

HPC Profiles in Leadership

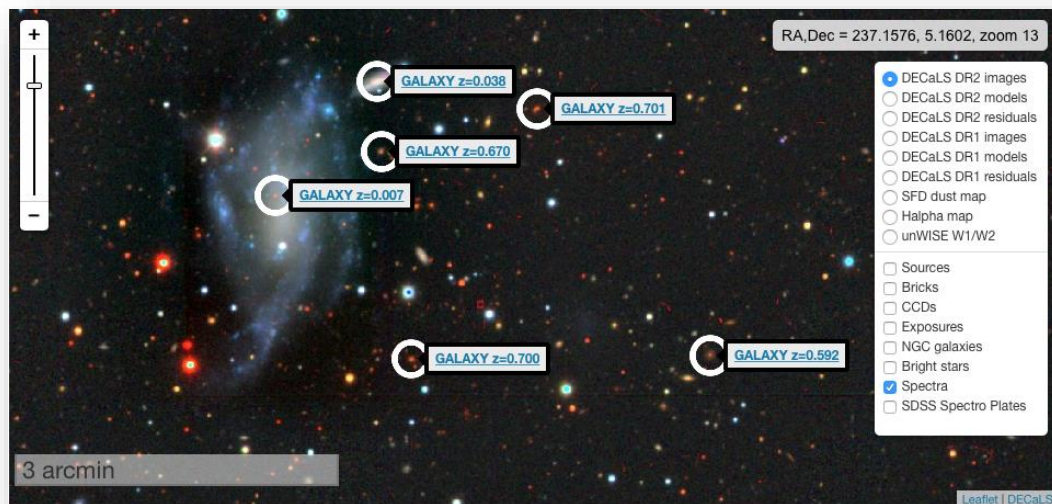
Celeste Project Seeks New Ways of Cataloging the Universe

Earl Joseph, Steve Conway, and Bob Sorensen

October 2017

HYPERION RESEARCH OPINION

Celeste, a hierarchical model designed to catalog stars, galaxies and other light sources in the universe, takes the concept of implementing statistical inference to build a fully generative model to mathematically locate and characterize light sources in the sky. Researchers from UC Berkeley, Intel, the National Energy Research Scientific Computing Center (NERSC), Lawrence Berkeley National Laboratory, Julia Computing and JuliaLabs@MIT developed a parallel version of the Celeste code, an innovation that leverages 8,192 Intel Xeon processors in Berkeley Lab's new Cori supercomputer and Julia, the high-performance, open-source scientific computing language. Lessons learned in this project about inferencing will be applicable to many advanced analytics problems.



Note: this page is intentionally blank.

SITUATION OVERVIEW

A Berkeley Lab-based research collaboration of astrophysicists, statisticians, and computer scientists is looking to shake things up with Celeste, a new statistical analysis model designed to enhance one of modern astronomy's most time-tested tools: sky surveys. A central component of an astronomer's daily activities, surveys are used to map and catalog regions of the sky, fuel statistical studies of large numbers of objects, and enable interesting or rare objects to be studied in greater detail. But the traditional ways image datasets from these surveys are analyzed today remains complex and time consuming.

The Celeste code, developed under Berkeley Lab's MANTISSA project, implements statistical inference to build a fully generative model to mathematically locate and characterize light sources in the sky. When it was first released in 2015, Celeste was limited to single-node execution on at most hundreds of megabytes of astronomical images. In the case of the Sloan Digital Sky Survey, which is the dataset used for this research, this analysis was conducted by identifying points of light in nearly 5 million images of approximately 12 megabytes each - a dataset of 55 terabytes. For this project, researchers from UC Berkeley, Intel, the National Energy Research Scientific Computing Center (NERSC), Lawrence Berkeley National Laboratory, Julia Computing and JuliaLabs@MIT developed a parallel version of the Celeste code, an innovation that leverages 8,192 Intel Xeon processors in Berkeley Lab's new Cori supercomputer and Julia, the high-performance, open-source scientific computing language.

WHY THIS HPC INNOVATION IS IMPORTANT

This is one of the largest graphical model problems in science that requires a supercomputing platform for running the inference algorithms. The biggest challenge was the amount of conjecture involved because the star has not been directly observed. This is where statistical inference comes in. The task starts from the pixel intensities in the images and works backwards to where the light sources and their characteristics are located.

About Hyperion Research, LLC

Hyperion Research, consisting of the former IDC high performance computing (HPC) analyst team, provides HPC information, analysis, and recommendations based on technology and market trends. Research includes market sizing and forecasting, share tracking, segmentation, technology and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). We provide thought leadership and practical guidance for users, vendors and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue

St. Paul, MN 55102

USA

612.812.5798

www.hpcuserforum.com and www.HPCatHyperion.com

Copyright Notice

Copyright 2017 Hyperion Research LLC. Reproduction is forbidden unless authorized. All rights reserved. Visit www.HPCatHyperion.com to learn more. Please contact 612.812.5798 and/or email ejoseph@hyperionres.com for information on reprints, additional copies, web rights, or quoting permission.