CSE & HPC Education – Within & Beyond Classrooms

Hans-Joachim Bungartz

Technical University of Munich
Dean of Informatics, Chair of Scientific Computing
TUM Graduate Dean
Board member, Leibniz Supercomputing Center (LRZ)
Chairman, German National Research & Education Network (DFN)

HPC User Forum, Austin, 7.9.2016
Preface: Why CSE & HPC together?

Computational Science & Engineering is the field with closest links to HPC

- **HPC drives CSE**: CSE – in particular that at extreme scales and for the famous grand challenges – essentially needs HPC
- **CSE drives HPC**: CSE – i.e. simulation, optimization, analytics, … – has always been a main driving force for HPC as well as a main justification for HPC investments
- CSE has its own programs – with frequent HPC modules
- Computational scientists and engineers are those who will most probably deal with HPC in their future careers, either in academia or industry, either as developers or practitioners/users – at a higher percentage than computer scientists, by the way

**Hence an integral look at CSE & HPC education… and data belongs to that context, too!**

**Focus:**

- University stuff
- HPC centers have a sophisticated course program (for HPC professionals, esp.)
Challenges

Neither CSE nor HPC “belong” to one specific discipline

• No field considers it as its own responsibility (look at the curricular recommendations of other fields’ professional societies, etc., which are missing here)
• If done within classical programs: courses tailored to the respective students’ needs
• If done via new/specific (master) programs: heterogeneous knowledge

No broad consensus on what’s needed – beyond advanced programming skills and some hardware affinity

• Computational X / simulation technology programs come close, but can also be rather “implementation-free” (focus on modelling, e.g.)
• Endless debates on “is it really a discipline?”, “does it really need new programs?”, “shouldn’t we rather just …?”

Raise the sex-appeal of HPC in times of Google & Co.

• “HPC is outdated, all is done by the cloud, …”

Hence: attracting new talents remains random-based and difficult
Recent Developments

Workshop “Future Directions in CSE Education & Research”

• August 2014, Breckenridge, CO
• Some 30+ participants from all over the world
• Co-organized by SIAM (its CSE branch) and EESI (European exascale)
• Goals:
  – In general: positioning – branding – marketing
  – More concrete: update of the report “Graduate Education in CSE” (2001)
• Understanding of “Computational”, esp. w.r.t. HPC & Data Science

Close-to-final version just distributed

(authors: Rüde, Willcox, Curfman McInnes, De Sterck, et al.)

• Cover page
• Table of contents
• Central findings
Just one Figure from that Upcoming Report

Shows the interplay HPC-HW – HPC-Algo – HPC-Apps/Data
⇒ Important for education!

Expected to appear via SIAM in the very new future … so stay tuned!
Contents

Novel programs

Novel courses … in classrooms & labs

Ferienakademie: a summer school in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
Contents

Novel programs

Novel courses ... in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
CSE- and HPC-Related Programs @ TUM

Existing specializations in many classical undergrad & grad programs

Two established “Computational X” master’s programs
• Computational Mechanics – since 2000 ➔ emerged from a domain’s needs
• Computational Science & Engineering – since 2001 ➔ emerged from core technology
• Both are trans-disciplinary (4 or 7 depts., resp.) and international (English language)

One established “honours program”
• Bavarian Graduate School of Computational Engineering – since 2005, with Erlangen
• Additional modules, such as a “Team Software Project”
• Awards the “BGCE Student Paper Prize” at the SIAM CSE meetings – since 2007

Two new data-related master’s programs
• Mathematics in Data Science – starts this fall
• Data Engineering and Analytics – starts this fall
• Both are pillars of TUM’s “Integrative Study Program in Big Data”
Master’s Program in CSE

300–400 applications every year, 25–50 students per year, 90% int’l

4 pillars of education

• *Applied Mathematics* (focus on numerical topics, less modelling)
• *Informatics* (ParComp, HPC, Visualization, Algorithmics, Software Engineering, …)
• *Scientific Computing* (providing the holistic view – it’s not only “A+B”)
• *Applications*: Computational {Fluid Mechanics, Mechanics, Electrodynamics, Physics, Chemistry, Finance, …}

Specialities

• At lot of external master’s projects: industry, research institutions in Munich (Max Planck, Helmholtz, …), universities abroad ➔ sending out your best is a win, not a loss!
• From the very beginning a mandatory tailored software engineering course – trying to increase the level of software professionalism also in HPC
• Graduates very successful in academia and industry
Master’s Program in Data Engineering & Analytics

Starts in October

Driving ideas

• Informatics is the core, but there is more
• Necessity of knowledge in both technology (data bases) and analytics
• Strong links between computing and data analytics

Embedded into TUM’s new “Integrative Study Program Big Data”

• Even more than simulation, data is a topic that has been popping up or pops up in each domain
• Provide an overarching structure for data-related educational activities (like an interdisciplinary research center), reduce the “silo problem” (programs are in departments)
• Different styles of components
  – Complete study programs (such as the two new ones MDS & DEA)
  – Single modules, labs, …
Contents

Novel programs

Novel courses ... in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
Example #1: Interdisciplinary Lab
“Turbulent Flow Simulation on HPC Systems”

Occasion:
• “Ernst-Otto-Fischer Teaching Award”
• Annual TUM-internal competition for new concepts & innovative course formats

Idea: Teamwork across disciplines (Mechanical Engineering & Informatics)

Format: Labs (experiments & simulations) with accompanying lectures

Goals:
• Team-oriented learning
• Cross-disciplinary setting: 2 student sub-groups with different background

Contents:
• Extension of a 3D Navier-Stokes solver with a turbulence model & MPI
• Doing experiments and simulations, comparing resulting data
Example #1: Interdisciplinary Lab
“Turbulent Flow Simulation on HPC Systems”

Implementation:
• Lectures with core facts from both sides
• Small soft-skills sessions
• Supervised team meetings (fixed point-of-contact for student teams, feedback to lecturers)

Reward:
• Well, the idea got the prize …

Further reading:
• Neumann, Kowitz, Schranner, Azyrnykh: HPC Meets Engineering, Euro-Par 2015, pp. 125-134
Example #1: Interdisciplinary Lab
“Turbulent Flow Simulation on HPC Systems”

Statistics:
• 14-20 participants, Informatics / CSE / Mechanical Engineering
• 2 weeks Intro (basic code), 6 weeks Extension (turbulence + MPI; mixed teams), 4 weeks Project (individual or team (preferred) – report up to 60 p.)

Reflection:
• Heterogeneity is a challenge – but it’s feasible, and also enriching
• Very positive feedback & evaluation
Example #2: BGCE Software Projekt

Idea: address the crucial software issue in HPC/CSE

Main features:
- **Team** of 3–8 students – acting as a small SW company
- 6–9 months, 10 ECTS credits
- **Task**: write a complete SW system – from the idea to the product
- **Customer**: typically from industry (Siemens, GE, Zeiss, …)
- **Realistic setting**: roles (project manager, experts), contracting, specification, architectural design, implementation, documentation, presentation, fines for breach of contract, …

Topics so far:
- Computational Steering (‘04), Molecular Dynamics (‘05)
- FSI (‘06), Visualisation, Finance (‘07)
- Grid Computing (‘08), Free surface flow with LBM (‘09)
- Clouds (‘10), MRI goes FRAVE! (‘11)
- Symbolic Reasoning, Seismic Wave Model (‘12)
- Scripting Framework (‘13), MR Spectroscopy (‘14)
- CAD: Integrated Topology Optimization (‘15)
- Data-intensive Distributed Computing Workflows (‘16)

International Visibility: CiSE, 2006, e.g.
Contents

Novel programs

Novel courses … in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
Ferienakademie: (cf. www.ferienakademie.de)

- Est. 1984, TUM + U’s Erlangen & Stuttgart
- Summer school – but students work & present
- 2 weeks, in South Tyrol, during Oktoberfest …
- Funded by industry and by the 3 universities
- Application-based, for our best students
- ~ 12 courses, broad range of topics (even ethics)
- Leisure (hiking) plus intense work

Here (2015 – 2016 similar):

- “Let’s play! Simulated physics for interactive games”
- 22 students (background: MA, CS, CSE, PH, ENG)
- 5 teams (models, simulation, I/O, interaction, parallel programming)
- Project mgt. by students (profs. are advisers)
Format

Technical details:

• Students were provided with template-code and a server for version management
• Server set up beforehand, i.e. installation of libraries, Git-server, documentation, …
• On-site installation of a LAN and WiFi for development
• Limited internet access: preparation in advance necessary
• Assignment of topics for the talks (incl. a supervisor) beforehand
Content

Development of concrete game idea:
• Actually done by the student team in the first 2 days
• Important: realistic pathway to a working prototype at the end (guidance!)

Layout:
• Boat-racing game for two players
• Simulation of flow (water) and rigid bodies incl. coupling (floating obstacles)
• Decision for Lattice Boltzmann for fluid simulation
• Integration of the Bullet game-physics engine for rigid body simulation
• Interesting steering methods, e. g. Xbox Kinect
• Rendering via OpenGL
• AGILE project management implemented by the students
• students give brief tutorials on relevant algorithms, libraries, and tools (i.e.: yes, the classroom is more than inverted!)
Impressions from the Game
Impressions from the Game
Evaluation

Positive:
• Extremely high enthusiasm
• High level of identification with the project → high productivity, deep learning experience
• Willingness to invest a lot of work, even after midnight, at hiking breaks, etc. → high motivation
• Efficiency counts (even if not classical “batch job HPC”); plus software engineering
• Very good team spirit, friendships formed

Negative:
• AGILE transformed to some variant of the good old waterfall
• Over-estimation of available time for development (or under-estimation of work …)

Overall:
• Course very intense due to ambitious project plus all the other activities
Impressions
Contents

Novel programs

Novel courses ... in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
SPPEXA Doctoral Retreat & Coding Weeks

SPPEXA:

- Priority Program “Software for Exascale Computing” of the German Research Foundation (DFG), see www.sppexa.de (nation-wide funding scheme)
- 17 project consortia funded – each inter-institutional and interdisciplinary, some international ones due to a joint call with agencies from France and Japan

Education:

- Has been one core argument for SPPEXA’s installation

Organized training for PhD candidates:

- Annual 1-to-2-week “doctoral retreat”, 12-20 participants
- Practical components (hands-on sessions) plus tutorials
- Networking – foster exchange across the SPPEXA consortia
- Each doctoral candidate within SPPEXA must attend at least one retreat

“The fast and the curious – exploring efficient algorithms on fast hardware”
(organized by P. Neumann, J. Weidendorfer, K. Fürlinger)

**Tutorials:**
- Christian Bischof (automatic differentiation)
- Gerhard Wellein / Georg Hager (performance engineering)

**Intros:** linear solvers, state-of-the-art hardware

**Hands-on:** multigrid for Poisson on Xeon Phi etc.

**Challenge again:** heterogeneous knowledge & expertise (even within an HPC funding program, there are quite different perspectives … – which underlines the necessity of such events)
Impressions from 2014

You see: similar pics as before …

but this time at PhD level 😊
Contents

Novel programs

Novel courses ... in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
Student Cluster Competition 2015 – Facts

9 teams of 6 students before first degree:
- Arizona Tri-University Team (USA)
- Illinois Institute of Technology (USA)
- National Tsing Hua University (Taiwan)
- Northeastern University (USA)
- Pawsey Supercomputing Centre (Australia)
- Technical University of Munich (Germany)
- Tsinghua University (China)
- Universidad EAFIT (Colombia)
- University of Oklahoma (USA)

4 applications + LINPACK + mystery application (HPCG):
- WRF: Weather Research and Forecasting Model (weather forecast)
- MILC: MIMD Lattice Computation (quantum chromodynamics)
- HPC Repast: (agent-based modeling)
- Trinity: (RNA sequencing)
Student Cluster Competition 2015 – Facts

Cluster: power limit of 3120 Watt (2x 1.5kW)

No other hardware restrictions
• If power limit exceeded: sound signal plus malus points
• If exceeded significantly: disqualification (not to destroy SC‘s local power system…)

Award categories:
• LINPACK
• Overall performance

Tasks
• Given scenarios must be simulated within 48 hours; results have to be submitted
• Points awarded for partial results
• Interviews with jury members on the codes and on general HPC issues
• Accompanying program: talks, social events
Team TUMuch Phun

First TUM team … definitely not the last one (SC ‘16 approaching …)

… plus advisers: M. Bader, R. Wittmann, S. Rettenberger
## Team Summaries

<table>
<thead>
<tr>
<th>Team Name</th>
<th>Number of nodes</th>
<th>Cores/node</th>
<th>mem (GB)/node</th>
<th>Total cores</th>
<th>Total mem (RAM)</th>
<th>Accelerators</th>
<th>CPU</th>
<th>Interconnect</th>
<th>Sponsors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Team Desert Heat</td>
<td>7</td>
<td>2</td>
<td>12</td>
<td>24</td>
<td>128</td>
<td>160</td>
<td>896</td>
<td>0</td>
<td>IT Partners, HP, ASU, NAU, UdA</td>
</tr>
<tr>
<td>Universidad EAFIT</td>
<td>10</td>
<td>2</td>
<td>7x14 cores</td>
<td>3x64</td>
<td>7x28 9x64</td>
<td>256</td>
<td>768</td>
<td>0</td>
<td>Matrix, EAFIT</td>
</tr>
<tr>
<td>Illinois Institute of Technology</td>
<td>4</td>
<td>4</td>
<td>16</td>
<td>64</td>
<td>512</td>
<td>256</td>
<td>2048</td>
<td>0</td>
<td>Intel, Argonne National Laboratory, IT</td>
</tr>
<tr>
<td>Northeastern University</td>
<td>3</td>
<td>4</td>
<td>16</td>
<td>64</td>
<td>512</td>
<td>192</td>
<td>1536</td>
<td>0</td>
<td>AMD, Symmetric Computing, Northeastern</td>
</tr>
<tr>
<td>National Tsing Hua University</td>
<td>10</td>
<td>2</td>
<td>16</td>
<td>32</td>
<td>128</td>
<td>320</td>
<td>1280</td>
<td>0</td>
<td>National Center for High Performance Computing, Quanta Cloud Technology, Russia</td>
</tr>
<tr>
<td>Pawsey</td>
<td>9</td>
<td>2</td>
<td>16</td>
<td>32</td>
<td>128</td>
<td>288</td>
<td>1152</td>
<td>4 K40</td>
<td>Cray, Pawsey Supercomputing Center</td>
</tr>
<tr>
<td>Team Diablo</td>
<td>8</td>
<td>2</td>
<td>18</td>
<td>36</td>
<td>128</td>
<td>288</td>
<td>1TB</td>
<td>0</td>
<td>Cray, Lenovo</td>
</tr>
<tr>
<td>Team TUMuch Phun</td>
<td>1</td>
<td>2</td>
<td>14</td>
<td>28</td>
<td>64</td>
<td>28</td>
<td>64</td>
<td>8 Xeon Phi 7120P</td>
<td>Intel, RSC, TUM, CCS</td>
</tr>
<tr>
<td>University of Oklahoma</td>
<td>9</td>
<td>2</td>
<td>12</td>
<td>24</td>
<td>64</td>
<td>216</td>
<td>576</td>
<td>0</td>
<td>IBM, Advisor, Dell, Intel, OU, SanDisk, Mellanox</td>
</tr>
</tbody>
</table>

[http://studentclustercompetition.us/2015/teams.html](http://studentclustercompetition.us/2015/teams.html)

(TUMuch Phun: 128 GB RAM instead of 64 GB RAM)
TUMuch Phun – Hardware Setup

Just one Host Node:
• 2x Intel Xeon E5-2697v3 - 2x14 Cores @ 2.3 – 3.6 GHz
• 128GB RAM
• Intel SSD DC S3710 – 1.2 TB Storage

8x Intel Xeon Phi Coprocessor 7120A/P
• each with 61 Cores @ 1.238 GHz
• each with 16GB GDDR5 RAM

Proprietary Interconnect from RSC (basis: external PCIe 3.0 switch)
• Xeon Phis directly communicate with each other, no detour via host
• Mellanox SB7790 Infiniband EDR Switch
• 2x Mellanox Connect X4 Dual Port EDR Adapter (Ein IB Adapter for 4 Xeon Phis)
• Host communicates via two PCIe 3.0 16x ports with all Xeon Phis via switch
Preparations in Munich …

First of all: find sponsors & supporters
- Hardware: RSC, Russia
- Travel funds: Gauss Centre for Supercomputing & TUM
- Close relations to Intel very helpful (IPCC, …)

Then: find the student team
- High visibility (SC) & attractiveness (trip to Texas …) ➞ excellent students

Get acquainted with the applications (input data, config files, …)

Regular meetings with the team (bi-weekly)

Installation of the codes on the test system (SuperMIC)
- special compiler flags for Xeon Phi architecture
- adapt software packages for Xeon Phi (some were too old)
- in case of Trinity: minimal code changes to switch from assembler optimization to generic optimization
... and in Austin

c.a. 36 hours for system setup and final tests
• first encounter with the final system
• energy tuning: one SSD had to be removed
Preparations in Munich ...
... and on the Way to Austin
Arrived in Austin … Finally

Hardware delivered by RSC just in time …
The SCC Corner

As remote as possible … almost had to take the train to get there …
Finally: the Ceremony!

#1 in LINPACK performance
#3 in overall performance
😊
Résumé

Students in 3rd / 5th semester, not too heterogeneous
• different sub-tasks: system administration, visualization, …
• experiences in competitions given (winner “Bundeswettbewerb Informatik”, e.g.)

Extremely high motivation, despite the big amount of work ( ! )

Technical problem: planned water cooling was not feasible ➔ access to final hardware only in Austin

Last-minute tuning of the cluster by RSC ➔ lack of preparation time with the applications (access one week before trip)

Ongoing hardware problems (system crashes, file system errors, …) required a lot of improvisation / manual interaction

Broad media impact
Contents

Novel programs

Novel courses … in classrooms & labs

Ferienakademie: a project in the mountains

Mountains again: SPPEXA doctoral retreat

Student Cluster Competition @ SC ’15

Concluding remarks
Concluding Remarks

Motivation makes the difference!

More hands-on / project-style training, less frontal-style courses

Go for special events: projects, competitions, hackathons, …

Go for “realism”: collaborations with computing centers, challenges from the applications (academia or industry), or the possibilities via an Intel Parallel Computing Center help a lot

Fight against the CS-spirit “efficient algorithms = sorting & searching & NP-strange stuff” – “we are / HPC is efficient algorithms”

⇒ this will increase HPC attractiveness for young brains

Far increased investment in planning & mentoring … but it pays off
Thanks for your attention!