Supercomputing in the Age of Discovering Superearths, Earths and Exoplanet Systems

Jon M. Jenkins
NASA Ames Research Center

Wednesday September 28, 2015

Ad Hoc Big Data Task Force of the NASA Advisory Council Science Committee
All the Known Planets In 1994
A More Recent Pictures of Planets

All 786 Known Planets

*3,338 planets as of 9/28/16
ALL 786 KNOWN PLANETS

TO SCALE

(AS OF JUNE 2012)

(SOME PLANET SIZES ESTIMATED BASED ON MAPS.)

THIS IS OUR SOLAR SYSTEM.

THE REST OF THESE ORBIT OTHER STARS
AND WERE ONLY DISCOVERED RECENTLY.

MOST OF THEM ARE HUGE, BECAUSE
THOSE ARE THE KIND WE LEARNED TO
DETECT FIRST, BUT NOW WE'RE FINDING THAT
SMALL ONES ARE ACTUALLY MORE COMMON.
WE KNOW NOTHING ABOUT WHAT'S ON ANY OF THEM.
Exoplanet Discoveries Over Time

Radii estimated for non-transiting exoplanets
Discovery data dithered randomly within discovery year
Exoplanet Discoveries

A Search for Earth-size Planets

2015

Planet Radius, Earth Units vs. Orbital Period, Days

- Radial Velocity
- Non-Kepler Transits
- Imaging
- Microlensing
- Kepler
Enabling Kepler

- Back illuminated CCDs (20 ppm photometric precision)
- Sophisticated algorithms
- Computational infrastructure
How Does Kepler Work?
How Hard is it to Find Good Planets?

Jupiter (~1%)

Earth (~0.01%)

Kepler Candidate KOI-351

Kepler Planet – Kepler-20e
Kepler’s Science Pipeline

- Artificial Transit & BEB Injection Machine
- CAL: Pixel level Calibration
- Raw Data → CAL → Calibrated Pixels
- PA: Photometric Analysis
- Artificial Transits and Eclipses
- Raw Light Curves & Centroids → PA → Calibrated Pixels
- PDC: Pre-search Data Conditioning
- Threshold Crossing Events (TCEs)
- DV: Data Validation
- TPS: Transiting Planet Search
- Planet, or dud?
- Auto-Vetting: Applying machine learning to candidate evaluation
- TCERT: Threshold Crossing Event Review Team + Robovetting

>1,000,000 Lines of Code; 26 different Modules
The Search Problem
The Search Problem
The Search Problem
Keeping Up with the Data
A Search for Earth-size Planets

Hardware Architecture: Kepler Science Operations Center

- 64 hosts, 712 CPUs,
- 3.7 TB of RAM,
- ~300 TB of raw disk storage
Hardware Architecture: NAS Pleiades Supercomputer

- 7.25 Pflop/s peak cluster
- 246,048 cores
- 938 TB of memory
- 29 PB of storage

Transiting Planet Search Running on Pleiades
A Search for Earth-size Planets

Processing Kepler Data on the NAS Pleiades

Processing scales from 100s of cores on local cluster to 10s of 1000s of cores on the NAS
A Search for Earth-size Planets

Kepler-452b

ARTISTIC CONCEPT
Light Curve
A Search for Earth-size Planets

Transit-like signals can be produced by a number of astrophysical phenomena:
- Background Eclipsing Binaries
- Triple star systems with an EB/planet
- Background/Foreground planet

BLENDER can assess statistical confidence in planetary nature of a candidate.

Computationally intensive: Supercomputer essential.
Blender Analysis for Kepler-452b

BEB odds: $1.21 \times 10^{-12}$
BP odds: $2.56 \times 10^{-10}$
HTP odds: $2.35 \times 10^{-6}$

Vs: (Expected) Planet odds: $9.97 \times 10^{-4}$

Therefore, odds ratio is $\sim 424:1$
A Search for Earth-size Planets
A Search for Earth-size Planets

Kepler Small Habitable Zone Planets Now Include One Orbiting a Sun-Like Star

[Diagram showing the energy received by planets as a function of surface temperature of the star.]
Searching for Exomoons

David Kipping and team have been searching for exomoons in ~400 light curves from Kepler on the NAS Pleiades supercomputer.

Each search consumes 50,000 CPU hours.

~40 light curves were searched as of 2014.

~300 were searched in 2015.

Exomoons remain elusive: None have been conclusively discovered.
TESS Elation!
• All sky transit survey to find Earth’s closest cousins

• 2 year primary mission

• Launch in December 2017 (tentative)

• TESS will identify best planets for follow up and characterization with James Webb and very large telescopes
• Processing TESS data on the NAS
Supercomputing has played an increasingly important role in exoplanet searches, validation and characterization.

The Kepler and TESS missions were and are not achievable without supercomputing.

The role of supercomputers in exoplanet science is sure to grow in the future as the amount of data and sophistication of the software continue to increase with future missions.