Progress of the SES-HPC Project at Uni Siegen

Jan Steiner
Zentrum für Informations- und Medientechnologie
Universität Siegen

Erlangen, October 9, 2018
Outline

• Background
• Project Status
• Summary
• Outlook
Outline

• Background
• Project Status
• Summary
• Outlook
Motivation:

- Permeability between Tier 3 and higher tiers
  - Code development on Tier 3 - productive runs on Tier 1
  - Less experienced users
  - Cheaper resources

→ Start early
→ Get people onto Tier 3
→ Facilitate movement to higher tiers
Background: existing infrastructure

- **HorUS**: Current Cluster at Uni Siegen
  - Installed in 2011
  - 134x12 + 20x16 CPUs

- Diverse users

- Future cluster
  - Proposal approved
  - Ca. 250-350 nodes
  - Additional users, e.g. Big Data

→ Good position to support users early
Background: project approach

- “Support for Experienced and Starting HPC Tier 3 Users”

- Support on Tier 3:
  - Starting: basic skills
  - Experienced: performance optimization

- Best practice

- Sustainability
## Project focus areas

<table>
<thead>
<tr>
<th>Teaching and Training</th>
<th>Performance Analysis</th>
<th>Third-party Code Support</th>
<th>Tier Change Support</th>
<th>Knowledge Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner and advanced devs</td>
<td>Experienced code developers</td>
<td>Users of commercial/open-source codes</td>
<td>Dev. teams who want to apply for higher tier hardware</td>
<td>All HPC users</td>
</tr>
<tr>
<td>• Hold classes</td>
<td>• Performance reviews</td>
<td>• Support in finding optimal settings</td>
<td>• Find most suitable hardware</td>
<td>• Establish and maintain wiki</td>
</tr>
<tr>
<td>• Advise on external courses</td>
<td>• Performance measurement tools</td>
<td>• Find most suitable hardware</td>
<td>• Test and evaluation of software</td>
<td>• Organize networking workshops</td>
</tr>
<tr>
<td>• Gauge demand for new courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Outline

• Background

• Project Status

• Summary

• Outlook
Project status (area 1/5)

Teaching and Training

Beginner and advanced devs

- Hold classes
- Advise on external courses
- Gauge demand for new courses

- Interviews with most HPC-related institutes
  - Basics of HPC (jobs, workspaces, ...)
  - Basics of Linux
  - Basics of Fortran, C++ (for HPC purposes)
  - Basics of HPC using Matlab

- Difference to higher tiers: mostly basics

- Parallelization: need rarely explicitly expressed
  - But I met people who would profit
Project status (area 1/5)

Teaching and Training

<table>
<thead>
<tr>
<th>Beginner and advanced devs</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Hold classes</td>
</tr>
<tr>
<td>• Advise on external courses</td>
</tr>
<tr>
<td>• Gauge demand for new courses</td>
</tr>
</tbody>
</table>

• Cluster introduction course
  – Well received
  – Feedback: in English

• Additional courses this semester
  – Introduction to Linux
  – Introduction to C++
Project status (area 2/5)

- Slow progress due to unfilled position
  - Now filled

- I did in-person consulting
  - Individual PhD students
  - Diverse fields
    - Economic CS
    - Linguistics
    - Geomathematics

- Simple tips go a long way ("basics" problem)
Third-party Code Support

Users of commercial/open-source codes

- Support in finding optimal settings
- Find most suitable hardware

- High number of Matlab users
  - Uni Siegen has Matlab Distributed Computing Server license
  - Nobody knows about it
  - Nobody knows how to use it

- Found testers
  - Again from interviews

- Goal: best practice guide
Project status (area 4/5)

- So far no test cases for Tier < 3

- But: several “tier changes” from PC to cluster
  - Again: lack of knowledge
  - Is what I am doing HPC?
Project status (area 5/5)

Knowledge Transfer

All HPC users

- Establish and maintain wiki
- Organize networking workshops

- Little advantage of wiki over website
  - Low number of authors
  - Editing still necessary

- Instead:
  - New cluster website
  - Pipeline for simple content addition (Markdown)

- Networking workshop this semester
Project status (area 5/5)

Knowledge Transfer

• Establish and maintain wiki
• Organize networking workshops

All HPC users

• Cluster website redesign
  – German+English
  – Description of cluster
  – Guides on Linux, SLURM, etc.
  – News, courses
  – “FAQ” section → questions from tickets

• “Getting Started” section - link included in welcome e-mail
• 30k words
Outline

• Background
• Project Status
• Summary
• Outlook
Summary

• Last year: “reach people on social level”

• Routinely meet people who:
  – Have problems but don’t report them
  – Use cluster wrong without knowing
  – Use cluster but we are not aware of them

• Interviews a crucial tool
  – Example: supervisor could not add student (wrong user ID), found oversight in documentation by coincidence during phone call
Summary

- Starting to develop “early warning system”
  - Supervisors know us
  - Tickets (reading between lines)
    • Example: job dies, but `rm slurm*` at end of job script $\rightarrow$ newbie

- Actively approach people

- Visibility increasing through courses and other activities

$\rightarrow$ Social approach starting to work
Outline

• Background
• Project Status
• Summary
• Outlook
Outlook

• New colleague: offer in-depth performance reviews of larger codes

• Additional courses

• Increase visibility further
  – (Networking) workshops
  – Build community

• Performance increase through mastery of basics: \(\rightarrow\) metric?
THANK YOU FOR YOUR KIND ATTENTION.