EDUCATION & WORKFORCE DEVELOPMENT

Melyssa Fratkin
Industrial Programs Director
IDC HPC User Forum
Sept 7, 2016
LIFELONG LEARNING IN ADVANCED COMPUTATION

Undergraduate Programs  
(Verizon, NSF REU, STAR Scholars, etc.)

K-12 Outreach

University Courses in Scientific Computation  
(Undergraduate/Graduate)

Topic-focused Training in Advanced Computation  
(HPC, Data, Vis, LifeSci)

TACC Institute Series  
(HPC, Data/Vis, SysAdmin, LifeSci)
UNDERGRADUATE PROGRAMS

- NSF Research Experiences for Undergraduates (REU) grants
  - 10 undergraduates majoring in science and engineering disciplines
  - Participants explore grand challenges such as climate modeling, weather forecasting, drug delivery, brain mapping, energy exploration and understanding the human genome, among others.
- STAR Scholars (Funded by TACC industry partners)
STAR SCHOLARS

- Graduate or undergraduate interns, depending on project
- Students work under a TACC mentor, on projects of interest to Partners
- Students submit papers or posters at conferences
- Partner funded ($50K per student for 2 semesters)
- Funding from BP, Shell, NASA
STAR SCHOLARS - OUTCOMES

- Interactive Parallelization Tool - paper by Ritu Aurora and student Madhav Gupta
  https://www.youtube.com/watch?v=L4a19kF6q48
  - IPT can be used by domain-experts and students to semi-automatically generate parallel programs based on multiple parallel programming paradigms (MPI, OpenMP, and CUDA) and learn about these paradigms through observation and comparison.

- Student paper accepted to SC16: "A Tool for Semi-Automatic Application-Level Checkpointing"

- Work on TACCStats – Todd Evans and student Vivek Pradhan
  - Student benchmarked a NoSQL database for analysis of all of the job-level data from our systems that is collected by TACCStats.
  - Working on visualization to facilitate detection of poorly written programs that are not using resources effectively
UNIVERSITY COURSES IN ADVANCED COMPUTATION

- UT Austin dept. of Statistics and Data Sciences
- Dual-listed undergraduate/graduate
- Five (5) courses covering many aspects of advanced computation
  - Introduction to Scientific Programming – programming concepts using C/C++ and Fortran
  - Scientific and Technical Computing – basic numerical methods, fundamentals of floating-point arithmetic, common tools for scientific software development (git, make), profiling and optimization, debugging
  - Parallel Computing – developing parallel applications using OpenMP and MPI
  - Visualization and Data Analysis – tools for data analysis (Hadoop, R) and visualization of scientific datasets (VisIt, Paraview)
  - Distributed and Grid Computing – executing data parallel ensembles which do not require MPI using grid-enabled tools (Globus, Condor, etc.)
- Content and instructor notes made publicly available through a grant from Chevron
  - https://www.tacc.utexas.edu/education/academic-courses
TOPIC-FOCUSED TRAINING IN ADVANCED COMPUTATION

- Offered in both spring and fall academic semesters
- 1 and 2-day courses in HPC, Data, and Visualization
  - Mix of lecture and hands-on exercises
- Open to TACC users, academic and industry partners
- In many cases courses are live-streamed and recorded on YouTube
- Upcoming and past training at: [https://portal.tacc.utexas.edu/training](https://portal.tacc.utexas.edu/training)

Courses have included:
- MPI/OpenMP
- Profiling and Optimization
- Debugging
- Programming the Intel Xeon Phi (KNC/KNL)
- Python
- R
- Using Hadoop/Spark
- Data Management Practices and Techniques
- Introduction to Scientific Visualization
  ...and more
SUMMER SUPERCOMPUTING INSTITUTE (SSI)

- Held each summer since 2007
- Originally an aggregation of TACC’s topic-focused training courses
  - MPI, OpenMP, profiling, debugging, visualization
- Offers one-on-one meetings with computational staff to discuss research-specific problems
- From 2007-2015, enrollment steady around 40
- Registration fee for attendees
SUMMER SUPERCOMPUTING INSTITUTE (SSI)

- In 2016, SSI was expanded
- Two tracks
  - Parallel Applications
    - MPI, OpenMP, debugging, profiling, many-core programming (Xeon Phi)
  - Data Analysis and Visualization
    - Python, R, Data Management, Hadoop/Spark, VisIt/Paraview, High Throughput Computing
- Plenary talks showcasing newest research
- 90 participants
  - 5 countries represented
  - 18% female participation
COMING SOON – TACC INSTITUTE SERIES

► Series of week long training events modeled on SSI
  ► High Performance Computing
  ► Visualization and Data Analysis
  ► Cluster design and administration
  ► Life Sciences Computing

► Registration fee TBD
► To be offered during the summer months
► More details coming at SC16 in November!
MEASURING IMPACT

- More than 1,000 registered for training between Aug. 2015 and July 2016
- We also track number of independent views of our recorded training events
  - More than 11,000 views between Aug. 2015 and July 2016
    - Up 25% over previous year
- In-person participants are asked to answer surveys after training events
  - Satisfaction with content: > 70% “very satisfied”
  - Satisfaction with instructors: > 90% “very satisfied”
  - Would recommend to others: > 70% “definitely”
CONTACT US

- Melyssa Fratkin, Industrial Programs Director, mfratkin@tacc.utexas.edu
- Lucas Wilson, Director of Training & Professional Development, lwilson@tacc.utexas.edu
- https://portal.tacc.utexas.edu/training