Progress of the SES-HPC Project at Uni Siegen

Jan Steiner, Jörg Hauser
Zentrum für Informations- und Medientechnologie
Universität Siegen

Paderborn, October 18, 2019
Motivation:

- Permeability between Tier 3 and higher tiers
  - Code development on Tier 3 - productive runs on Tier 1/2
  - 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 userbase
  - Usually first-time HPC users

- Future cluster
  - Additional users, e.g. Big Data

→ Good position to support users early
Upcoming cluster

- 3.5 M€
- 200-300 nodes
- Intended to begin operations in Q1/2020
- Status:
  - Initial offers and negotiation phase complete
  - Requests for final offers sent
- Variety of use cases
  - Classical HPC (CFD etc.)
  - Data Science etc.
  - Worldwide LHC computing grid (WLCG)
- GPU, ARM partitions
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>Type</th>
<th>Focus Area</th>
<th>Details</th>
</tr>
</thead>
</table>
| Teaching and Training       | Beginner and advanced devs                                                  | - Hold classes
                             |                                                                             | - Advise on external courses
                             |                                                                             | - Gauge demand for new courses |
| Performance Analysis        | Experienced code developers                                                 | - Performance reviews
                             |                                                                             | - Performance measurement tools |
| Third-party Code Support    | Users of commercial/open-source codes                                       | - Support in finding optimal settings
                             |                                                                             | - Find most suitable hardware |
| Tier Change Support         | Dev. teams who want to apply for higher tier hardware                       | - Find most suitable hardware
                             |                                                                             | - Test and evaluation of software |
| Knowledge Transfer          | All HPC users                                                               | - Establish and maintain docum.                                        |
                             |                                                                             | - Organize networking workshops |

Oct. 18, 2019
Project status: training and user interaction

• In regular contact with users
  → Overview for what users are doing, and how

• Training program
  – Courses held once per semester
    • Basics of HPC, Linux etc.
  – Continuous evaluation of demand for new courses

• Workshops: exchange experience among more experienced users
  – Matlab Workshop: good basis for Best Practice Guide
Project status: in-person consulting

- Project completed:
  - PhD student, Computer-Assisted Legal Linguistics
  - Analysis of language in law and legal texts
  - Pattern analysis on ~43000 legal documents
  - Java application, connected MySQL database

- Support on:
  - Migration from Windows desktop to HorUS
  - Domain decomposition of intended analysis
  - Job setup
Project status: in-person consulting

• Lessons learned
  – Computational effort was underestimated
  – Used Java universal library structures (e.g. HashMap)
  – Efficiency of using MySQL doubtful

• Conclusions:
  – Users select technologies that make entry easier
    • Available technology expertise among colleagues
  – Technology decisions happen early, are then hard to change
  – Early contact with our support crucial
Project status: documentation and cooperation

- Cluster website growing
  - Users now much more aware of it
  - Integration of Gauss-Allianz HPC Calendar
  - Currently: live display of cluster status (queues etc.)

- Increasing visibility
  - Advice from university press office
  - Publicity due to new cluster

- Looking towards HPC wiki
Project status: implications of new cluster

• Chance to think about things ahead of time
  – Strategy for environment module structure
  – Track use of software
  – Best Practice or at least Lessons Learned for next cluster purchases

• More diverse use cases
  – Reception of new users
  – New course topics (and presenters)

• Larger cluster
  – Reception of new users
Summary

• Training courses well received
  – Demand mostly for beginner-level topics

• Visibility now good among HPC users
  – People who already used cluster
  – Users starting now

• Visibility less good beyond that
  – Cover more general SciComp topics (C++, Git)
  – Advertise university-wide
Summary

• In-person consulting
  – Demand also higher for general topics
  – Few HPC-specific performance tools used

• Best practice guides
  – Matlab
  – To be made available
Outlook

• Additional courses
  – Invite external trainers (C++)
  – More workshops
    • Software engineering

• Increase visibility further
  – Press work (with university’s press office)
  – “Stammtisch” (Regular meetup)

• “Agnostic” Intro to HPC/Parallel Programming course
THANK YOU FOR YOUR KIND ATTENTION.