Developing IEEE-TCPP Parallel/Distributed Curriculum and NSF CyberTraining Program

Sushil K Prasad

National Science Foundation
Georgia State University
Former Chair, IEEE Technical Committee on Parallel Processing (TCPP)

SCEC-18
TCPP Curriculum Initiative:
http://www.cs.gsu.edu/~tcpp/curriculum/
Outline

• IEEE-TCPP Curriculum
  – Why this curriculum initiative and what are the opportunities for the audience?
  – Key Activities and Milestones
    • ACM/IEEE 2013 CS Curriculum Taskforce
      – provided direct link to us for rigorous coverage
    – How was the curriculum formulated?
    – How is it getting evaluated?
  – Current Activities

• NSF CyberTraining Program
  – Computational and Data-driven Science for All
  – Goals; Communities of Concern
  – Award Framework
Who are we?

- Chtchelkanova, Almadena - NSF
- Dehne, Frank - University of Carleton, Canada
- Gouda, Mohamed - University of Texas, Austin, NSF
- Gupta, Anshul - IBM T.J. Watson Research Center
- JaJa, Joseph - University of Maryland
- Kant, Krishna – George Mason University
- La Salle, Anita - NSF
- LeBlanc, Richard, Seattle University
- Lumsdaine, Andrew - Indiana University
- Padua, David - University of Illinois at Urbana-Champaign
- Parashar, Manish - Rutgers
- Prasad, Sushil - Georgia State University
- Prasanna, Viktor - University of Southern California
- Robert, Yves - INRIA, France
- Rosenberg, Arnold - Northeastern
- Sahni, Sartaj - University of Florida
- Shirazi, Behrooz - Washington State University
- Sussman, Alan - University of Maryland
- Weems, Chip, University of Massachussets
- Wu, Jie - Temple University
Why now?

- Computing Landscape has changed
  - Mass marketing of multi-cores
  - General purpose GPUs even in laptops (and handhelds)
- A student with even a Bachelors in Computer Science (CS) or Computer Engineering (CE) must acquire skill sets to develop parallel software
  - No longer instruction in parallel and distributed computing primarily for research or high-end specialized computing
  - Industry is filling the curriculum gap with their preferred hardware/software platforms and “training” curriculums as alternatives with an eye toward mass market.
Stakeholders

- CS/CE Students
- Educators – teaching core courses as well as PDC electives
- Universities and Colleges
- Employers
- Developers
- Vendors
- Authors
- Researchers
- NSF and other funding agencies
- IEEE Technical Committees/Societies, ACM SIGs,
- Curriculum Task Forces such as CS2013 ACM/IEEE
How was the curriculum formulated?

Why would they come?

*Field of Dreams (1989):* "If you build it, he will come"
Curriculum Planning Workshops at DC (Feb-10) and at Atlanta (April-10)

• Goals
  – setup mechanism and processes which would provide periodic curricular guidelines
  – employ the mechanism to develop sample curriculums

• Agenda:
  – Review and Scope
  – Formulate Mechanism and Processes
  – Preliminary Curriculum Planning
    • Core Curriculum
    • Introductory and advanced courses
  – Impact Assessment and Evaluation Plan

Main Outcomes

- Priority:
  Core curriculum revision at undergraduate level

- Preliminary Core Curriculum Topics

- Sample Intro and Advanced Course Curriculums
Weekly Tele-Meetings on Core Curriculum (May-Dec’10; Aug’11-Feb’12)

**Goal:** Propose core curriculum for CS/CS graduates

- Every individual CS/CE undergraduate must be at the proposed level of knowledge as a result of their required coursework

**Process:** For each topic and subtopic

1. Assign *Bloom’s classification*
   
   K = Know the term (basic literacy)
   
   C = Comprehend so as to paraphrase/illustrate
   
   A = Apply it in some way (requires operational command)

1. Write *learning outcomes*
2. Identify core CS/CE courses impacted
3. Assign number of hours
4. Write suggestions for “how to teach”
### TCPP Curriculum Example

**4 Curriculum Areas**
- Architecture, Programming, Algorithms, Cross-cutting

<table>
<thead>
<tr>
<th>Algorithms Topics</th>
<th>Bloom</th>
<th>Course</th>
<th>Learning Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Algorithmic problems</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Communication</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>broadcast</td>
<td>C/A</td>
<td>Data Struc/Algo</td>
<td>The important thing here is to emphasize the parallel/distributed aspects of the topic</td>
</tr>
<tr>
<td>multicast</td>
<td>K/C</td>
<td>Data Struc/Algo</td>
<td>represents method of exchanging information - one-to-all broadcast (by recursive doubling)</td>
</tr>
<tr>
<td>scatter/gather</td>
<td>C/A</td>
<td>Data Structures/Algorithms</td>
<td>Illustrate macro-communications on rings, 2D-grids and trees</td>
</tr>
<tr>
<td>gossip</td>
<td>N</td>
<td></td>
<td>Not in core</td>
</tr>
</tbody>
</table>

**Asynchrony**

- **K**
- **CS2**
- asynchrony as exhibited on a distributed platform, existence of race conditions

**Synchronization**

- **K**
- CS2, Data Struc/Algo
- aware of methods of controlling race condition,

**Sorting**

- **C**
- CS2, Data Struc/Algo
- parallel merge sort,

**Selection**

- **K**
- CS2, Data Struc/Algo
- min/max, know that selection can be accomplished by sorting

Prasad/SCEC-18
How is the Curriculum being evaluated?

Early Adopter Program
EduPar/EduHPC/Euro-EduPar Workshop series
Early Adopter Program

• Over 100 institutions worldwide
  – Spring-11: 16 institutions; Fall’11: 18;
  – Spring-12: 21; Fall-12: 25 institutions, Fall-13: 25 institutions, Fall-14: 25, Fall-15: 13
  – Most from US (4 year to research institutions, one high school)
  – Some from South America, a few from Europe, fewer from Asia (India, China, Indonesia, Singapore), Middle East

• Next competition: Deadline Feb 12, 2019
  – NSF/Intel funded Cash Award/Stipend up to $1500-5000/proposal
  – Which course(s), topics, evaluation plan?

• Instructors for core CS/CS courses such as CS1/2, Systems, Data Structures and Algorithms – department-wide multi-course multi-semester adoption preferred
  – Elective courses; graduate courses
Edu* Workshop Series

- EduPar-11 at Alaska, IPDPS-2011
  - Receive feedback from the Adopters
  - Stimulate discussion of curricular and other educational issues.
- EduPar-12 at Shanghai, IPDPS-2012
  - A regular satellite workshop of IPDPS
- EduPar-13 in Boston + EduHPC Workshop at SC-13 + BOF at SIGCSE-14
- EduHPC-14 @ SC-14, Nov – New Orleans; EduHPC-15 in Austin, EduHPC-16, EduHPC-17, EduHPC-18 in Dallas
- EduPar-15 @IPDPS, May, India; EduPar-16, Chicago, EduPar-17 in Orlando; EduPar-18 in Vancouver
- EduHiPC 2018 @ HiPC in Bangalore – for India and the region
  - Monday, Dec 2018
- EduPar-19 @ IPDPS  in Rio  in May’19
  - Deadline Jan 18, 2019
NOW OPEN - CDER Courseware Website

Upload and Search Course Material

• **Type:**
  – Slides, Syllabus, Tutorial, Video
  – Animation, Article, Award, Blog, Book, Competition
  – Course Template, Course Module, Data
  – Hardware Access, Software/Tools
  – Proposal, Report

• **Courses:**
  – CS1, CS2, Systems, Data Structures and Algorithms, …

• **NSF/TCPP Topic/Subtopic Classification:**

  ALGORITHMS
  – Parallel and Distributed Models and Complexity
  – Algorithmic Paradigms
    – Divide & conquer (parallel aspects)
    – Algorithmic problems

  ARCHITECTURE
  PROGRAMMING
  CROSS-CUTTING
CDER Book Project

• Lack of suitable textbooks to integrate PDC topics into the core courses
  – CS1, CS2, Systems, and Data Structures and Algorithms

• Part I - For instructors: Basic Concepts and References on what and how to teach

• Part 2: For students: Supplemental teaching material for core courses

• 9 chapters
  – over 27K chapter downloads – free downloads

• 2nd Volume – Published Nov’19
  – Vol 3 – Early Adoptor course and topic exemplars and accompanying resources
<table>
<thead>
<tr>
<th></th>
<th>Areas</th>
<th>Architecture</th>
<th>Algorithms</th>
<th>Programming</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>New Aspects</strong></td>
<td><strong>Area Lead/Aspect Lead</strong></td>
<td>Chip Weems</td>
<td>Arnold Rosenberg</td>
<td>Alan Sussman</td>
</tr>
<tr>
<td><strong>Exemplars</strong></td>
<td>Sushil Prasad</td>
<td>Karen Karavanic,</td>
<td>Erik Saule, Duane Merril, David Bunde</td>
<td>David Brown, Eric</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Eric Freudenthal</td>
<td></td>
<td>Freudenthal</td>
</tr>
<tr>
<td><strong>Distributed</strong></td>
<td>Vaidyanathan Ramachandran</td>
<td>Vaidyanathan</td>
<td>Vaidyanathan Ramachandran, Costas Busch,</td>
<td>Alan Sussman, Chi</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Ramachandran,</td>
<td>Manish Parashar, Denis Trystram</td>
<td>Shen</td>
</tr>
<tr>
<td><strong>Big Data</strong></td>
<td>Trilce Estrada</td>
<td>Craig Stunkel</td>
<td>Cynthia Phillips, Debzani Deb</td>
<td></td>
</tr>
<tr>
<td><strong>Energy</strong></td>
<td>Krishna Kant, Craig Stunkel</td>
<td>Craig Stunkel,</td>
<td>Denis Trystram</td>
<td>John Dougherty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Karen Karavanic</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Crosscutting</strong></td>
<td>Sheikh Ghafoor, Arny Rosenberg, Anshul</td>
<td>Craig Stunkel,</td>
<td>Robert Robey, Martina Barnas</td>
<td>Sheikh Ghafoor, Eric</td>
</tr>
<tr>
<td></td>
<td>Gupta</td>
<td>Eric Freudenthal</td>
<td></td>
<td>Freudenthal</td>
</tr>
</tbody>
</table>
Sponsorship Acknowledgements

- NSF
  - NSF/TCPP Curriculum Initiative
  - Early adopter competitions (stipend, travel)
  - EduPar/EduHPC workshop series
  - CRI-ADDO CDER (2012-15)

- Intel
  - international early adopter institutions (stipend, travel)

- nVIDIA
  - GPU cards to all the 50+ early adopters from Spring'11, Fall'11 and Spring'12 rounds.

- IEEE TCPP, IBM
  - Keynotes in the past
Innovations in NSF Advanced Cyberinfrastructure Research Workforce Development and Education Programs

Office of Advanced Cyberinfrastructure (OAC)
Computer and Information Science & Engineering (CISE)
National Science Foundation

Sushil K Prasad,
Questions: sprasad@nsf.gov
Dec 2018
NSF Office of Advanced Cyberinfrastructure

Program Staff

- Computing: Bob Chadduck, Ed Walker
- Data: Amy Walton, Stefan Robila
- Software: Vipin Chaudhary, TBD
- Networking & Cybersecurity: Micah Beck
- Learning & Workforce Development: Kevin Thompson
- Bill Miller
- Beth Plale
- Alejandro Suarez
- Julie Stalhut

Join NSF/OAC: Multiple Program Officer openings

* IPA Appointment
My Journey as a NSF Program Director

• What does ACI/OAC do – OAC’s Mission?
  1. Advanced CI – cyberinfrastructure funding in HW, SW, Data, Networking, Security
  2. Forward looking research and education, but...

• Status of research and education programs in OAC
  – Dwindling when arrived in 2015 – participation in CAREER, CRII, REU site; NRT
  – Multidisciplinary, use-inspired focus

• My IEEE TCPP experience
  – Massive Outreach
  – Connecting with diverse, multidisciplinary research communities

=> OAC CAREER: Twice as many proposals in 2016; Thrice in 2017
My Journey as a NSF Program Director

• What does ACI/OAC do – OAC’s Mission?
  – Advanced CI – cyberinfrastructure funding in HW, SW, Data, Networking, Security
  – Forward looking research and education, but...

• Status of research and education programs in OAC
  – Dwindling when arrived in 2015 – participation in CAREER, CRII, REU site; NRT
  – Multidisciplinary, use-inspired focus

• My IEEE TCPP experience
  – Massive Outreach
  – Connecting with diverse, multidisciplinary research communities

=> OAC CAREER: Twice as many proposals in 2016; Thrice in 2017
My Journey as a NSF Program Director

• What does ACI/OAC do – OAC’s Mission?
  – Advanced CI – cyberinfrastructure funding in HW, SW, Data, Networking, Security
  – Forward looking research and education, but...

• Status of research and education programs in OAC
  – Dwindling when arrived in 2015 – participation in CAREER, CRII, REU site; NRT
  – Multidisciplinary, use-inspired focus

• My IEEE TCPP experience
  – Massive Outreach
  – Connecting with diverse, multidisciplinary research communities

=> OAC CAREER: Twice as many proposals in 2016; Thrice in 2017
My Journey as a NSF Program Director

• What does ACI/OAC do – OAC’s Mission?
  – Advanced CI – cyberinfrastructure funding in HW, SW, Data, Networking, Security
  – Forward looking research and education, but...

• Status of research and education programs in OAC
  – Dwindling when arrived in 2015 – participation in CAREER, CRII, REU site; NRT
  – Multidisciplinary, use-inspired focus

• My IEEE TCPP experience
  – Massive Outreach
  – Connecting with diverse, multidisciplinary research communities

=> OAC CAREER: Twice as many proposals in 2016; Thrice in 2017
OAC Research and Education Scope

• OAC mission: 2. Forwarding looking research and education
  – Multidisciplinary, use-inspired focus – ambiguous
• Continual Internal discussions within OAC and NSF
• Studied current and past programs
• Workshops; NSCI, NAS study
• Converged on a key gap in training/education=> CyberTraining Program
  – Computational and data-driven science for all
  – 2 competitions in 2017 and 2018
    • extraordinary response and growth
OAC Research and Education Scope

• OAC mission: 2. Forwarding looking research and education
  – Multidisciplinary, use-inspired focus

• Continual Internal discussions within OAC and NSF

• Studied current and past programs

• Workshops; NSCI, NAS study

• Converged on a key gap in training/education=> CyberTraining Program
  – Computational and data-driven science for all
  – 2 competitions in 2017 and 2018
    • extraordinary response and growth
OAC Research and Education Scope

• OAC mission: 2. Forwarding looking research and education
  – Multidisciplinary, use-inspired focus
• Continual Internal discussions within OAC and NSF
• Studied current and past programs
• Workshops; NSCI, NAS study
• Converged on a key gap in training/education=> CyberTraining Program
  – Computational and data-driven science for all
  – 2 competitions in 2017 and 2018
    • extraordinary response and growth
OAC Research and Education Scope

- OAC mission: 2. Forwarding looking research and education
  - Multidisciplinary, use-inspired focus
- Continual Internal discussions within OAC and NSF
- Studied current and past programs
- Workshops; NSCI, NAS study
- Converged on a key gap in training/education => CyberTraining Program
  - Computational and data-driven science for all
  - 2 competitions in 2017 and 2018
    - extraordinary response and growth
Planning for the Future CI Ecosystem

Key Drivers

- Changing application landscape & workload profile
- Changing technology, services landscape
- Increasing availability of (exp., obs.) data
- Growing role of ML, data-driven approaches

Prasad/SCEC-18
Communities of Concern

CI Contributors
Cyber Scientists
to develop new capabilities

CI Professionals
Professional Staff
to deploy & support new capabilities

CI Users
Area Scientists
to exploit new capabilities

OAC

Prasad/SCEC-18
Training-based Workforce Development for Advanced Cyberinfrastructure (CyberTraining)
NSF 19-524
(replaced NSF 18-516)

Submission Deadline: Feb 6, 2019
**Overarching and Solicitation Goals**

- **Overarching Goal:** prepare, nurture and grow scientific research workforce
- **Goal 1:** ensure **broad adoption** of CI tools, methods, and resources, OR
- **Goal 2:** integrate skills into educational **curriculum/instructional material fabric** in
  - advanced cyberinfrastructure (CI) +
  - computational and data science and engineering (CDS&E)
  - spanning undergraduate and graduate courses.
- **Innovative, scalable training, education, and curricular** programs addressing
  - targeting one or both of the solicitation goals
  - Emerging needs and Unresolved bottlenecks
  - Undergrads, grad students, instructors, faculty, research CI professionals
Overarching and Solicitation Goals

• **Overarching Goal:** *prepare, nurture and grow* scientific research workforce

• **Goal 1:** ensure *broad adoption* of CI tools, methods, and resources, OR

• **Goal 2:** integrate skills into educational *curriculum/instructional material fabric* in
  • advanced cyberinfrastructure (CI) +
  • computational and data science and engineering (CDS&E)
  • spanning undergraduate and graduate courses.

• **Innovative, scalable training, education, and curricular programs** addressing
  – targeting *one or both* of the solicitation goals
  – Emerging needs and Unresolved bottlenecks
  – Undergrads, grad students, instructors, faculty, research CI professionals
NSF-wide Participation

- CISE/OAC - Office of Advanced Cyberinfrastructure – lead
  - Sushil K Prasad (Includes BD Hub)
- CISE/CCF Computing and Communication Foundation
  - Almadena Chtchelkanova
- EHR/DGE - Division of Graduate Education
  - Victor Piotrowski; Chun-Hsi (Vincent) Huang
- ENG - Directorates of Engineering
  - Joanne Culbertson, ENG/CMMI
  - Ronald Joslin; Christina Payne, ENG/CBET
  - Anthony Kuh, ENG/EECS
- GEO - Directorate for Geosciences
  - Eva Zanzerkia
- MPS - Directorate for Mathematical & Physical Sciences
  - Nigel A. Sharp, MPS/AST; Daryl W. Hess, MPS/DMR; Bogdan Mihaila, MPS/PHY
- SBE - Social Behavioral and Economic Sciences
  - Sara Kiesler and Kenneth C. Land

• Intent: stimulate co-funding between OAC and one or more domains
• Consult OAC + other Cognizant Program Officers
  • At least one month in advance of the submission deadline

Prasad/SCEC-18
NSF-wide Participation

- CISE/OAC - Office of Advanced Cyberinfrastructure – lead
  - Sushil K Prasad
    (Includes BD Hub)
- CISE/CCF Computing and Communication Foundation
  - Almadena Chtchelkanova
- EHR/DGE - Division of Graduate Education
  - Victor Piotrowski; Chun-Hsi (Vincent) Huang
- ENG - Directorates of Engineering
  - Joanne Culbertson, ENG/CMMI
  - Ronald Joslin; Christina Payne, ENG/CBET
  - Anthony Kuh, ENG/EECS
- GEO - Directorate for Geosciences
  - Eva Zanzerkia
- MPS - Directorate for Mathematical & Physical Sciences
  - Nigel A. Sharp, MPS/AST; Daryl W. Hess, MPS/DMR; Bogdan Mihaila, MPS/PHY
- SBE - Social Behavioral and Economic Sciences
  - Sara Kiesler

Intent:
stimulate co-funding between OAC and one or more domains
- Consult OAC + other Cognizant Program Officers
  - At least one month in advance of the submission deadline
FY 19: Award Framework

• Excellent community response
  • 40% additional submissions in 2nd round!
  • About 25 awards made in FY 16 and FY17

• Three project classes:
  • Pilot: Exploratory activities
    • $300K, 2 yrs
  • Implementation: Broadly accessible to community
    • Small: $500K, 4 yrs
    • Medium: foster a community,
      • $1M, 4 yrs
  • Large-scale Project Conceptualization:
    • Planning grants for potential future institute-like CyberTraining projects
    • $500k, 2 yrs

• No separate tracks, still 3 communities of concerns
  • CI Professionals, CI Contributors, and CI Users

• Next Deadline:
  • Feb 6, 2019
  • Webinar on Nov 26
FY 19: Award Framework

• **Excellent** community response
  • 40% additional submissions in 2nd round!
  • About 25 awards made in FY 16 and FY17

• Three project classes:
  • *Pilot:* Exploratory activities
    • $300K, 2 yrs
  • *Implementation:* Broadly accessible to community
    • *Small:* $500K, 4 yrs
    • *Medium:* foster a community,
      • $1M, 4 yrs
  • *Large-scale Project Conceptualization:*
    • Planning grants for potential future institute-like CyberTraining projects
      • $500k, 2 yrs

• No separate tracks, still 3 communities of concerns
  • CI Professionals, CI Contributors, and CI Users

• Next Deadline:
  • Feb 6, 2019
  • Webinar on Nov 26
Developing IEEE-TCPP Parallel/Distributed Curriculum and NSF CyberTraining Program

Sushil K Prasad

National Science Foundation
Georgia State University
Former Chair, IEEE Technical Committee on Parallel Processing (TCPP)

SCEC-18
TCPP Curriculum Initiative:
http://www.cs.gsu.edu/~tcpp/curriculum/