

# Progress of the SES-HPC Project at Uni Siegen

Jan Steiner

Zentrum für Informations- und Medientechnologie  
Universität Siegen

Erlangen, October 9, 2018

# Outline

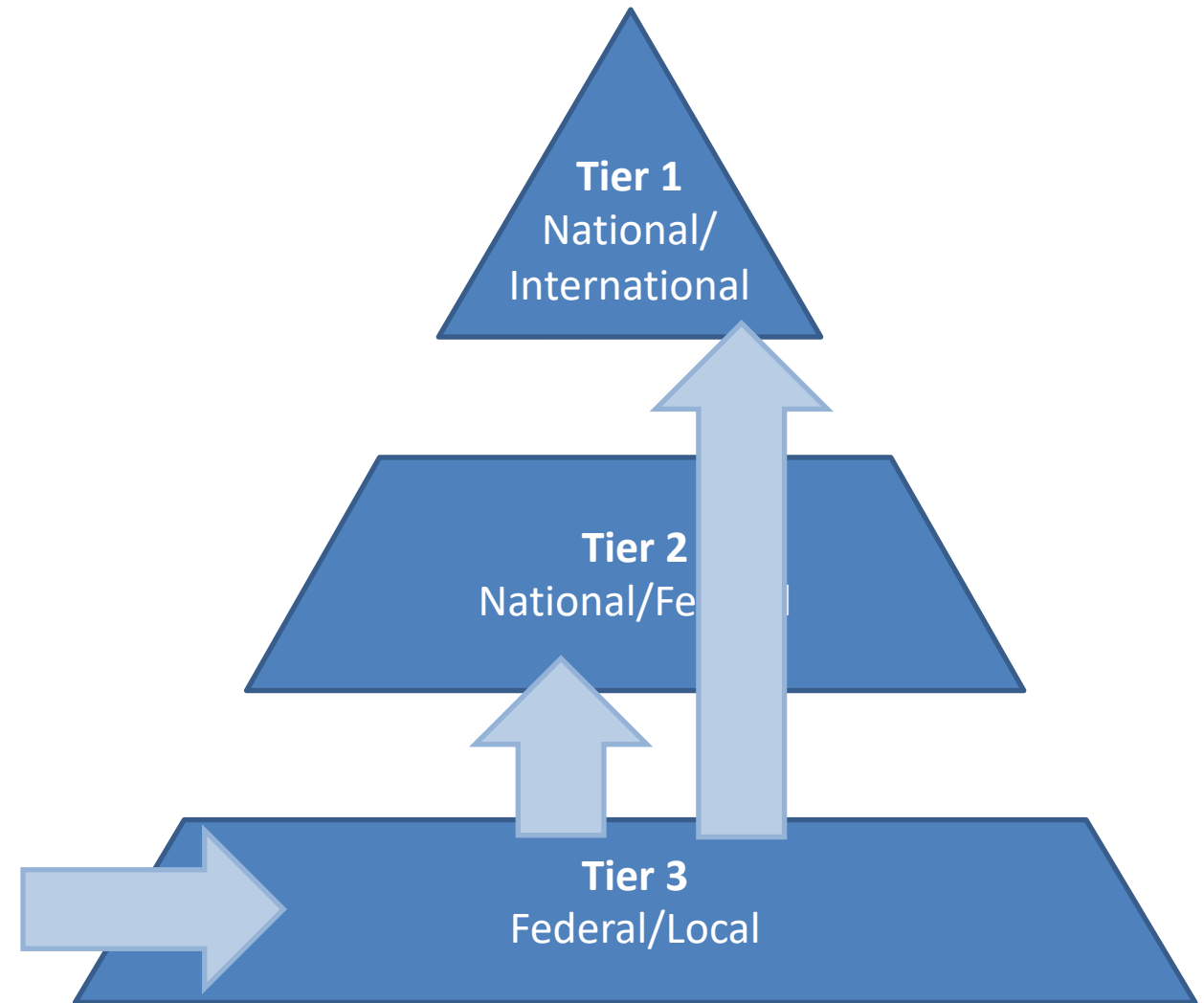
- Background
- Project Status
- Summary
- Outlook

# 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

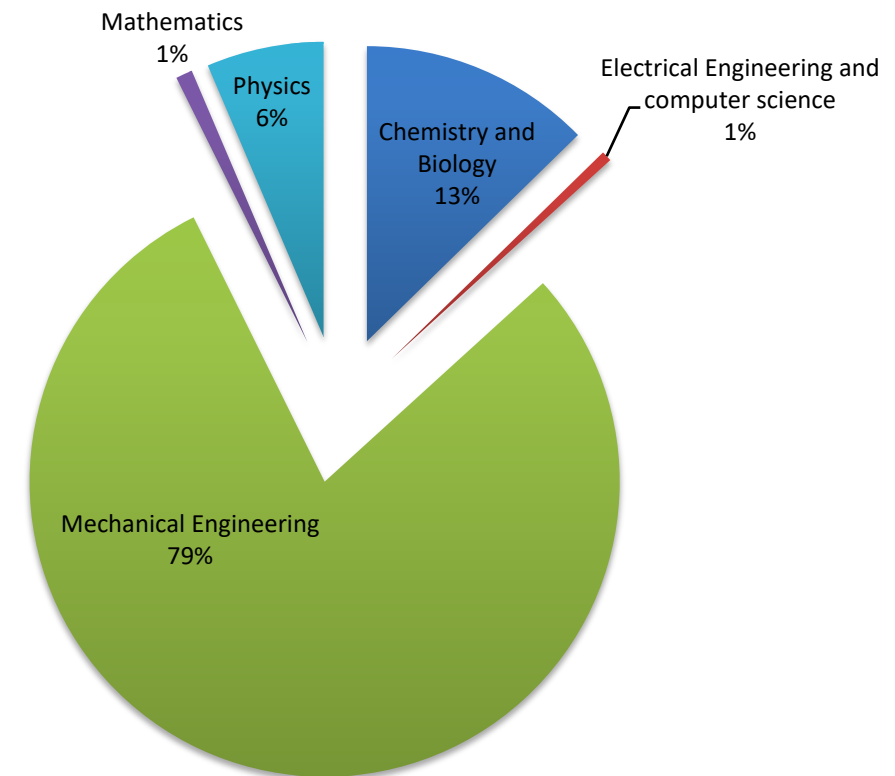


# Background: existing infrastructure

- 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

User groups on the cluster



# 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

Teaching and Training	Performance Analysis	Third-party Code Support	Tier Change Support	Knowledge Transfer
Beginner and advanced devs	Experienced code developers	Users of commercial/open-source codes	Dev. teams who want to apply for higher tier hardw.	All HPC users
<ul style="list-style-type: none"> <li>• Hold classes</li> <li>• Advise on external courses</li> <li>• Gauge demand for new courses</li> </ul>	<ul style="list-style-type: none"> <li>• Performance reviews</li> <li>• Performance measurement tools</li> </ul>	<ul style="list-style-type: none"> <li>• Support in finding optimal settings</li> <li>• Find most suitable hardware</li> </ul>	<ul style="list-style-type: none"> <li>• Find most suitable hardware</li> <li>• Test and evaluation of software</li> </ul>	<ul style="list-style-type: none"> <li>• Establish and maintain wiki</li> <li>• Organize networking workshops</li> </ul>

# 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

Beginner and advanced devs

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

- Cluster introduction course
  - Well received
  - Feedback: in English
- Additional courses this semester
  - Introduction to Linux
  - Introduction to C++

## Project status (area 2/5)

### Performance Analysis

Experienced code developers

- Performance reviews
- Performance measurement tools

- 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)

## Project status (area 3/5)

### 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

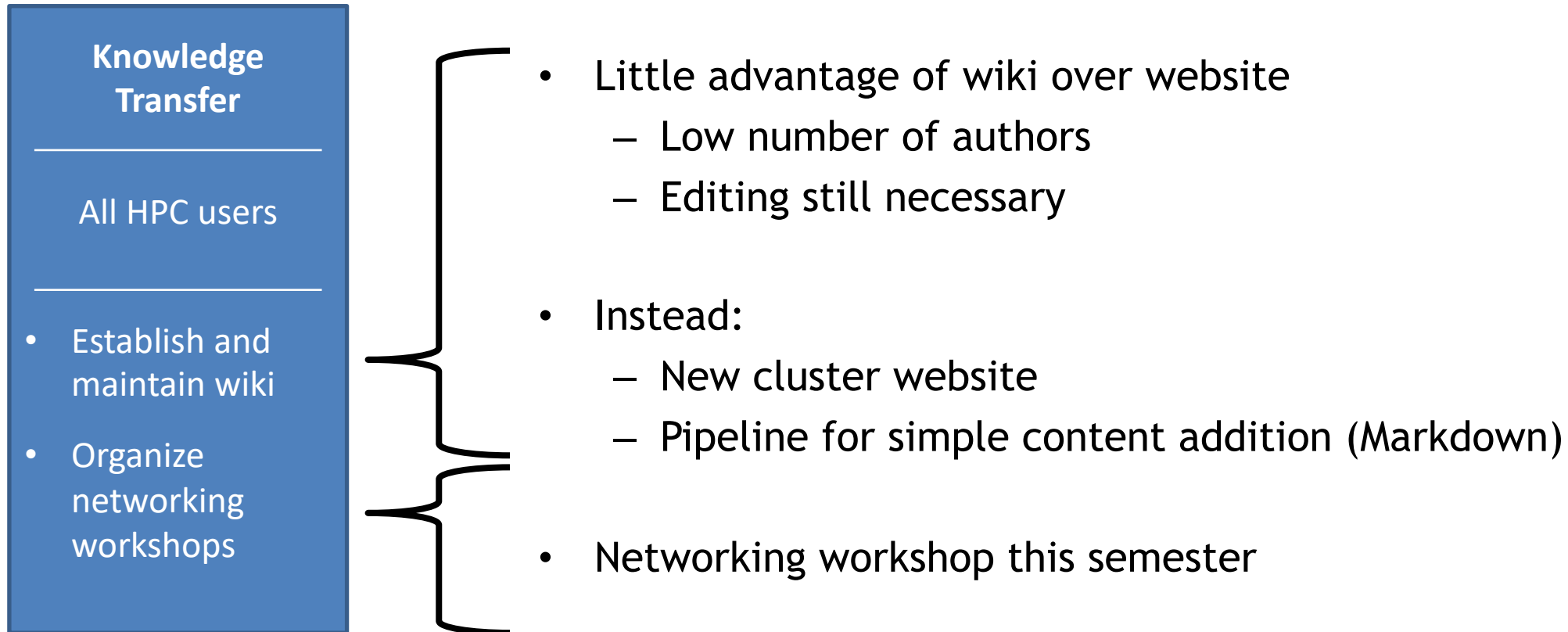
# Project status (area 4/5)

## Tier Change Support

Dev. teams who want to apply for higher tier hardw.

- Find most suitable hardware
  - Test and evaluation of software
- 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)



## Project status (area 5/5)

### Knowledge Transfer

All HPC users

- Establish and maintain wiki
- Organize networking workshops

- 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 → newbie
  - Actively approach people
  - Visibility increasing through courses and other activities
- 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: → metric?

**THANK YOU FOR YOUR KIND ATTENTION.**