Software Productivity and Sustainability track

Anshu Dubey, David E. Bernholdt, Greg Becker, and Jared O’Neal

Argonne Training Program on Extreme-Scale Computing summer school

Slides, late-breaking updates, etc. at: https://bssw-tutorial.github.io/ and click the link for today’s tutorial
License, Citation and Acknowledgements

License and Citation

• This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).


• Individual modules may be cited as Speaker, Module Title, in Tutorial Title, …

Acknowledgements

• This work was supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research (ASCR), and by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.

• This work was performed in part at the Argonne National Laboratory, which is managed by UChicago Argonne, LLC for the U.S. Department of Energy under Contract No. DE-AC02-06CH11357.

• This work was performed in part at the Lawrence Livermore National Laboratory, which is managed by Lawrence Livermore National Security, LLC for the U.S. Department of Energy under Contract No. DE-AC52-07NA27344.

• This work was performed in part at the Los Alamos National Laboratory, which is managed by Triad National Security, LLC for the U.S. Department of Energy under Contract No.89233218CNA000001

• This work was performed in part at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract No. DE-AC05-00R22725.

• This work was performed in part at Sandia National Laboratories. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.
About Us

• Anshu Dubey, ANL
• David E. Bernholdt, ORNL
• Greg Becker, LLNL
• Jared O’Neal, ANL

• Current of former members of the IDEAS Productivity Project: http://ideas-productivity.org

• Or the Spack project (Greg): https://spack.io
The IDEAS-ECP team works with the ECP community, and beyond, to improve developer productivity and software sustainability as key aspects of increasing overall scientific productivity.

1. Customize and curate methodologies
   - Target scientific software productivity and sustainability
   - Use workflow for best practices content development

2. Incrementally and iteratively improve software practices
   - Determine high-priority topics for improvement and track progress
   - *Productivity and Sustainability Improvement Planning (PSIP)*

3. Establish software communities
   - Determine community policies to improve software quality and compatibility
   - Create Software Development Kits (SDKs) to facilitate the combined use of complementary libraries and tools

4. Engage in community outreach
   - Broad community partnerships
   - Collaboration with computing facilities
   - Webinars, tutorials, events
   - *WhatIs* and *HowTo* docs
   - Better Scientific Software site ([https://bssw.io](https://bssw.io))

For more about our work see this report: [https://doi.org/10.2172/1606662](https://doi.org/10.2172/1606662)
Building an Online Community

https://bssw.io

- New community-based resource for scientific software improvement
- A central hub for sharing information on practices, techniques, experiences, and tools to improve developer productivity and software sustainability for computational science & engineering (CSE)

Goals

- Raise awareness of the importance of good software practices to scientific productivity and to the quality and reliability of computationally-based scientific results
- Raise awareness of the increasing challenges facing CSE software developers as high-end computing heads to extreme scales
- Help CSE researchers increase effectiveness as well as leverage and impact
- Facilitate CSE collaboration via software in order to advance scientific discoveries

Site users can…

- Find information on scientific software topics
- Contribute new resources based on your experiences
- Create content tailored to the unique needs and perspectives of a focused scientific domain
Follow IDEAS and BSSw

• IDEAS Productivity mailing list: http://eepurl.com/cQCyJ5
  – Announcements of IDEAS-organized events
    • Best Practices for HPC Software Developers webinar series
    • Strategies for Working Remotely panel series
    • Software-focused events at major scientific meetings (e.g., SIAM, ISC, SC, etc.)
  – Typically 2-3 messages per month

• BSSw Digest: https://bssw.io/pages/receive-our-email-digest
  – Updates on BSSw content
    • New blog posts, events, and resources
    • BSSw Fellowship
  – Typically 1-2 messages per month
  – Also: RSS feed: https://bssw.io/items.rss
The Importance of Naming

• Computing is rife with terminology that many consider harmful and exclusionary
  – Examples include: whitelist/blacklist, master/slave, and master (standalone)

• We support efforts to replace such language with more inclusive language

• In this tutorial, we strive to use inclusive language
  – Example: we use “main” for the default git branch, even where outside sources we reference may use “master”

• We welcome suggestions for further improvements in our tutorial

• Additional information:
  – The Inclusive Naming Initiative
  – The BSSw.io resource on inclusive naming provides some additional context and links
BSSw Tutorial Web Site

- **https://bssw-tutorial.github.io/** is the one URL you need to find all the resources for this tutorial
- Each tutorial event has its own page
- Each tutorial page is considered archival
  - All of the materials used in that tutorial (or links to them)
  - Materials may be updated if we find mistakes
Explaining Slide 2

• Slide 2 in all of our presentations contains the license, citation, and acknowledgements for the tutorial

• (Software) best practice to make your license and preferred citation(s) easily findable

• Sponsor acknowledgements rarely hurt!

License, Citation and Acknowledgements

License and Citation

- This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY 4.0).
- The requested citation the overall tutorial is: David E. Bernholdt, Anshu Dubey, Patricia A. Grubel, Rinku K. Gupta, and Gregory R. Watson, Better Scientific Software tutorial, in the International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC21), St. Louis, MO, USA and online, 2021. DOI: 10.6084/m9.reshare.14556628
- Individual modules may be cited as Speaker, Module Title, in Better Scientific Software tutorial...

Acknowledgements

- This work was supported by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research (ASCR), and by the Exascale Computing Project (17-SC-25-SC), a collaborative effort of the U.S. Department of Energy Office of Science and the National Nuclear Security Administration.
- This work was performed in part at the Argonne National Laboratory, which is managed by UChicago Argonne, LLC for the U.S. Department of Energy under Contract No. DE-AC02-06CH11357.
- This work was performed in part at the Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.
- This work was performed in part at the Lawrence Livermore National Laboratory, which is managed by Lawrence Livermore National Security, LLC for the U.S. Department of Energy under Contract No. DE-AC52-07NA27344.
- This work was performed in part at the Los Alamos National Laboratory, which is managed by Triad National Security, LLC for the U.S. Department of Energy under Contract No. 89233218CNA000001.
- This work was performed in part at Sandia National Laboratories. Sandia National Laboratories is a multi-mission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525.
We Want to Interact with You!

- We find these tutorials most interesting and informative (for everyone) if you ask questions and share experiences!
  - We learn too!
- Please raise your hand at any time to ask a question
- After the tutorial email us at bssw-tutorial@lists.mcs.anl.gov
  - With questions or feedback
  - The list moderator will allow your messages to be posted
- Refer to bssw-tutorial.github.io page for all tutorial materials
## Agenda (morning)

<table>
<thead>
<tr>
<th>Time (CDT)</th>
<th>Title</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 AM</td>
<td>Introduction</td>
<td>David E. Bernholdt (ORNL)</td>
</tr>
<tr>
<td>8:35 AM</td>
<td>Motivation and Overview of Best Practices in HPC Software Development</td>
<td>David E. Bernholdt (ORNL)</td>
</tr>
<tr>
<td>9:15 AM</td>
<td>Scientific Software Design</td>
<td>Anshu Dubey (ANL)</td>
</tr>
<tr>
<td>10:00 AM</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>10:30 AM</td>
<td>Spack: Package Management for HPC</td>
<td>Greg Becker (LLNL)</td>
</tr>
<tr>
<td>11:30 AM</td>
<td>Spack Hands-On</td>
<td>Greg Becker (LLNL)</td>
</tr>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
<td></td>
</tr>
</tbody>
</table>

The agenda is also available on the tutorial web page. Visit [https://bssw-tutorial.github.io](https://bssw-tutorial.github.io) and click on the link for today’s tutorial.
# Agenda (afternoon)

<table>
<thead>
<tr>
<th>Time (CDT)</th>
<th>Title</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>12:30 PM</td>
<td>Lunch</td>
<td></td>
</tr>
<tr>
<td>1:30 PM</td>
<td>Software Testing and Verification</td>
<td>Anshu Dubey (ANL)</td>
</tr>
<tr>
<td>2:30 PM</td>
<td>Refactoring Scientific Software</td>
<td>Anshu Dubey (ANL)</td>
</tr>
<tr>
<td>3:00 PM</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>3:30 PM</td>
<td>Lab Notebooks for Computational Mathematics, Sciences, &amp; Engineering</td>
<td>Jared O'Neal (ANL)</td>
</tr>
<tr>
<td>4:30 PM</td>
<td>Managing Computational Experiments</td>
<td>Anshu Dubey (ANL)</td>
</tr>
<tr>
<td>4:45 PM</td>
<td>Improving Reproducibility Through Better Software Practices</td>
<td>David E. Bernholdt (ORNL)</td>
</tr>
<tr>
<td>5:45 PM</td>
<td>Summary</td>
<td>David E. Bernholdt (ORNL)</td>
</tr>
<tr>
<td>6:00 PM</td>
<td>Q&amp;A, including your software experiences and challenges</td>
<td>All</td>
</tr>
<tr>
<td>6:30 PM</td>
<td>Adjourn</td>
<td></td>
</tr>
</tbody>
</table>

The agenda is also available on the tutorial web page. Visit [https://bssw-tutorial.github.io](https://bssw-tutorial.github.io) and click on the link for today’s tutorial.