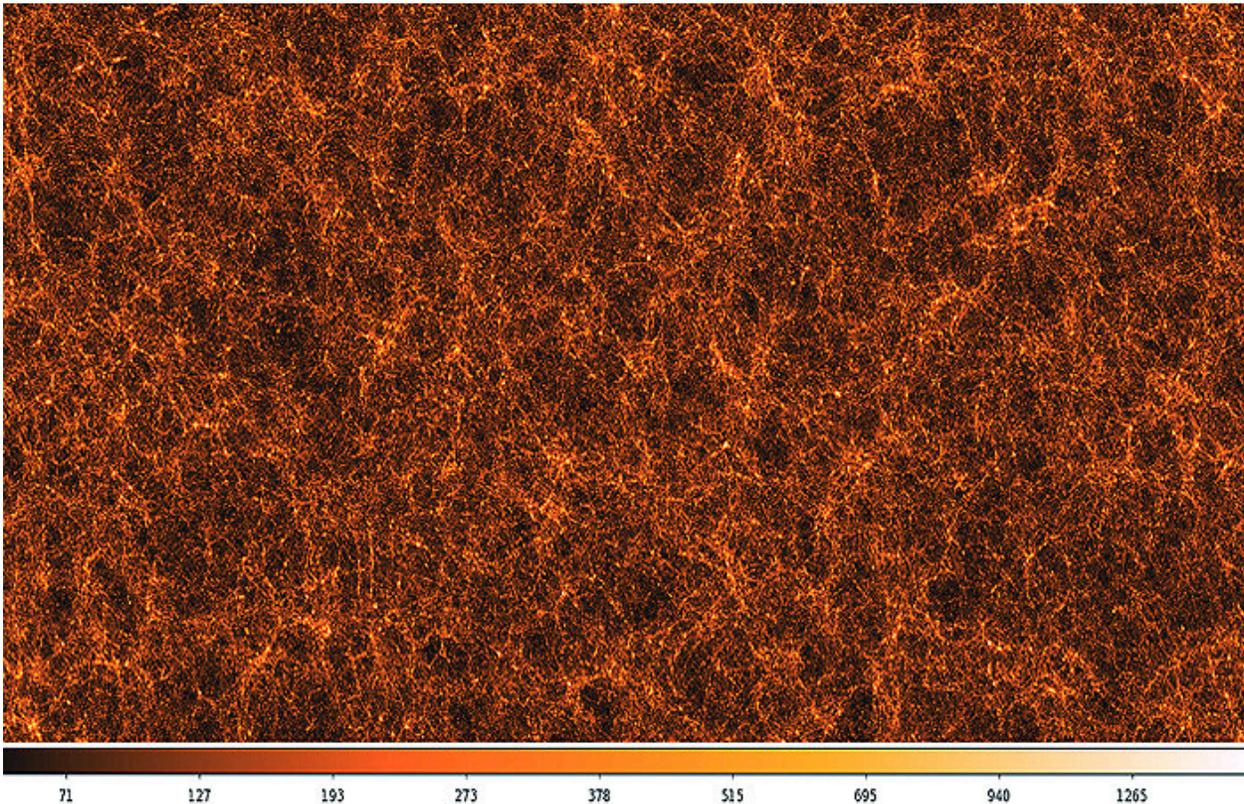

First trillion particle cosmological simulation completed

January 8, 2015



Cosmological simulations are the cornerstones of theoretical analyses of structure in the Universe from scales of kiloparsecs to gigaparsecs. A team of astrophysicists and computer scientists, including Los Alamos National Laboratory researchers, has created high-resolution cyber images of our cosmos. The scientists completed the first-ever complete trillion-particle cosmological simulation and have made an initial 55 Tbyte (trillion bytes) public data release. A [paper](#) describes the research and data release.

Significance of the research

The Dark Sky Simulations are an ongoing series of cosmological simulations designed to provide a quantitative and accessible model of the evolution of the large-scale Universe. Simulations are essential for many aspects of the study of dark matter and dark energy, because scientists lack a sufficiently accurate analytic model of non-linear gravitational clustering. The simulations form a major pillar of our understanding of

the cosmological standard model; they are an essential link in the chain that connects particle physics to cosmology, and similarly between the first few seconds of the Universe to its current state. Modeling must keep advancing for researchers to interpret what they see in the skies. Current and upcoming optical surveys need support from numerical simulations to guide observational campaigns and interpret their results. During the next few years, projects such as Pan-STARRS, the South Pole Telescope, and the Dark Energy Survey will measure the spatial distribution of large-scale structure in enormous volumes of space across billions of years of cosmic evolution. The rigorous statistical and systematic demands of upcoming surveys requires high quality simulations to further researchers' understanding of cosmological theory and the large scale structure of the Universe. The trillion-particle simulation in the current release is state of the art in terms of mass resolution for its cosmological volume and offers insight into the largest scales and structures in the Universe.

This is the first time this type of information is being made available for everyone, not just scientists. A primary goal of this project is to adopt some of the fundamental concepts of the open source community, and translate them to open data for state-of-the-art cosmological simulations. The team aims to decrease the barrier to entry for accessing and analyzing data, and increase the speed of iteration and pace of scientific inquiry. The scientists created an interface to the large datasets that is novel, simple, and extensible to allow people with a wide range of technical abilities and interests to explore the data.

In principle, high school students interested in physics and/or computation should be capable of accessing subsets of a trillion particle dataset and studying the structure of the dark matter potential in a sample galaxy cluster. Researchers in large scale data visualization should be able to load halo catalogs into 3-D models of the Universe and explore alternate representation methods for high-dimensional datasets. The simulations are available in coLaboratory, a new tool created by Google and Project Jupyter, which allows multiple people to analyze the data simultaneously.

Research achievements

The team performed the calculations with 2HOT, a purely tree-based adaptive N-body method developed at LANL, running on 200,000 processors of the Titan supercomputer at Oak Ridge National Laboratory. The team received time on the Titan supercomputer through the DOE's Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program for leadership computing. They used two-thirds of Titan for 33 hours to direct a trillion virtual particles to follow the laws of gravity in a simulated universe over the past 13.7 billion years. Three months later, the team provided the 55 Tbyte data release with an overview of the derived halo catalogs, mass function, power spectra, and light cone data.

The team created an interface to datasets that is based on an abstraction of the World Wide Web (WWW) as a file system, remote memory-mapped file access semantics, and a space-filling curve index. This provides a means to not only query stored results such as halo catalogs, but also to design and deploy new analysis techniques on large distributed datasets.

The research team

Researchers include Michael S. Warren and Alexander Friedland of LANL's Nuclear and Particle Physics, Astrophysics and Cosmology group, Daniel Holz of University of Chicago and collaborators from Stanford University, Paris Institute of Astrophysics,

Ohio State University and the National Center for Computing Applications, University of Illinois.

Laboratory Directed Research and Development (LDRD) and the DOE Office of Science, Office of High Energy Physics supported different aspects of the research. The Laboratory's Institutional Computing Program provided the computing resources used for the initial development and simulations for the project. The work supports the Lab's Information, Science, and Technology and Nuclear and Particle Futures science pillars.

Los Alamos National Laboratory

www.lanl.gov

(505) 667-7000

Los Alamos, NM

Operated by Los Alamos National Security, LLC for the Department of Energy's NNSA

