THE FIRST GALAXIES AND QUASARS IN THE BLUETIDES COSMOLOGICAL SIMULATION

Allocation: NSF PRAC/2.625 Mnh
PI: Tiziana Di Matteo[1]
Co-PIs: Yu Feng[1,2], Rupert Croft[1]
Collaborators: Nick Battaglia[3], Steven Wilkins[1,4], Ananth Tenneti[1]

[1]Carnegie Mellon University
[2]University of California, Berkeley
[3]Princeton University
[4]University of Sussex (U.K.)

EXECUTIVE SUMMARY:

Numerical simulations are now the dominant tool in theoretical cosmology. Our team has led the development of codes adapted to petascale supercomputers to study how supermassive black holes and galaxies formed, from the smallest to the rarest and most luminous. Our BlueTides simulation of nearly 1 trillion particles is the largest-ever simulation with full physics (hydrodynamics, star formation, black holes) and targeted the early universe of galaxies and quasars. The simulation agreed well with observations from the Hubble Space Telescope when the universe was only 5% of its current age. It covers a sky area 300 times larger than the largest survey with Hubble, allowing predictions of what Hubble's successors will see. Among the millions of galaxies that formed in the simulation were large disk galaxies as massive as the Milky Way and the first quasars; these objects will be exciting targets for the frontier of observations.

INTRODUCTION

The ability to model the complex physics of structure formation in the universe is increasing, driven by advances in both computer power and software. To continue to support observational and theoretical discovery, we need highly detailed simulations that follow the evolution of the matter, energy, and radiation in model universes from the Big Bang to the present.

In the coming decade, a new generation of astronomical instruments, all in the billion-dollar class, will observe the universe during the period of the first stars and quasars. Petascale computing resources running the newest codes are the theoretical equivalent to these programs and maximize the return on comparable observational and experimental studies. Using our simulation, we hope to make predictions for what should be seen in future observations and surveys as well as to imitate future surveys.

METHODS & RESULTS

The actions of gravity, hydrodynamics, forming stars, black holes, molecular gas, inhomogeneous ionizing radiation, and more were included in the BlueTides cosmological simulation [1]. We carried out a full-machine run on Blue Waters with our newly developed MP-Gadget. The simulation aims to understand the formation of the first quasars and galaxies from the smallest to the rarest and most luminous, and the role of these processes in the reionization of the universe.

Computation

BlueTides followed the evolution of 0.7 trillion particles in a large volume of the universe (600 co-moving Mpc on a side) over the first billion years of the universe's evolution with a dynamic range of 6 orders of magnitude in space and 12 in mass. This makes BlueTides by far the largest cosmological hydrodynamic simulation ever run.

We improved the cosmological code Gadget, now MP-Gadget, to make it suitable for the petascale and beyond:

- Threading efficiency: We replaced global critical sections with per-particle (per-node) spin locks. The code is now scalable to any number of threads and allows us to use fewer domains, which reduces the complexity of domain decomposition and inter-domain communication.
- Mesh gravity solver [2]: Domain decomposition was an issue with the huge Fourier transforms (up to $16{,}384^3$) needed to carry out large-scale mesh computation of the gravitational force. We improved the speed by a factor of 10 by switching from a slab decomposition to a "pencil" decomposition, which allows work to be distributed among 20,000+ XE6 nodes on Blue Waters.
- MP-sort [3]: We implemented and published with Blue Waters staff a new sorting algorithm that exchanges each data item exactly once (O(N) communication). It is only limited by the network bandwidth [4].

We also improved the physical modelling in MP-Gadget:

- A new pressure–entropy SPH (Smoothed Particle Hydrodynamics) formulation [9] replaced the old density–entropy formulation, which suppressed phases mixing in a non-physical way.
- A patchy reionization model from [10] introduced a UV field based on a predicted time of reionization at different spatial locations in the simulation.
- Improved stellar feedback models to include mass/halo dependent supernova wind efficiency (also metal cooling and H2 mass fraction) [11].

Science Results

The core scientific data set of BlueTides summed to 2 PB, consisting of 20 snapshots (47 TB each) of the properties of mass elements used in the simulation. We performed on-the-fly and offline data analysis to extract scientific results from this data set. Recent work included predictions about the first galaxies, first quasars, and their contribution to reionization.

First galaxies: We applied observational selection algorithms (SourceExtractor) to the simulated sky maps and created catalogs of millions of galaxies. This uncovered a striking and unexpected population of large Milky-Way-sized disk galaxies present when the universe was 5% of its present age [5].

First quasars: We computed statistical measures of the population of galaxies and quasars at different redshifts, which were consistent with early data from the Hubble Space Telescope and made predictions for upcoming surveys. We also can make predictions for the massive black holes that assemble at these early epochs and have masses 100 million times that of the Sun [7].

Reionization from galaxies and quasars: We found that the luminosity functions of the first galaxies and quasars indicate that a high escape fraction is required if galaxies dominate the ionizing photon budget during reionization. With smaller galaxy escape fractions, a large contribution from the faint quasars is feasible [8].

WHY BLUE WATERS?

Previous calculations have either used a small simulation volume or large particle masses, both of which resolve only large galaxies. BlueTides used more than 20,000 nodes and about 90% of the memory on Blue Waters and included both a large simulation volume and small particle masses to capture the rarest and brightest objects, the first quasars.

Blue Waters staff was critical to our success. They helped us run our full-system job in a timely fashion, helped with MPI+OpenMP development of the Gadget code, and assisted with fine-tuning file handling using HDF and Lustre. They also aided development of some radical new ways to process petabytes of simulation output.

Our simulations blaze a trail for future calculations. The next Track-1 system will make detailed 14-million-year simulations (early universe to today) possible. This would facilitate detailed comparisons between our simulated galaxy and quasar population and predictions that are relevant to the large-scale upcoming observational projects.

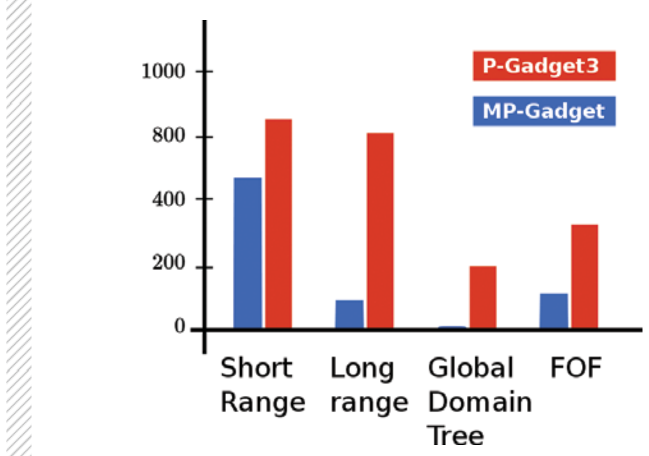

FIGURE 1: Comparing the wall time in a single step in MP-Gadget and P-Gadget in the BlueTides simulation. Even with more physics modules, the short-range force in MP-Gadget still spends less time per step than P-Gadget3. The long range force module in MP-Gadget out-performs P-Gadgets by a factor of 8. The global domain builder a factor of 10 and the galaxy catalogue (FOF) builder a factor of 2.

PUBLICATIONS

Feng, Y., et al., The formation of Milky Way-mass disk galaxies in the first 500 million years of a cold dark matter universe. *Astrophys. J. Lett.*, 808:1 (2015), L17, doi:10.1088/2041-8205/808/1/L17.

Croft, R., T. Di Matteo, N. Khandai, and Y. Feng, Petascale cosmology: simulation of structure formation. *Comput. Sci. Eng.*, 17:2 (2015), pp. 40–46, doi:10.1109/MCSE.2015.5.

Feng, Y., M. Straka, R. Croft, and T. Di Matteo, Sorting At Scale on Blue Waters. *Cray User Group Conf. 2015*, Chicago, Ill., April 26–30, 2015. [Best Paper finalist]