

I've been to the summer camp, now what?

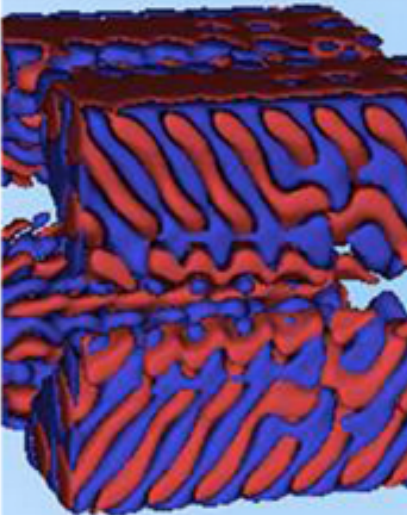
June 4, 2015

Sharon Broude Geva
Director of Advanced Research Computing (ARC)

sgeva@umich.edu
arc.umich.edu

What is ARC?

Computational Science



Michigan Institute for
Computational Discovery
and Engineering (MICDE)

Data Science



Michigan Institute for
Data Science (MIDAS)

Technology Services



Advanced Research
Computing - Technology
Services (ARC-TS)

Consulting Services



Center for Statistical
Consultation and Research
(CSCAR)

Advanced Research Computing (ARC) enables data-intensive and computational research at U-M

Advanced Research Computing (ARC):

- Provides [Flux](#), the shared, campus-wide high-performance computing cluster through [Advanced Research Computing - Technology Services \(ARC-TS\)](#)
- Provides or facilitates access to [other research computing resources](#) on and off the U-M campus, including running a **free** data science [Hadoop](#) cluster, through ARC-TS
- Affiliates the [Michigan Institute for Computational Discovery and Engineering \(MICDE\)](#) and the [Michigan Institute for Data Science \(MIDAS\)](#) to support academic programmatic initiative and multi-disciplinary collaboration
- Promotes training and support for users of computational research resources, through the [Center for Statistical Consultation and Research \(CSCAR\)](#), and a variety of other learning opportunities available to the U-M community.

Is advanced research computing relevant to me?

- NSF HPC+ Strategy high-level goal:
“Provide computational infrastructure to advance computational- and data-enabled science and engineering across all scientific and engineering disciplines”
- ACI-1341698, Michael Norman, UCSD, “Gateways to Discovery: Cyberinfrastructure for the Long Tail of Science” (Comet system), 10/1/2013, 5 years, \$12M
- ACI-1341711, Daniel Stanzione, UT-Austin, “Wrangler: A Transformational Data Intensive Resource for the Open Science Community” (Wrangler system), 11/1/2013, 2 years, \$6M

Funding for Big Data Core Technologies

- In 2012 & 2013, NSF & NIH awarded 45 projects ranging from \$250K/year for up to 3 years to \$1M/year for up to 5 years
- 51% by number of projects went to “Data Collection, Management, Mining and Machine Learning”
- An additional 10% went to “Social Networks”

The sky's the limit (currently [Blue Waters](#) is)...



Where can I find information about advanced research computing?

- The ARC website: arc.umich.edu
- ARC weekly email: to subscribe, <http://arc.umich.edu/news-events/subscribe-to-the-arc-newsletter/>
- Research Computing Symposia (Spring, Fall)
- Research Computing Symposium poster sessions (prizes!)
- My Twitter: @sbroudegeva (relevant retweets from various sources, no cats)
- ARC's Twitter: @ARCCatUM





ADVANCED
RESEARCH COMPUTING
UNIVERSITY OF MICHIGAN

... and training?

- Flux 100, Flux 101 and others - every couple of months
- <http://arc-ts.umich.edu/training-workshops/>
- Flux open user meetings
- ARC website + weekly email
- ARC Twitter (advance notice for training!)
- Online resources, for example:
 - Python - <http://www.codecademy.com/>
 - SQL - <http://www.sqlcourse.com>

More involved training and learning

- VSCSE Science Visualization (August 24-25) <https://portal.xsede.org/course-calendar/-/training-user/class/382/session/700> (Free, onsite at U-M from TACC)
- VSCSE Supercomputing for Everyone Series: Performance Tuning Summer School (August 17-21) <https://portal.xsede.org/course-calendar/-/training-user/class/420/session/701> (Free, onsite at U-M, from IU)

Info about events is always posted on  **ARC** website and sent out in the periodic email update  **ARC** ADVANCED RESEARCH COMPUTING UNIVERSITY OF MICHIGAN

Graduate Data Science Certificate Program

- Through the Michigan Institute for Data Science (MIDAS)
- The Rackham-approved Data Science Certificate program aims to provide core experiences in:
 - (Modeling) Understanding of core Data Science principles, assumptions & applications;
 - (Technology) Data management, computation, information extraction & analytics;
 - (Practice) Hands-on experience with modeling tools and technology using real data.

For more information, <http://midas.umich.edu/certificate/>

Contact: Ivo D. Dinov (dinov@umich.edu)

Where can I find more compute power?

- Flux - the on-campus shared computing cluster (provided by ARC; a for-fee service)

<http://arc-ts.umich.edu/flux/>

*

Some schools and departments have also bought allocations for shared use

- XSEDE - 16 supercomputers and high-end visualization and data analysis resources across the country (Provided by the NSF; free with a short proposal) www.xsede.org

Contact: Brock Palen, hpc-support@umich.edu

Where can I find people to help me?

- **ARC Liaisons:** Charles Antonelli (cja@umich.edu) (LSAIT) for LSA; Todd Raeker (raeker@umich.edu) for Ross and other Central Campus units
- **XSEDE** - Brock Palen (hpc-support@umich.edu)
- **UM3D lab** - Advanced visualization
- **CSCAR** - Statistics consulting (<http://cscar.research.umich.edu/consulting>)
- **Visualization Librarian** - Justin Joque
- **Spatial and Numeric Data Librarians** (assist in finding, manipulating and analyzing diverse types of data, GIS) (<http://www.lib.umich.edu/clark-library/services/sand>)

Besides social media, where else can I find data online?

- HathiTrust - Millions of digitized library collections (Jeremy York, MLibrary) <http://www.hathitrust.org/>
- DPLA - Digital Public Library of America dp.la
- EEBO-TCP - Early English Books 1475-1700 (Rebecca Welzenbach, MLibrary) <http://www.textcreationpartnership.org/tcp-eebo/>

Advanced Research Computing

Questions?

sgeva@umich.edu

arc.umich.edu