March 4, 2014

Extended Collaborative Support Services

John Cazes
cazes@tacc.utexas.edu

XSEDE
Extreme Science and Engineering Discovery Environment

NSF

XSEDE offers huge variety of resources

- Leading-edge distributed memory systems
- Very large shared memory systems
- High throughput systems
- Visualization engines
- Accelerators and co-processors

Many scientific problems have components that call for use of more than one architecture.

Extended Collaborative Support Services Strategic Objective -To Help Users Make More Productive Use of XSEDE Resources

- Support people who understand the discipline as well as the systems (perhaps more than one support person working with a project).
- 37 FTEs, spread over ~80 people at almost a dozen sites.

How do you get Extended Collaborative Support?

- You have to ask for it -it's an allocated resource.
- You can always ask for it, i.e. midstream or even as part of a startup request.
- Lasts up to a year.
- Must have specific goal in mind- can't just say I want additional programming support.

ECSS has 5 support services

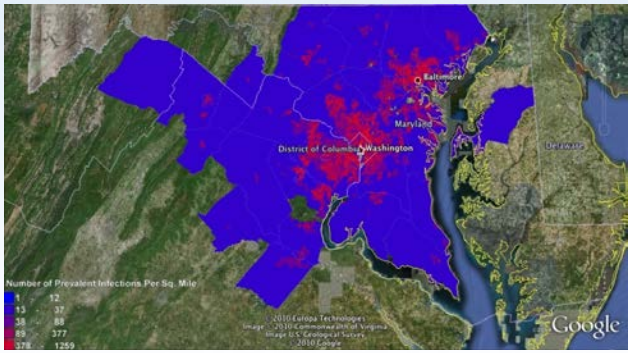
- Support for
  - Research Teams
  - Novel and Innovative Projects
  - Community Capabilities
  - Gateways
  - Training and Outreach

Support for Research Teams
(led by Mark Fahey, NICS (U. of Tennessee))

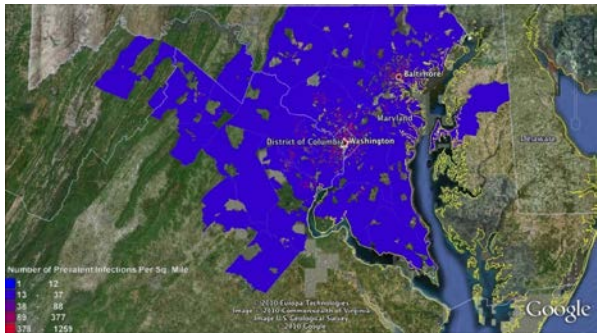
- Optimization
  - Profiling
  - Scaling to higher core count
  - Improving IO
  - Porting to GPUs
  - Finding better solvers (what's better often depends on the degree of parallelization)
- Visualization
- Workflows

FRED- Epidemiology Modeling Code

- Agent-based modeling system that uses synthetic populations derived from census data to capture geographic and demographic distributions.
- Recently used in real time by the Dept. of Health and Human Services in managing the H1N1 outbreak
- Usually applied to a region. When scaled to the entire US, the serial implementation required approximately 96 hours and 540 GB of memory to complete a 100-day simulation.
- An ECSS consultant reduced this to 3-4 hours using less than 200 GB of memory.



15% serologic attack rate, no mitigation



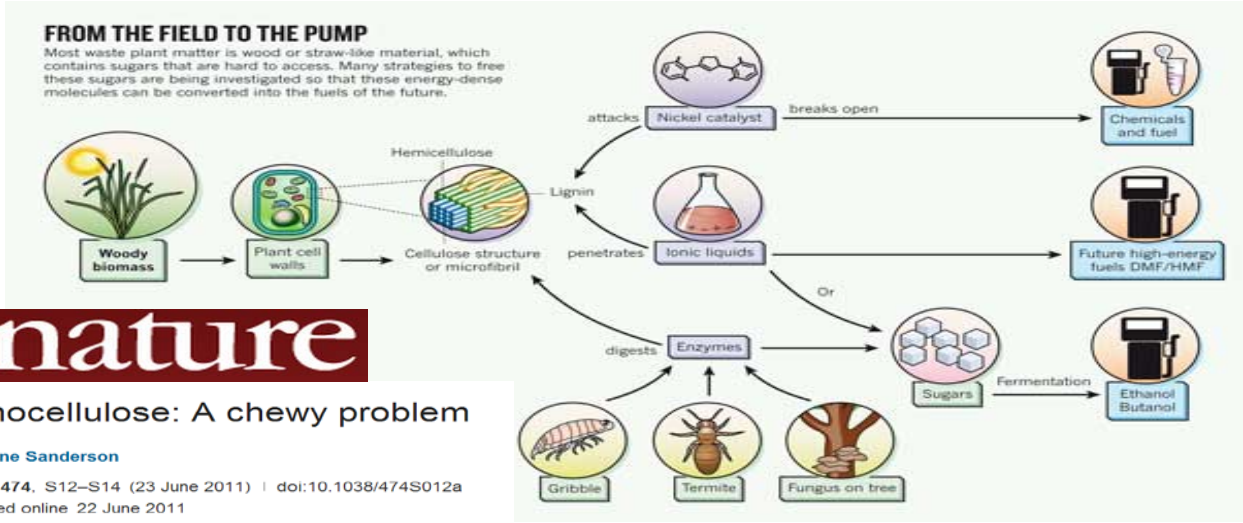
Vaccinate a month before the peak

Support for Novel and Innovative Projects
(led by Sergiu Sanielevici, Pittsburgh Supercomputing Center)

- Pro-actively reaching out to communities new to advanced computing, e.g.
  - social science
  - social network analysis
  - language processing
  - genomics
  - digital humanities (occasioned recent temporary hire)
  - economics
  - library science
  - public health
- Look for pioneers, work with them to craft the project, and then help them begin to execute it.

Large Genome Assemblies

- ECSS Staff working with leading researchers and code developers
- Largest ever metagenome assembly, using 3.5 TB RAM on PSC Blacklight



nature

Lignocellulose: A chewy problem

Katharine Sanderson

Nature 474, S12–S14 (23 June 2011) | doi:10.1038/474S012a
Published online 22 June 2011

"I wouldn't have been able to do anything on Blacklight without ECSS staff… (consultant) took a real interest and solved a lot of things that were hard for me. He found bugs in the software and got them resolved with the software authors. I'd worked for months and not made that progress. Without his expertise, I might have given up…"

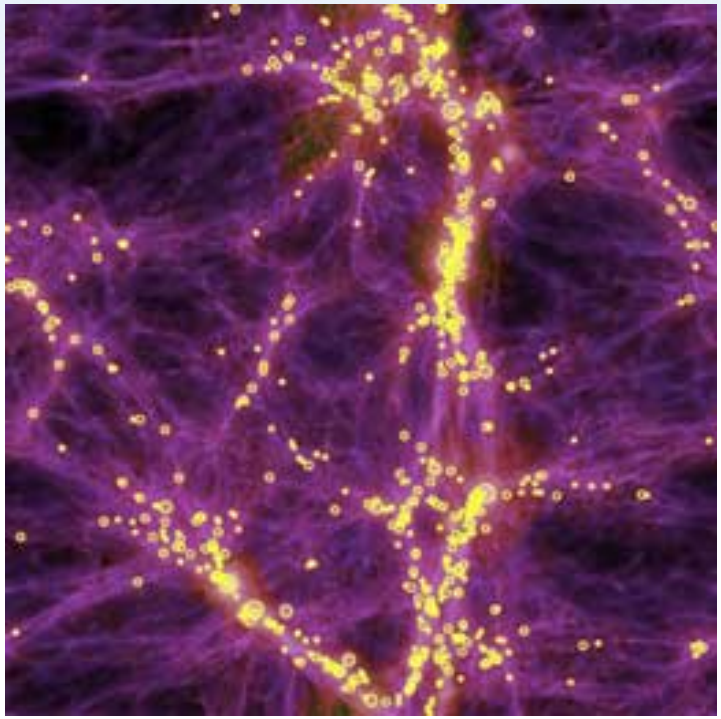
Support for Community Codes
(led by John Cazes, Texas Advanced Computing Center)

Community codes – applications, tools, and libraries used by multiple research groups

- Deploying, hardening, and optimizing useful software systems
- Assisting users with community codes and tools
- Establishing relationships between XSEDE and developer communities
- Extending XSEDE documentation to cover community code implementations

Gadget- A Widely Used Astrophysics Code

- Sometimes when you change systems, and especially compilers, things that used to work, no longer do.
- Usually, this is because the author of the code had used what they thought was a default feature of a language, but it really wasn't.
- In this case the consultant tracked down a serious error in Gadget code and eliminated an unsafe MPI call.
- The fixed version is now in the Gadget release, which is still in wide use.

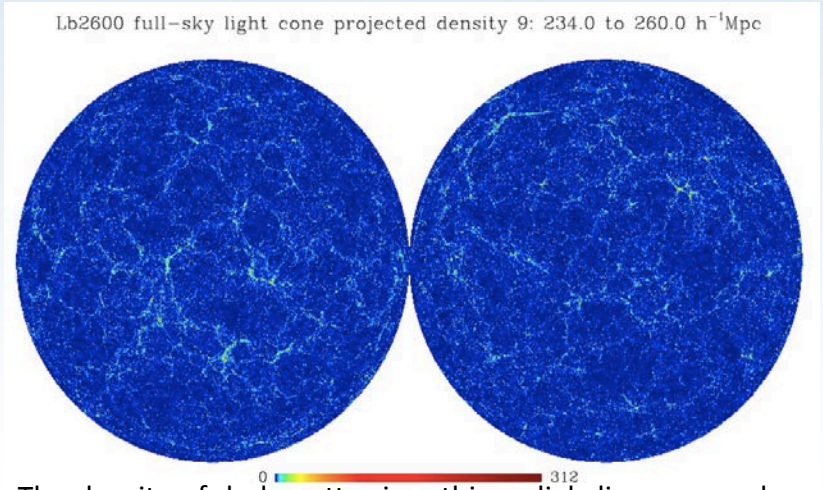


Support for Science Gateways
(led by Suresh Marru, Indiana U)

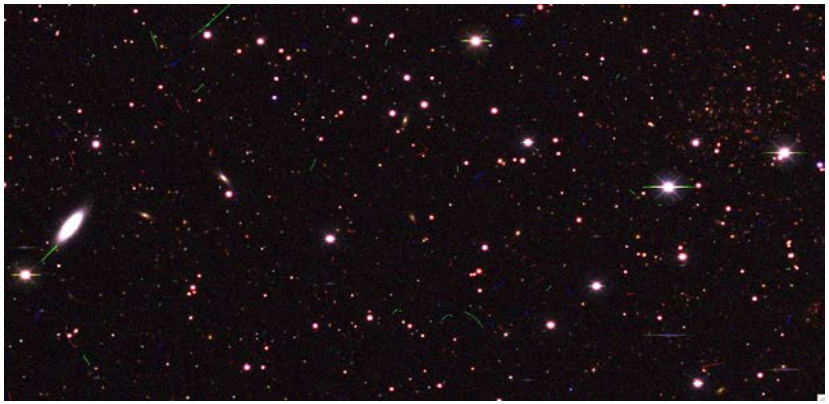
- Science Gateways enable communities of users associated with a common discipline to use computational resources through a familiar and simpler interface e.g. a web interface.
- Examples of ECSS Science Gateway Support
  - Assist with job submission and data movement to XSEDE
  - Automation of scientific processes through workflows
  - Integration of XSEDE resources into a portal/gateway
  - Support with grid security and community accounts

Realizing the Universe for the Dark Energy Survey (DES)

PI's: August Evrard (UM) and Andrew Kravtsov (UC)



The density of dark matter in a thin radial slice as seen by a synthetic observer located in the 8 billion light-year computational volume. Courtesy: Matthew Becker, University of Chicago.



A synthetic 2x3 arcmin DES sky image showing galaxies, stars, and observational artifacts. Courtesy Huan Lin, FNAL.

- ECSS provided support to the Dark Energy Survey Simulation Working Group in developing large, multi-staged computations (workflows) to develop Blind Cosmology Challenge (BCC) catalogs.
- These executions Catalog a variety of sky realizations in different cosmologies are analyzed, in a blind manner, by science teams generate expectations for galaxy yields in various cosmologies.
- Analysis of these simulated catalogs offers a quality assurance capability for cosmological and astrophysical analysis of upcoming DES telescope data.

Support for Training, Education, and Outreach

(led by Jay Alameda, NCSA)

- Collaboration with XSEDE Training, Education, and Outreach Services
- Training support
  - On-line course development and update
  - Synchronous training (live, web, …)
  - Collaboratively identify new training areas
- Education
  - Online course support
- Outreach
  - Conferences, speakers bureau, campus visits

Education: Applications of Parallel Computers

- Adapted from Jim Demmel's CS267 Course
  - Video, Slides in Cornell's Virtual Workshop
  - Quizzes in every class
- 3 Homework (programming) problems
  - ECSS staff helped port to XSEDE resources
  - ECSS staff part of support team for students
  - Autograding by Berkeley team
- Capped enrollment at 300 students
  - Limit reached in < 3 days
- Goal of supporting more courses, with more students


Our reach will forever
exceed our grasp, but,
in stretching our horizon,
we forever improve our world.
XSEDE
Extreme Science and Engineering
Discovery Environment