

- **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 the ~1% level
- **The Challenge:** Theory and simulation must satisfy stringent criteria for inverse problems and precision cosmology not to be theory-limited!

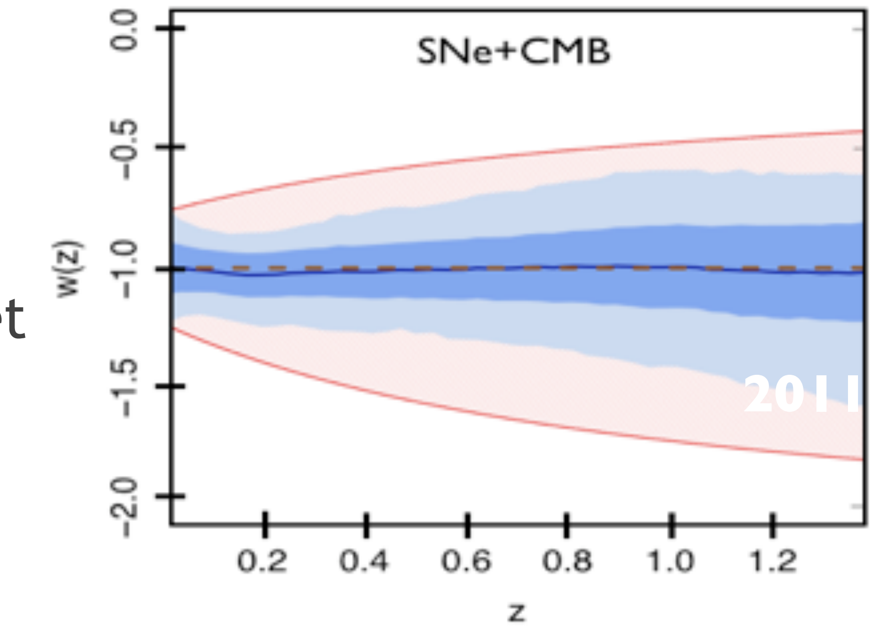


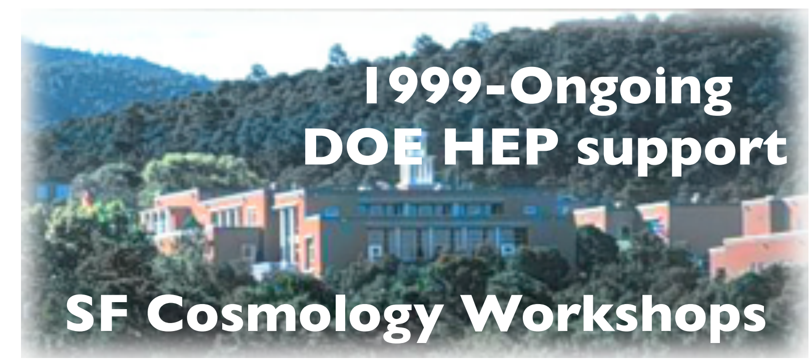


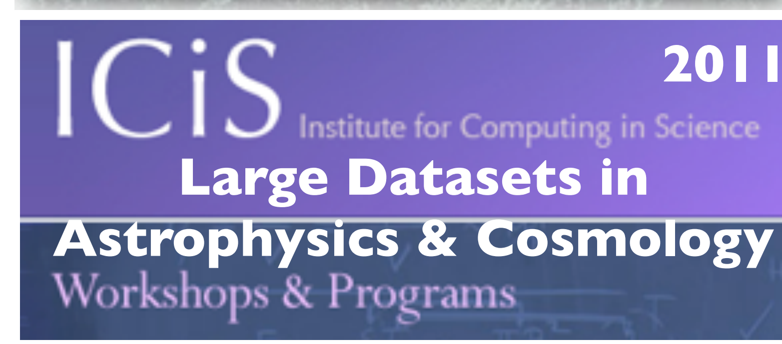


Recent Research Highlights/Activities

- **Nonparametric Dark Energy Reconstruction:** Holsclaw et al. PRL 2010, PRD 2010, PRD 2011 (submitted)
- **Dark Matter Halos and Structure Formation:** Bhattacharya et al. Ap. J. 2011, Lukic et al. Ap. J. 2009, Evrard et al. Ap. J. 2008, Lukic et al. Ap. J. 2007
- **Understanding the 'Cosmic Web':** Shandarin et al. PRD 2010
- **BAO with the Ly-α Forest:** White et al. Ap. J. 2010
- **Fossil Groups:** Voevodkin et al. Ap. J. 2010, 2008
- **Cosmic Calibration:** Habib et al. PRD 2007, Schneider et al. PRD 2008
- **Community Cosmology Code Comparison:** Heitmann et al. Comput. Sci. Dis. 2008
- **'Coyote Universe' Project:** Lawrence et al. Ap. J., 2010, Heitmann et al. Ap. J. 2010, 2009
- **Advanced Cosmological Simulations:** Habib et al. J. Phys. Conf. 2009, Pope et al. Comp. Sci. Eng. 2010





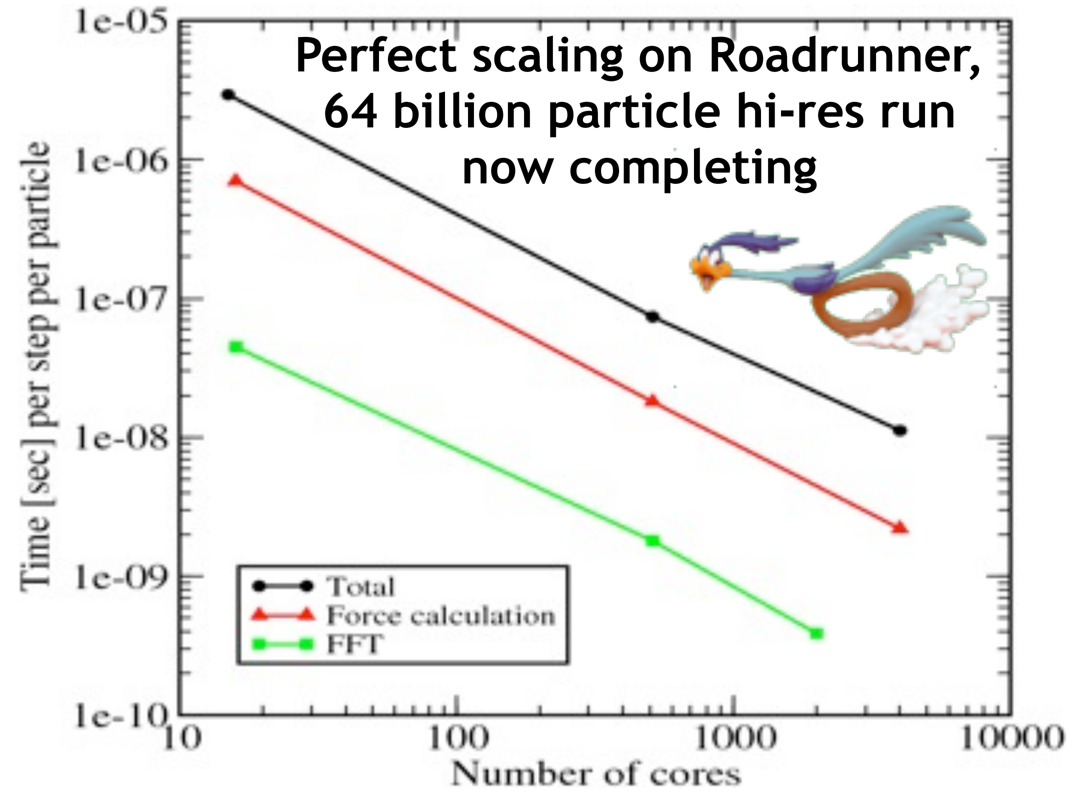


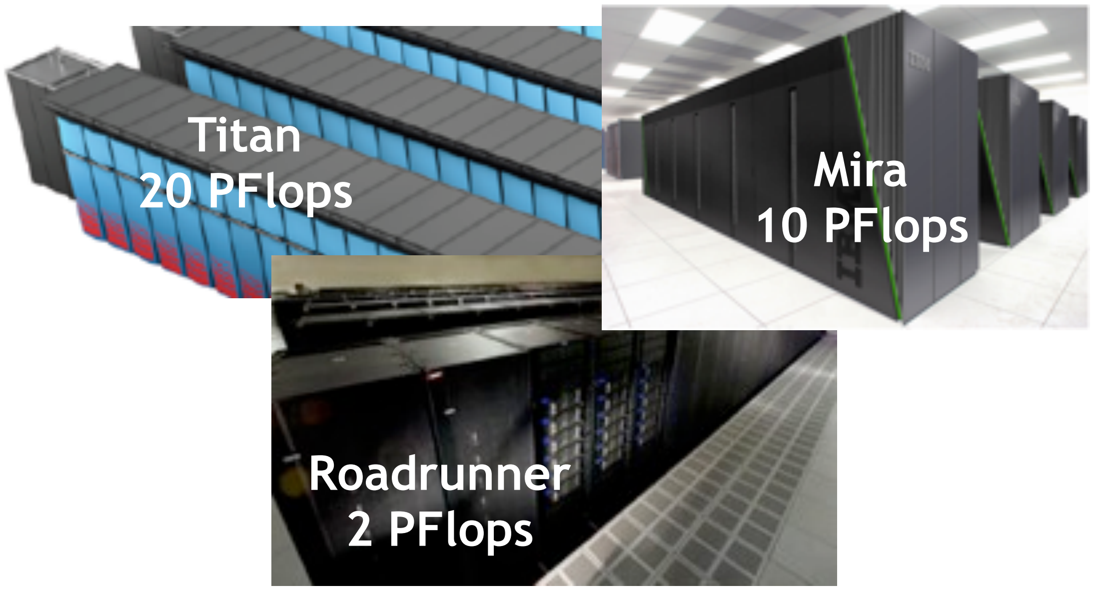
Computing the Universe: Simulating Surveys

- **Simulation Volume:** Large survey sizes impose simulation volumes ~ (3 Gpc)^3, with memory requirements ~**100** TB
- **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 cross-platform 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, high-order spectral operators, ~50% of peak Flops
- **Embedded Analysis:** High performance with low I/O and storage requirement
- **Early Science Project on Mira:** 150M CPU-hrs on ANL BG/Q (summer 2012)

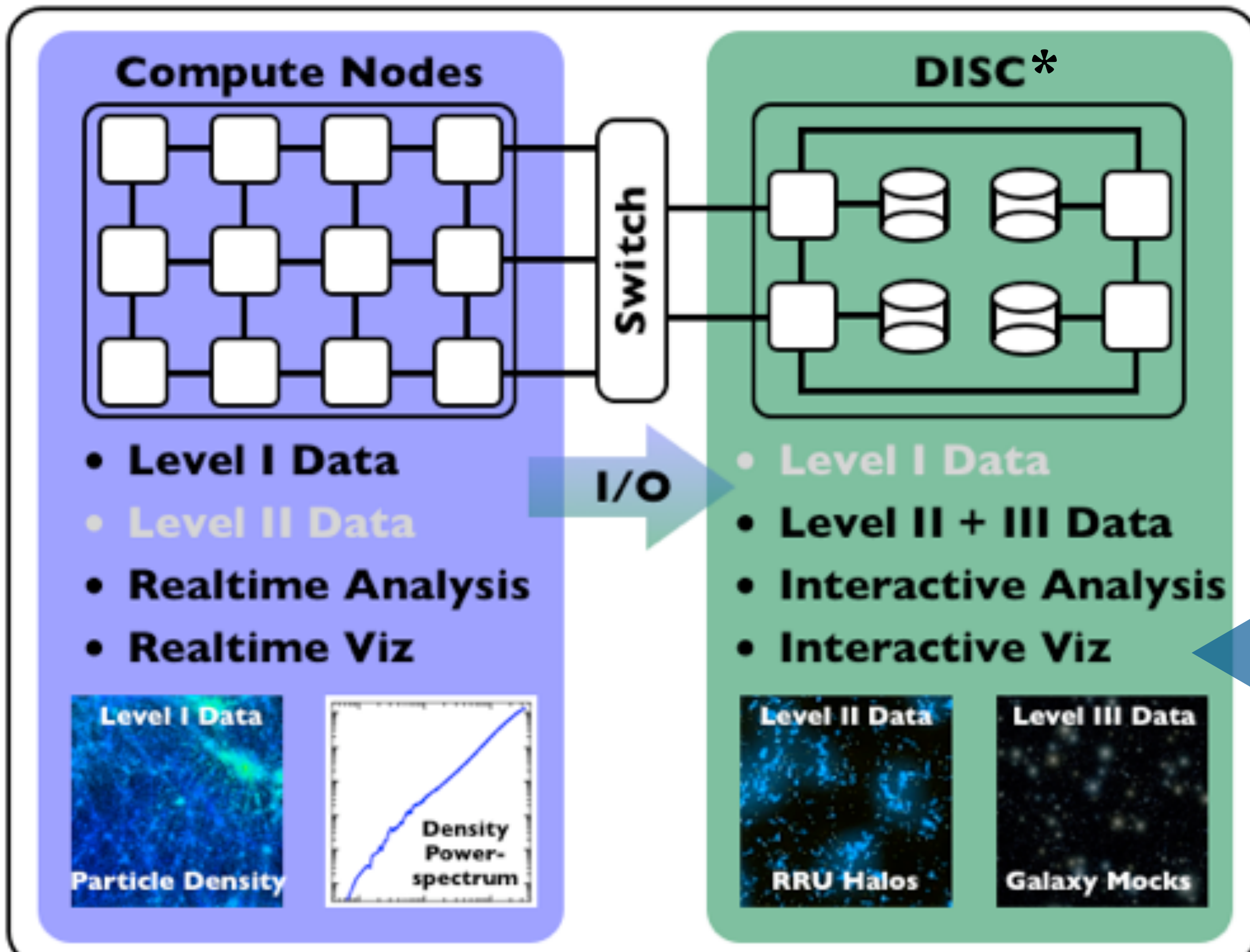




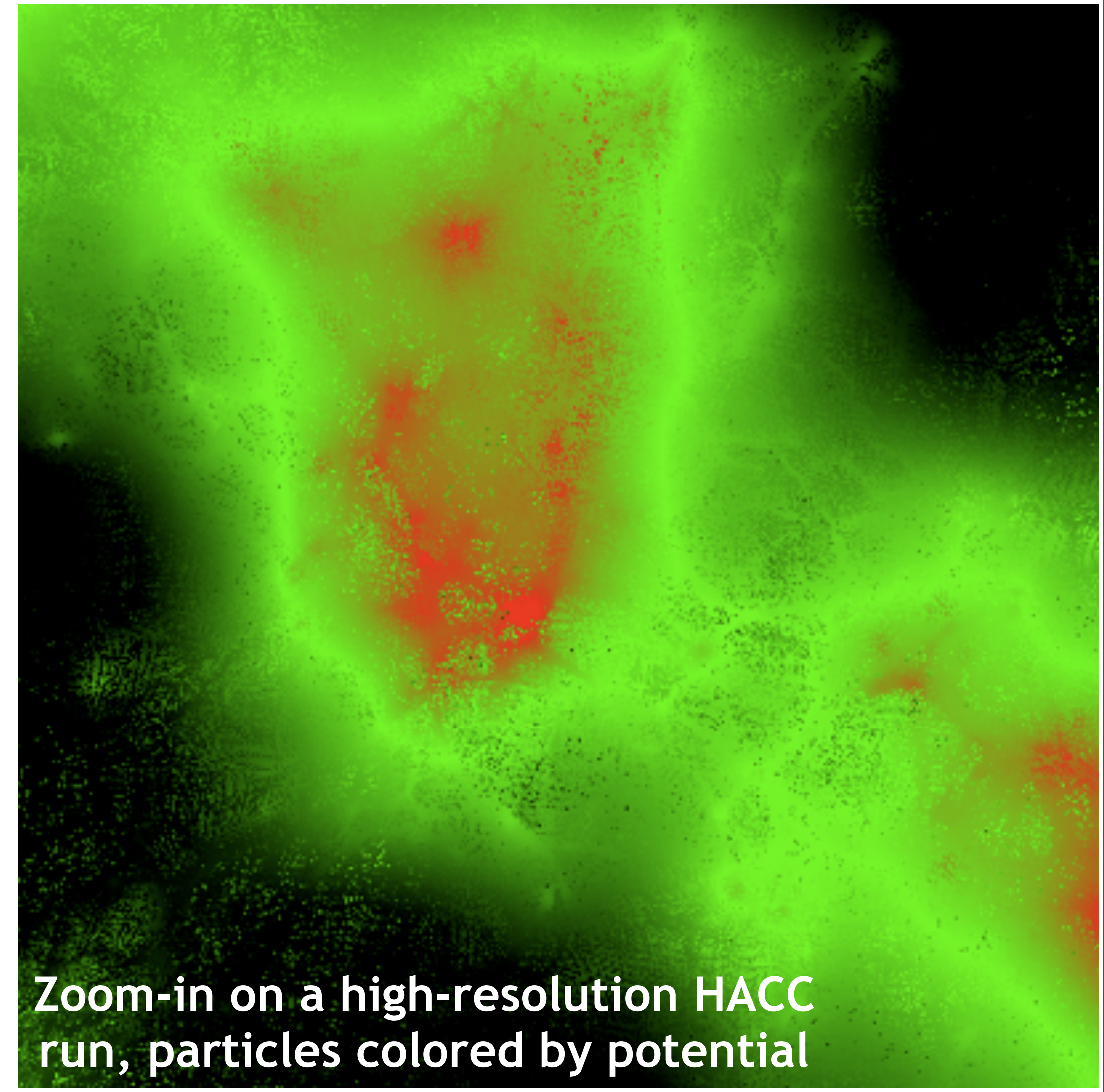
Habib et al. 2009, Pope et al. 2010

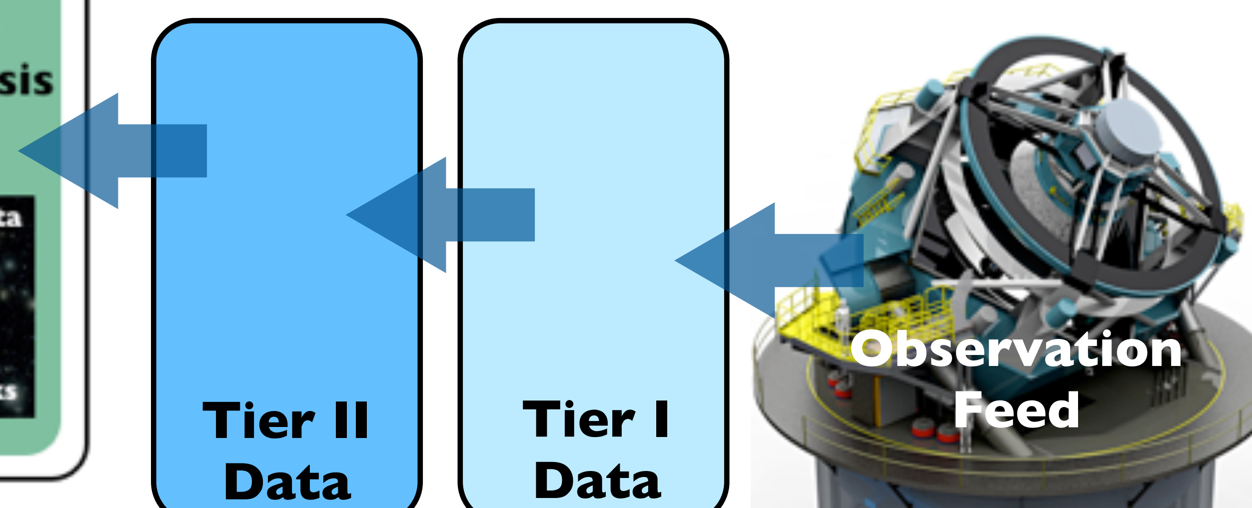
HACC in the HPC/DISC Future

- **HACC as Exascale Co-Design Driver:** Most codes cannot meet future science requirements and HPC constraints, HACC capabilities already demonstrated on Cell and GPU-accelerated systems
- **Synergies with HEP:** Accelerators, LQCD



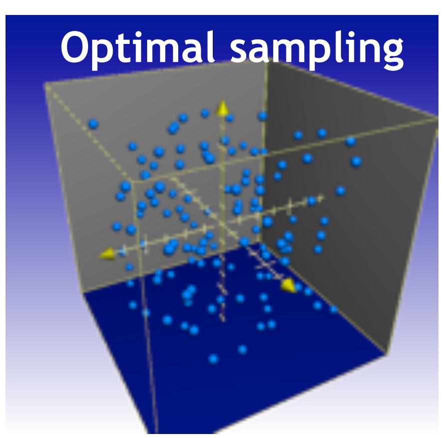
***DISC=Data-Intensive SuperComputer**

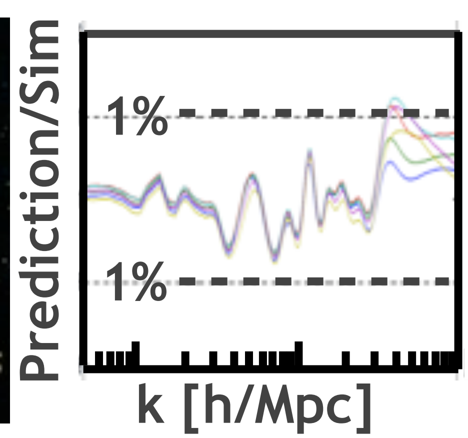

Zoom-in on a high-resolution HACC run, particles colored by potential



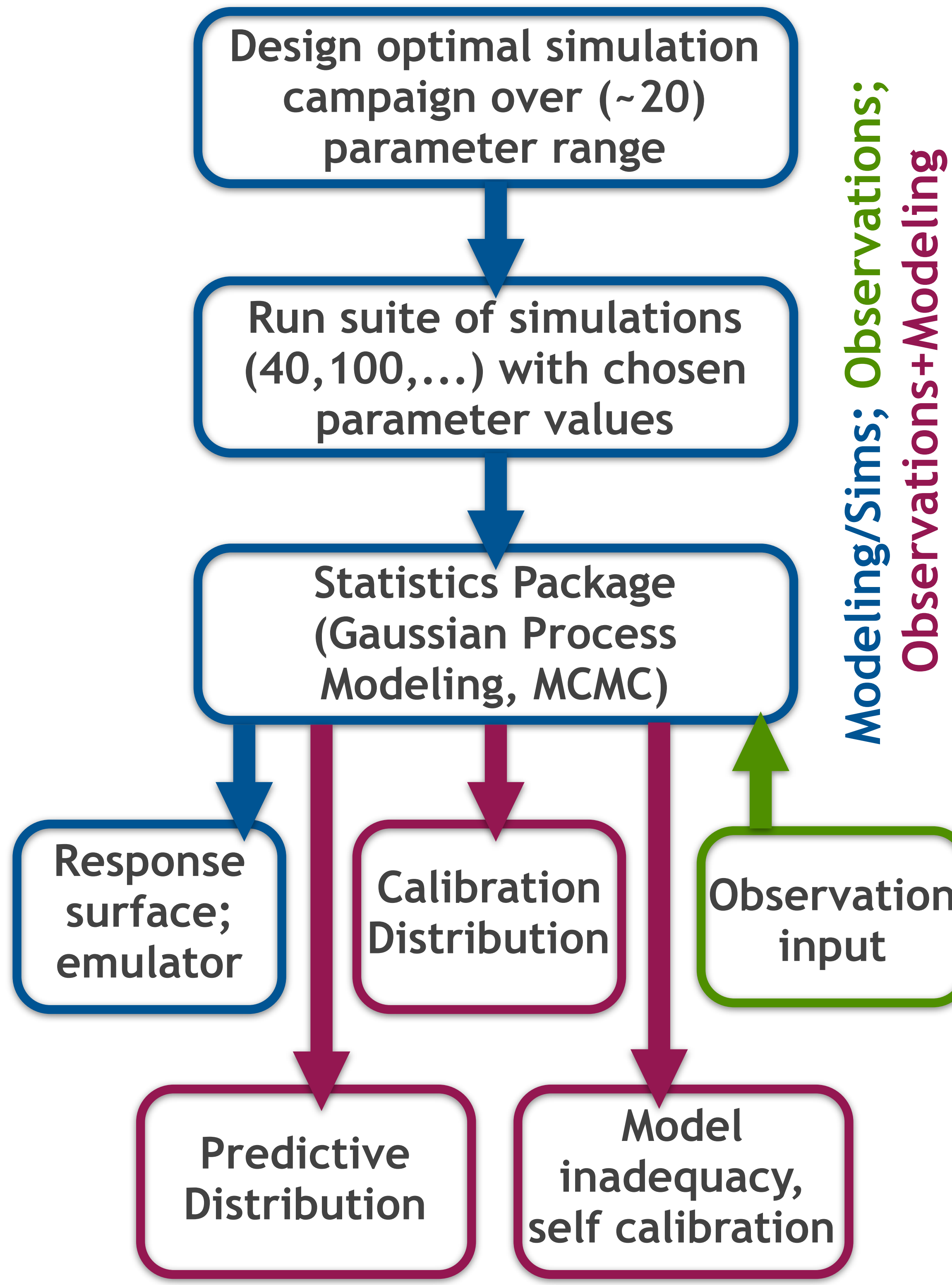
Cosmic Calibration: Solving the Inverse Problem

- **Challenge:** To extract cosmological constraints from observations in the nonlinear regime, need to run Markov Chain Monte Carlo; input: 10,000 - 100,000 different models
- **Brute Force:** Simulations, ~30 years on 2000 processor cluster ---
- **Current Strategy:** Fitting functions, e.g. for P(k), accurate at 10% level, not good enough!
- **Our Solution:** Precision emulators





Heitmann et al. 2006, Habib et al. 2007



Cosmic Emulator in Action

- Instantaneous 'oracle' for nonlinear power spectrum, easy to use, reduces run time from weeks to 'zero', 1% accurate to k~1/Mpc for wCDM cosmologies -- based on ~1000 simulation runs for 38 cosmologies
- For the first time enables direct MCMC with results from full simulations

Cosmic Emulators for Future Surveys

- **Extension Beyond wCDM:**

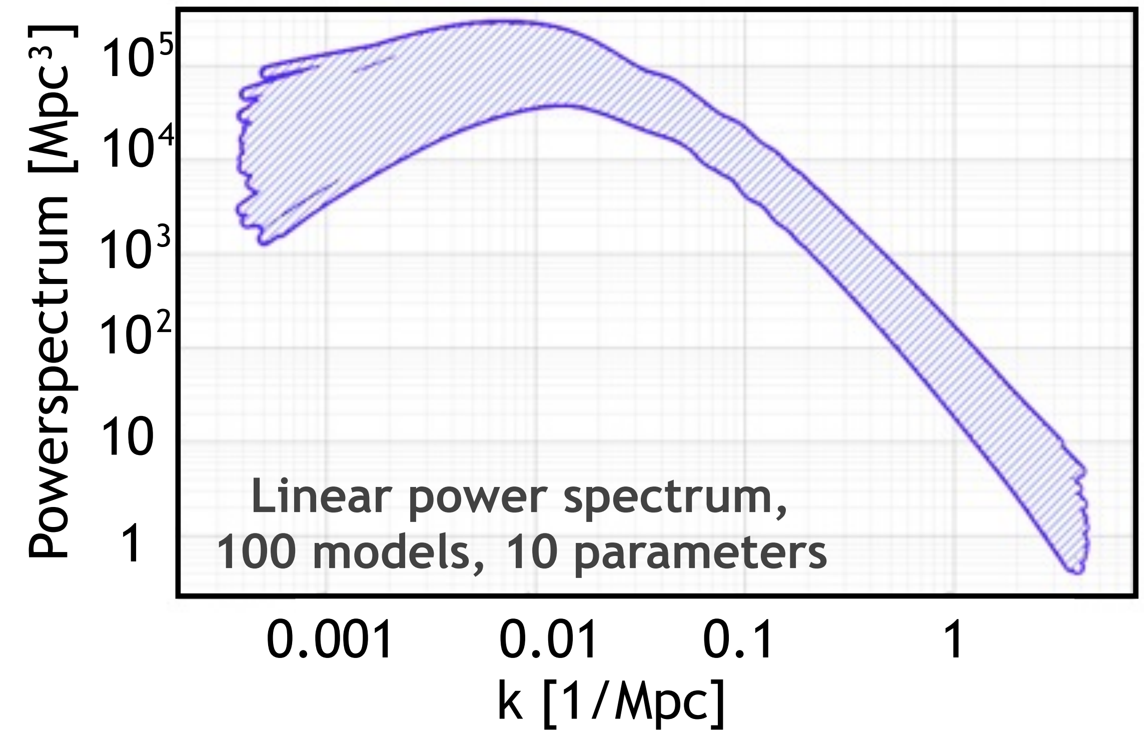
$$\theta = \{\omega_m, \omega_b, n_s, w_0, \sigma_8,$$
$$w_a, dn_s/d\log k, h, \Omega_k, f_\nu\}$$

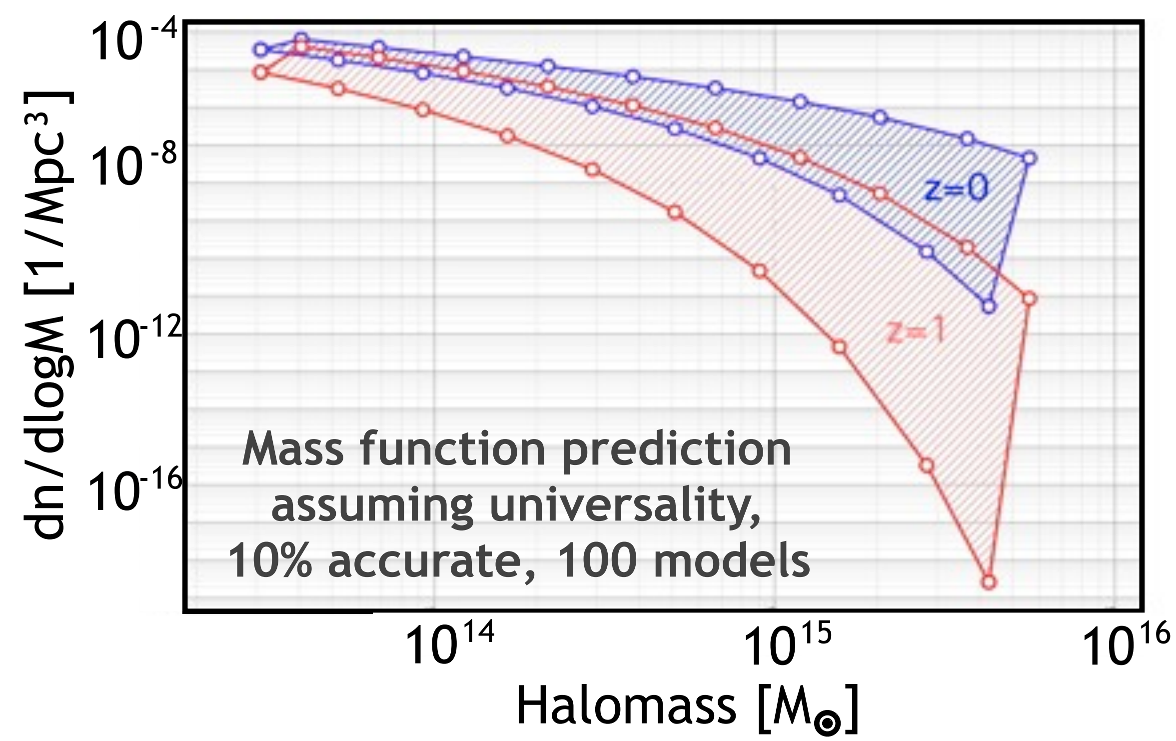
  Currently fine-tuning number of models and parameter ranges with surrogates, input from community

- **Emulators for a Variety of Observables:** power spectrum, mass function (different mass definitions), shear power spectrum, peak statistics, bias, ---

- **Extension of Range of Validity:** Higher resolution, baryonic physics

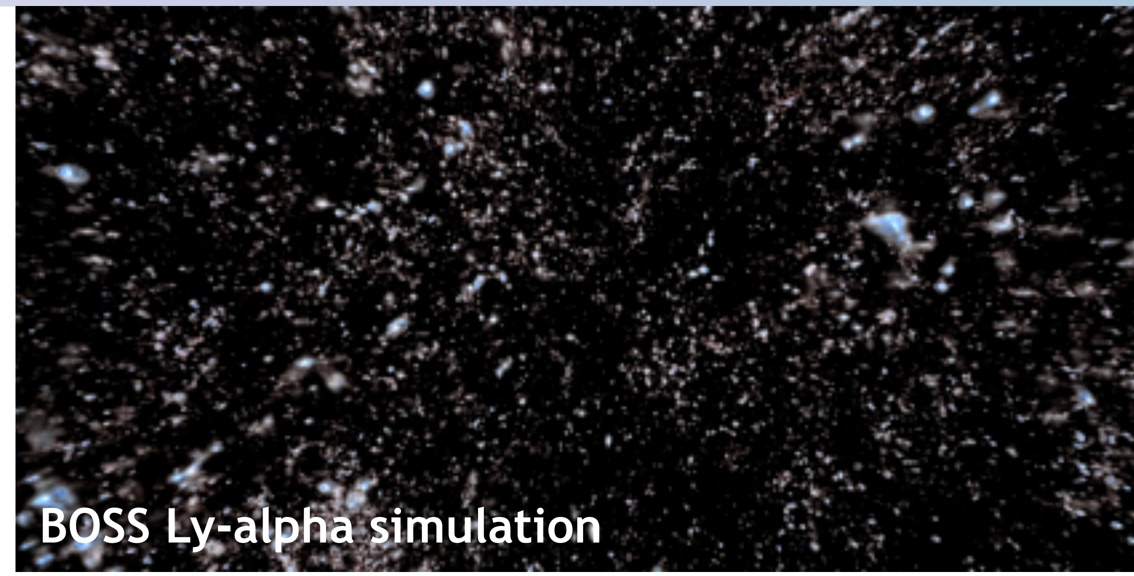
- **Covariances**



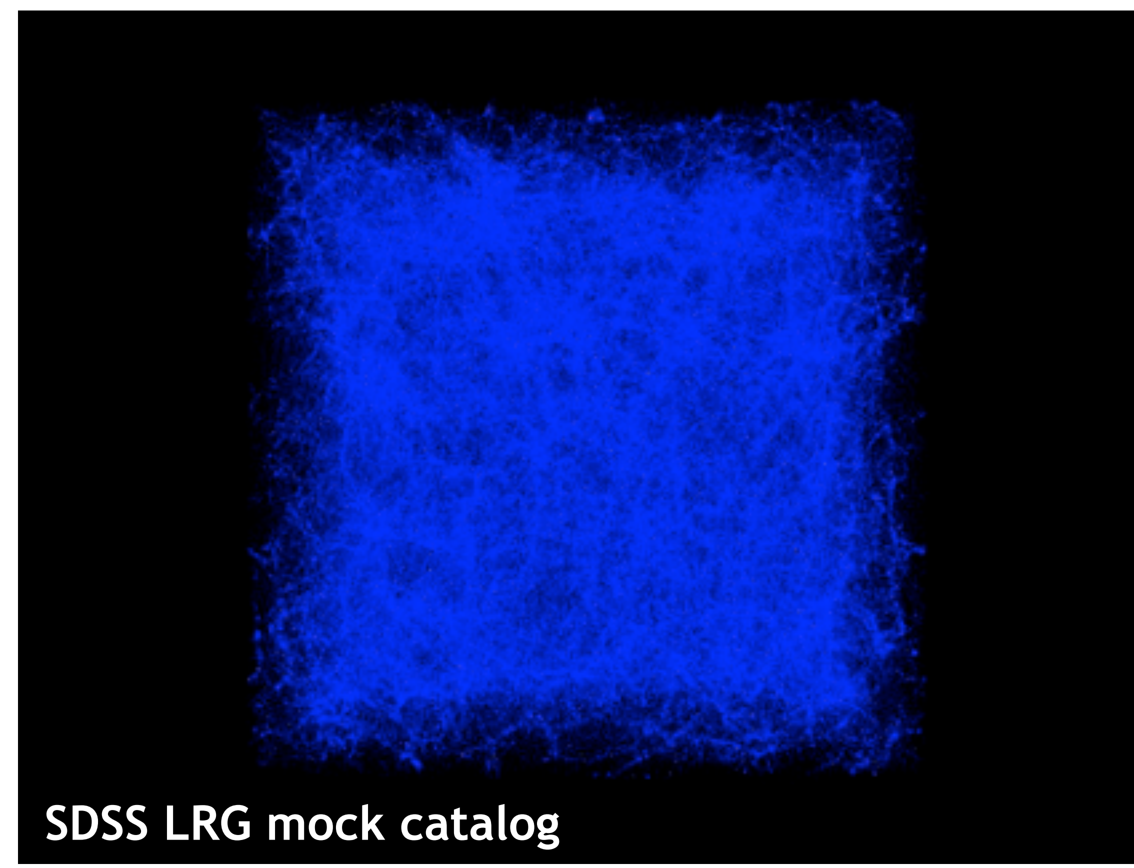


Summary and Outlook

- New ‘Cosmic Frontier’ Theory Group at ANL focused on cosmological probes of physics beyond the Standard Model
- Expertise covers quantum field theory, physical cosmology, high-performance computing, and cosmological surveys
- Strong connections to the Argonne Leadership Computing Facility
- Strong University and DOE Lab collaborations, opportunities for students
- Science-rich, large-scale simulations and analysis tools will be made available to the community



Roadrunner view (halos) of the Universe at z=2 from a 64 billion particle run

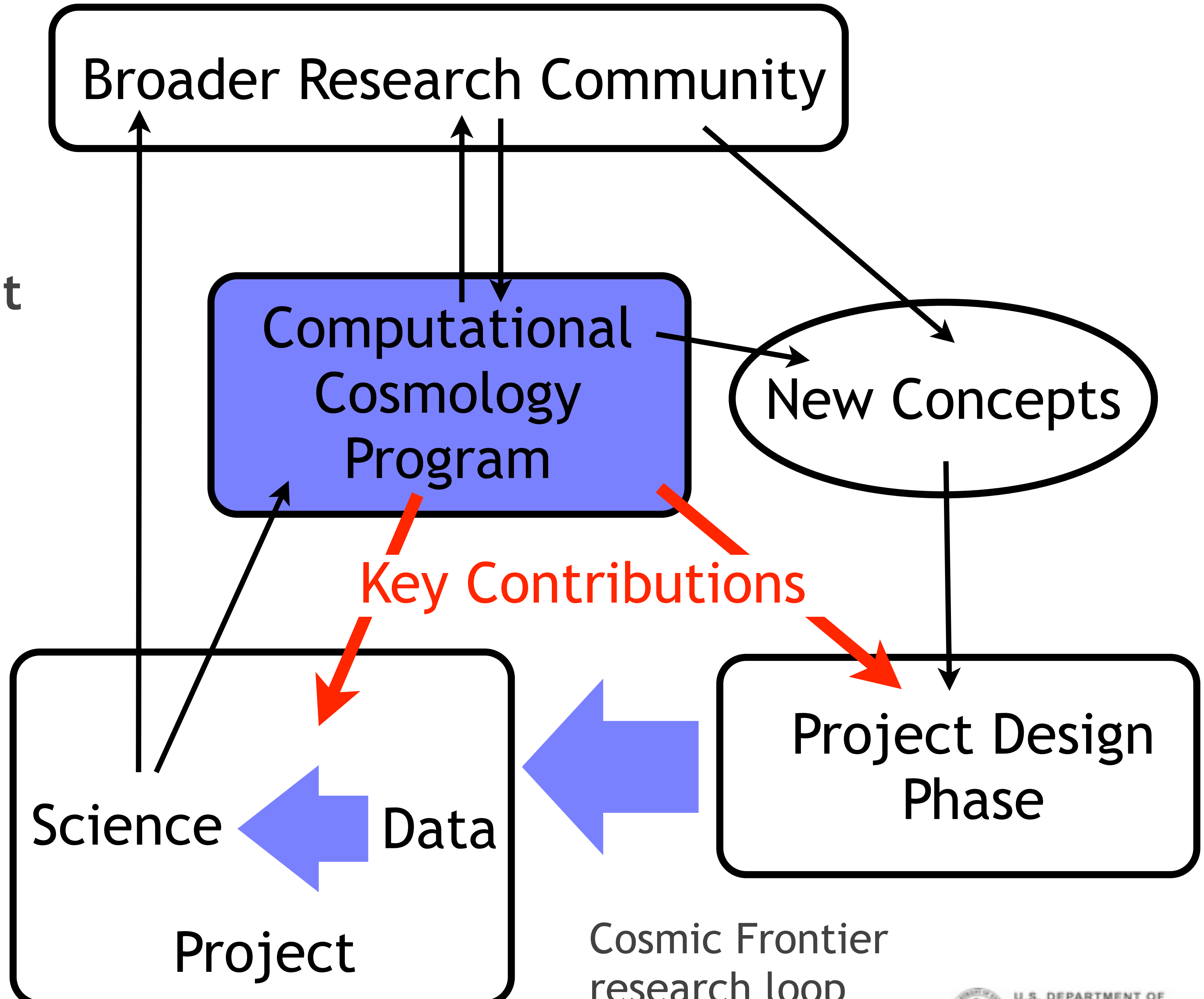


Mock catalog for SDSS luminous red galaxies (orange) and satellite galaxies (green), in coll. with M. White

Back-up Slides

Locating the DOE HEP Computational Cosmology Program

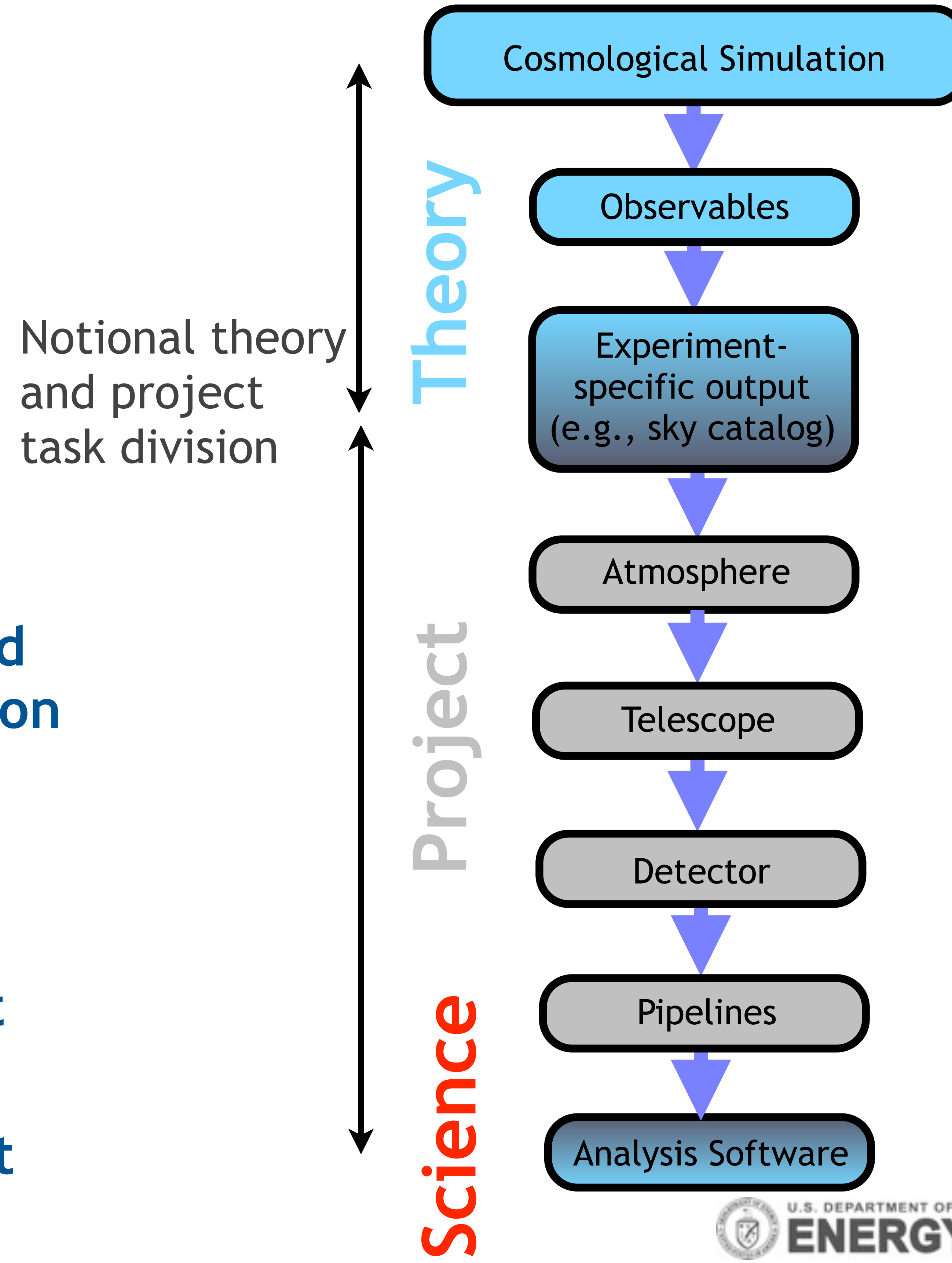
- Resides as a core capability program within DOE HEP
- Contributes to 'discovery space'
- Catalyzes development of concepts into projects
- Plays a key role in project optimization
- Is an essential component of the **'Data to Science'** step for projects
- Functions as a major community resource



Cosmic Frontier research loop

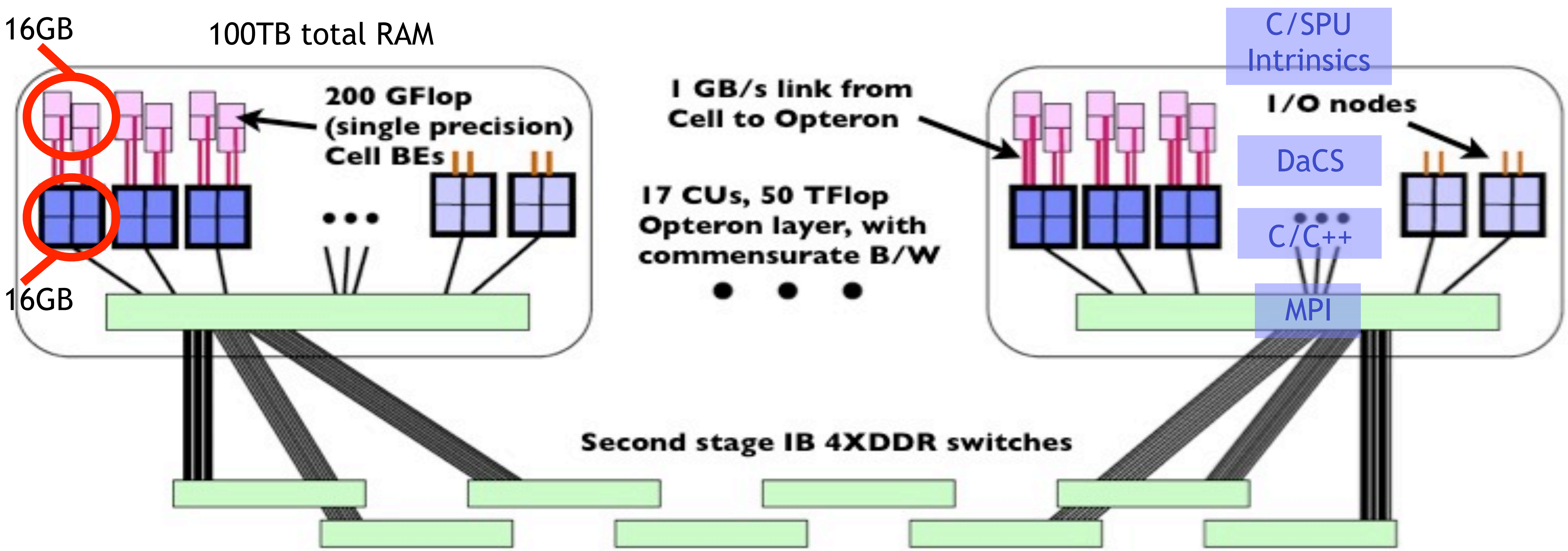
DOE HEP Computational Cosmology Program Advantages

- Key Roles
  - One-point contact for scientists, projects, and programs
  - In-house theory, modeling, and simulation capability
  - Connection to HEP computing
  - Efficient collaboration, ability to work to milestones/time tables
  - Repository of ‘Lessons Learnt’ and ‘Best Practices’ (crucial in precision cosmology)
  - Continuous development paths
  - Develop and maintain simplified ‘detector model’ views of project space (hunt for subtle signals)
  - Connections across projects (joint analyses)



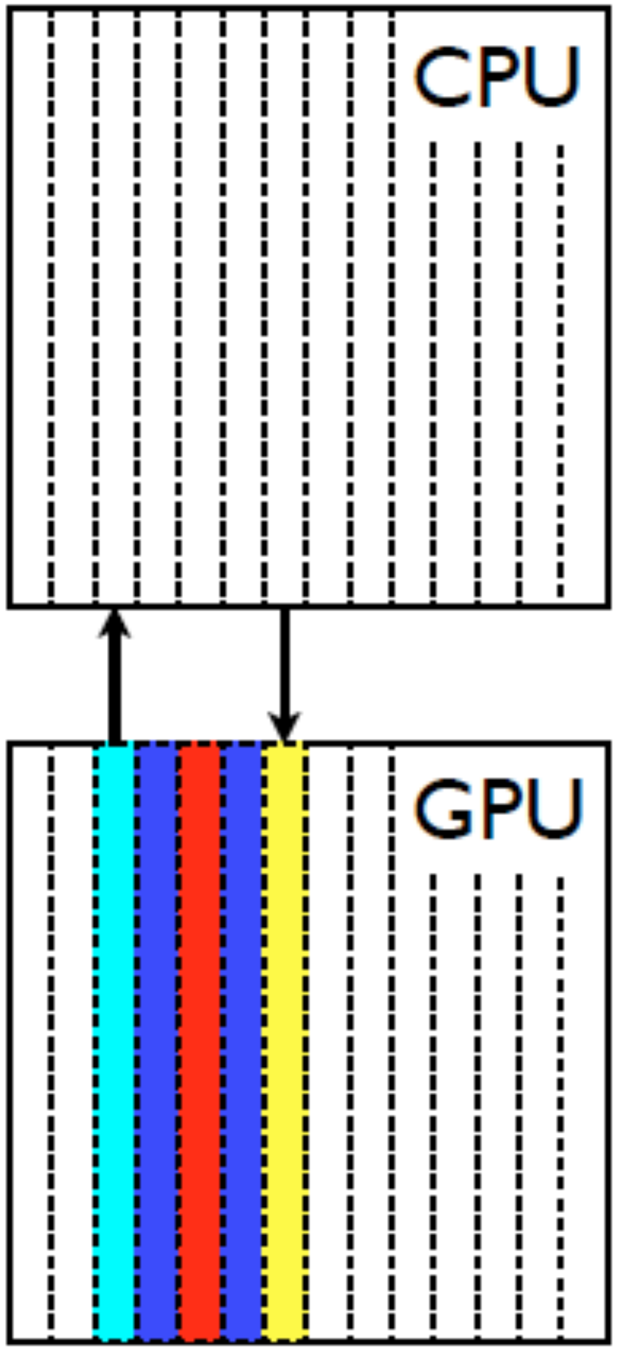
HACC Example I: Roadrunner (CPU+Cell)

- Hybrid machine architecture, out of balance communication (50-100) and performance (20)
- Multi-level programming paradigm
- ‘On the fly’ analysis to reduce I/O
- Prototype for exascale code design problems
- Scalable approach extensible to all next-generation architectures (BG/Q, CPU/GPU, --)



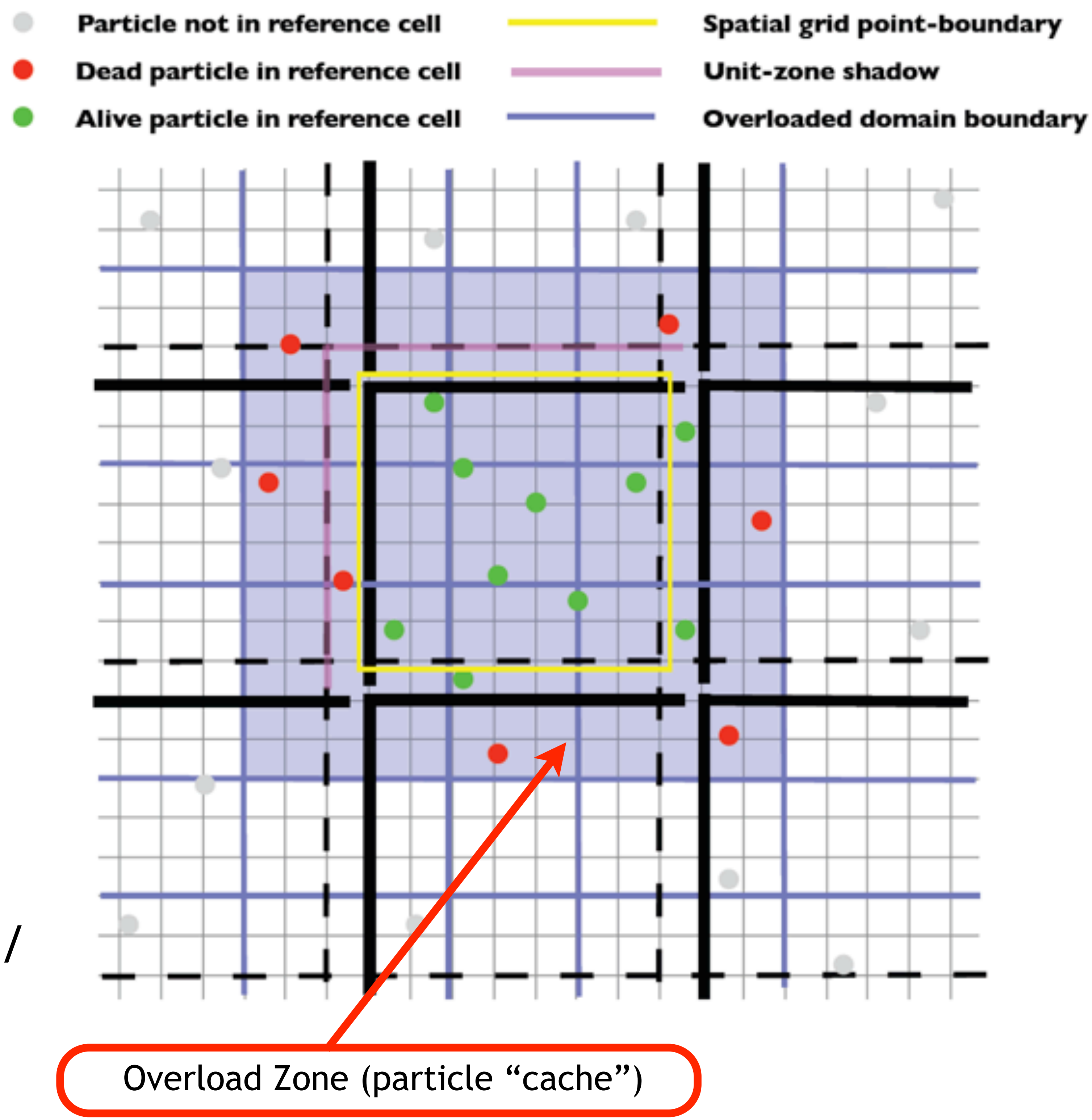
HACC Example 2: CPU+GPU

- 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)