Multidisciplinary Online Training Program in Spring 2020 on Big Data + High Performance Computing + Atmospheric Sciences

Information and apply by 01/01/20 at http://cybertraining.umbc.edu

Introduction

CyberTraining at UMBC is an NSF-funded initiative to create a training program in big data applied to atmospheric sciences as application area and using high-performance computing as indispensable tool. The training consists of instruction in all three areas of "Big Data + HPC + Atmospheric Sciences", followed by faculty-guided project research in a multidisciplinary team of participants from each discipline. Participants can be **graduate students, post-docs, and junior faculty** from around the nation who will be exposed to multidisciplinary research and have the opportunity for significant career impact. Both instruction and research are mentored by faculty and supported by teaching assistants from each discipline. **Each participant will receive $1,500 stipend.**

We successfully trained 16 participants in 2018 and 17 participants in 2019 (5 teams each year). Each team delivered a technical report and most reports have been extended for peer-reviewed conference or journal papers, and faculty and participants are presenting there. Detailed information can be found at http://cybertraining.umbc.edu/2018.html and http://cybertraining.umbc.edu/2019.html.

Training program syllabus

Module	Topic	Goal
1	Introduction of Python/C, Linux and HPC environment	Running their own jobs on HPC
2	Numerical methods for Partial Differential Equations (PDE)	Model as PDE and solve them using numerical methods
3	Message Passing Interface (MPI)	Write MPI jobs and performance studies
4	Basics of earth-atmosphere radiative energy balance and global warming	Understand basic concepts and principles of radiative energy balance and global warming
5	Basics of radiative transfer simulation framework	Understand the basic physics underlying the transport of radiation in atmosphere
6	Global Climate Model (GCM) simulation and satellite observations	Understand the importance of GCM and satellite remote sensing
7	Introduction of Big Data	Understand the basics of Big Data and demo programs
8	Big Data system: Hadoop/Spark	Write Hadoop/Spark jobs and run them on HPC
9	Big Data Machine learning	Write machine learning programs using Spark MLlib
10	Introduction of Deep Learning	Write deep neural network programs using Keras
11	Project introduction and assignment	Each interdisciplinary team will be assigned one project
12-14	Project progress report from each team and feedback	15 minutes report from each team + Q&A + rating
15	Final project presentation	Report, software, and a final presentation from each team

Online training structure

The training will adopt a flipped classroom approach for the instruction in modules 1-10. Participants watch the lecture videos individually first, then work on homework as a team. Help is available at all stages from instructors and TAs through a discussion board. During modules 11-15, each team will present their research progress online and get feedback. Participants, instructors, and TAs communicate online every week on Fridays 3:30-5:00 ET. Each participant is expected to spend 9-12 hours per week on the training.

Who should participate

- Students/researchers interested in interdisciplinary research and how Big Data and HPC techniques can be applied to Computational Physics and other Computational Sciences.
- Graduate students and post-doctoral researchers / junior faculty who want to participate in project-based multidisciplinary research to further their career.

Ideal prerequirite knowledge for different majors

Major	*Ideal Prior Expertise*
Computing related	• Programming • Distributed Systems
Mathematics / Statistics	• Partial Differential Equations • Computational Mathematics and Programming
Physics	• Computational Physics

Instructors

- Dr. Jianwu Wang, Department of Information Systems
- Dr. Matthias Gobbert, Dept. of Mathematics and Statistics
- Dr. Zhibo Zhang, Department of Physics
- Dr. Aryya Gangopadhyay, Dept. of Information Systems

How to apply for Spring 2020 training (01/2020 - 05/2020)

There is no fee to apply to the training program. **Each participant who successfully finishes the program and completes all requirements will receive $1,500 stipend.** We expect to have 15 participants in 5 multidisciplinary teams of three, one participant from each discipline. Due to the capacity limit, there will be a selection process. Application is due 01/01/20.

The application package should include applicant's CV, personal statement, transcript, and at least two letters of recommendation. The personal statement needs to address specifically why the participant is interested in interdisciplinary research, the background in software tools and languages, how participation will promote his/her career goals, and how he/she can contribute to a team of participants from each discipline.