Humans in Scaling HPC
Facilitation and Education

IEEE-HPEC
September 2019

Bob Freeman, PhD
Director, Research Technology Operations
Harvard Business School
Why is Scaling with People an Important Problem?

- Science and research programs are no longer an isolated, work-in-silos problem
- Increase in complexity is driving an increased need for better trained workforce
  - "HPC was evolving from traditional modeling & simulation into complex workflows of data collection, modeling, simulation, analytics, and AI" – Michael Rosenfeld, IBM
- Changes in technology, research approaches, business landscape, and staffing
- RC Consultants, Educators, & Facilitators
  - Work with users – researchers and educators – to help them improve their research and/or education productivity and aspirations via advanced cyberinfrastructure.

- Briefly discuss 4 topic areas on
  - Create Community in Education
  - Bring HPC/HTC Computing to All Disciplines
  - Bring Facilitation Approaches to Everyone
  - Building Communities for Enabling Research
The impact of MOOC methodology on the scalability, accessibility and development of HPC education and training

Julia Mullen
MIT Lincoln Laboratory
Lexington, MA
jms@mit.edu

Lauren Mulechini
Massachusetts Institute of Technology
Cambridge, MA
lauren.mulechini@mit.edu

Weronika Filinger
EPCC, The University of Edinburgh
Edinburgh, United Kingdom
w.filinger@epcc.ed.ac.uk

David Hearty
EPCC, The University of Edinburgh
Edinburgh, United Kingdom
d.hearty@epcc.ed.ac.uk

ABSTRACT
This work explores the applicability of Massive Open Online Courses (MOOCs) for scaling High Performance Computing (HPC) training and education. Most HPC centers recognize the need to provide their users with HPC training; however, the current educational structure and accessibility present many difficulties to non-HPC specialists. In this paper, MOOCs are developed for HPC practitioners. MOOCs enable more accessible and scalable learning paths toward HPC expertise, facilitating domain experts to create MOOCs and shorten the learning curve. In this paper, we present a preliminary assessment of the impact of MOOC methodology on the scalability, accessibility, and development of HPC education and training.
Create Community in Education

Harvard University Research Computing* Trainers group (HURT)

TORCH: Trainers Optimizing Research Computing across Harvard

FAS RC  FAS HumRC  HMS RITG

HBS RCS  HSPH BioInfo  FAS / SEAS

IEEE-HPEC, Bob Freeman, PhD, 25-Sep-2019
Bring HPC/HTC Computing to All Disciplines

Science Gateways are community-provided interfaces to XSEDE resources. This section provides a general overview and describes XSEDE support for science gateways. The section provides links to more detailed information on the XSEDE user portal and external sites.

A Science Gateway is a community-developed set of tools, applications, and data that are integrated via a portal or a suite of applications, usually in a graphical user interface, that is further customized to meet the needs of a specific community. Gateways enable entire communities of users associated with a common discipline to use national resources through a common interface that is configured for optimal use. Researchers can focus on their scientific goals and less on assembling the cyberinfrastructure they require. Gateways can also foster collaborations and the exchange of ideas among researchers.

Using Science Gateways

Gateways are independent projects, each with its own guidelines for access. Most gateways are available for use by anyone, although they usually target a particular research audience. XSEDE Science Gateways are portals to computational and data services and resources.

- High- Throughput Computing (HTC)
- Virtual Organizations
- Collaborations
- Job Scheduling
- Education

XSEDE Start-up and Educational allocations require only a one paragraph project description.

Key Points
- Gateways provide a higher level user interface for XSEDE resources that are tailored to specific scientific communities.
- XSEDE supports gateways through community accounts, gateway hosting, and extended collaborative support services.

Related Links
- Science Gateways Listing
- XSEDE Gateways Resources
- Science Gateways for PBS
- Science Gateways for Developers

IEEE-HPEC, Bob Freeman, PhD, 25-Sep-2019

Design Patterns
Elements of Reusable HPC/HTC Workflows
Julia Mullen
Lauren Milechin
Bob Freeman
Your name here

Copyrighted Material

ADDISON-WESLEY PROFESSIONAL COMPUTING SERIES

HARVARD BUSINESS SCHOOL
HPC Carpentry:
Teaching basic skills for high-performance computing.

HPC Carpentry is a set of teaching materials designed to help new users take high-performance computing systems. No prior computational experience is needed. The lessons are ideal for either an in-person workshop or independent study.

NOTE: This is the draft HPC Carpentry release. Comments and feedback are welcome.

WHAT IS HPC CARPENTRY?

Each of these lessons is designed to teach how to interact with a compute cluster in a specific way. The first two lessons, Intro to HPC, and Shell on HPC are meant to be used together as a general introduction to the bash command line and submitting jobs on a typical HPC cluster. The second two lessons focus on two separate use cases for HPC: running large numbers of compute jobs (often with complex inter-job dependencies), or creating parallel programs that execute across one or more compute nodes. Each lesson takes roughly a full day to teach (or work through independently). It is intended that the Intro to HPC lesson be taught together with either the Python or Chapel component to form a two-day workshop or the Intro to HPC + Shell can be taught as a standalone.

Next steps & How can you get involved??

In progress:
Tech wizardry to facilitate scheduler-agnostic lessons with site-specific (forked & rendered) scheduler & resource details
Updated Outline.md document with lesson topics, objectives, & goals to help direct lesson development

If you wish to help develop lessons, join discuss-hpc https://carpentries.topicbox.com/groups/discuss-hpc
If you’ve run these lessons, please give us feedback! (discuss-hpc)

Look for upcoming details about an HPC-Carpentry Sprint!!
Bring Facilitation Approaches to All Disciplines

NSF-Funded Project – ACI-REF

$5.3M NSF Award supports the project leadership team and 2 Facilitators for each of the 6 partner sites for 2 years.

PI: Jim Bottum, Clemson

Project Leadership:
• James Cuff, Harvard
• Maureen Dougherty, USC
• Gwen Jacobs, Hawai'i
• Paul Wilson, Wisconsin
• Tom Cheatham, Utah (PI Chair)
• Marcin Ziolkowski, Clemson

Facilitator Leads: Bob Freeman, Harvard

Chief Scientist: Miron Livny, Wisconsin

• "Work smarter, better, faster": be more efficient
• ... and to Think Differently: ask bigger questions and not be constraint by current (desktop) resources
• Push for answers – always find out why; just don’t “fix the problem”

Empowering Would-Be Computational Researchers

We are dedicated to forging a nationwide alliance of educators to empower local campus researchers to be more effective users of advanced cyberinfrastructure (ACI). In particular, we seek to work with the "long tail" of ACI users—those scholars and faculty members who traditionally have not benefited from the power of massively scaled cluster computing but who recognize that their research requires access to more compute power than can be provided by their desktop machines.

To do so, the consortium is building a coordinated network of ACI-REFs, campus champions whose mission is to leverage existing resources and “make a difference” in supporting their local campus researchers, while also unifying member institutions under common objectives.

Aligned with strategies and models for advancing the nation's research capabilities, the ACI-REFs will be computational scientists with broad backgrounds in computer instruction, uniquely qualified to bridge the gap between researchers and campus IT.

A Startup Framework for Building Digital Research Capacity and Community at UCLA

Annelie Rugg, UCLA Humanities Technology

IEEE-HPEC, Bob Freeman, PhD, 25-Sep-2019
Building Communities for Enabling Research

Our community of over 600 Campus Champions promotes and facilitates the effective participation of a diverse national community of academic and not-for-profit institutions in the application of advanced digital resources and services to accelerate discovery, enhance education, and foster scholarly achievement.

There are over 600 Campus Champions: Including directors, faculty, researchers, students, as well as research-enabling and systems professional At over 300 academic, non-academic, and not-for-profit research-focused institutions

https://www.xsede.org/community-engagement/campus-champions

---

Henry Neeman, University of Oklahoma
Director, OU Supercomputing Center for Education & Research (OSCER)
Associate Professor, College of Engineering
Adjunct Faculty, School of Computer Science
XSEDE Campus Engagement Joint Co-Manager
Virtual Residency Introductory/Intermediate Workshop 2019
Sunday June 2 2019

---

This Is So New, We Don’t Know How to Teach It

- For the Introductory workshops, we were able to find speakers for most of the topics we covered.
- For this combined Introductory/Intermediate workshop, very few of the topics are issues that any of us know enough about to be able to teach it to others at the Intermediate level.
- So, most of the Intermediate sessions are panels – we’ll learn from each other!
Building Communities for Enabling Research


Initially, a group of 28 institutions intent on developing charter, membership rules, etc.

Now producing deliverables for the community and open to all interested and willing to contribute.

https://carcc.org/ info@carcc.org

Ecosystem Workshop Participating Orgs

- Association of Research Libraries (ARL)
- Big Data Hubs
- Campus Champions (CC)
- Campus Research Computing Consortium (CaRCC)
- Carpentries
- Coalition for Academic Scientific Computation (CASC)
- Coalition for Networked Information (CNI)
- Education Opportunities (HPC University, SIG HPC Education)
- EDUCAUSE
- EPOC/CI Engineers
- Global Environment for Network Innovations (GENI)
- HPC Systems Professionals
- Midscale Experimental Research Infrastructure Forum (MERIF)
- Minority Serving Institution/ Historically Black Colleges & Universities
- Open Science Grid (OSG)
- Quilt (Regional Networks)
- Research Data Access & Preservation Association (RDAP)
- Women in HPC (WHPC)
- XSEDE (Extreme Science and Engineering Discovery Environment)
Building Community: the People Network

Initially organized around the five “facings”, each track provides for focused discussion. Each has (or will have) monthly videoconferences and mailing list.

- **Researcher-facing**: Launched Spring 2018
- **Systems-facing**: Launched Jan 2019
- **Data-facing**: Launched May 2019
- **Emerging-Centers**: Launching Sept 2019!
- **Software-Facing**: Launching at a future date
- **Sponsor/Stakeholder-facing**: To be developed

More info and join: [https://carcc.org/people-network/](https://carcc.org/people-network/)

Researcher- & Staff-facing, & Professional Development
- Effective Meetings
- Promoting Good User Behavior
- Crowdsourcing Training Friction Points / Open Mic
- Communicating a Problem / Open Mic
- Getting the researcher turnout you want
- Software, containers, libraries, ... oh my!
- XSEDE ECSS: How we can help you
- Lightning talks / Open Mic
- Creating effective training materials
- PEARC 2019 Follow-up
- "A Startup Framework for Building Digital Research Capacity and Community at UCLA"

Facilitation beyond the RC Center:
- Partnerships & collaborations with other service groups
- IT Department
- Research Administration
- Behavioral Research
- Security & Compliance
Questions?