



Summary



David E. Bernholdt (he/him)
Oak Ridge National Laboratory



Better Software for Reproducible Science tutorial @ SC23


Contributors: David E. Bernholdt (ORNL), Patricia A. GruContributors:
David E. Bernholdt (ORNL), Anshu Dubey (ANL), Patricia Grubel, David
M. Rogers (ORNL)



See slide 2 for
license details

License, Citation and Acknowledgements

License and Citation

- This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0). 
- **The requested citation the overall tutorial is:** David E. Bernholdt, Patricia A. Grubel, David M. Rogers, and Gregory R. Watson, Better Software for Reproducible Science tutorial, in The International Conference for High-Performance Computing, Networking, Storage, and Analysis (SC23), Denver, Colorado, 2023. DOI: [10.6084/m9.figshare.24226105](https://doi.org/10.6084/m9.figshare.24226105).
- 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-AC05-00OR22725.
- 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.

Science through computing is,
at best,
as credible as the software that produces it!

Today, We Covered Many Topics...

- Designing software for flexibility and extensibility
- Refactoring software
- Collaborative software development
- Software packaging
- Testing strategies for complex software systems
- Reproducibility
- Lab notebooks and managing computational experiments

And there are Many More We Didn't Have Time For

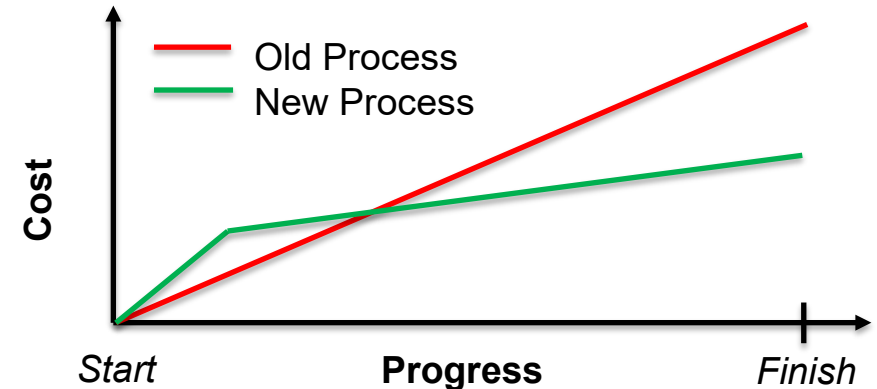
- Licensing
 - Continuous integration testing
 - Distribution
 - Issue tracking
 - Configuration and building
 - Debugging strategies
 - Building and sustaining communities around software
 - Software publication and citation
 - Packaging
 - Requirements gathering
 - Understanding and debugging floating-point math
 - Performance and performance portability
 - ...
- **Also important topics, but...**
 - Less distinction between research software and other software
 - More informational resources available
 - Next-level concerns for starting researchers
 - There's only so much time in the day!

But you're a researcher.
You can't afford to spend
“all” of your time on
software engineering.

A Final Recommendation: Continual, Incremental Software Process Improvement

Target: your project should include “just enough” software engineering so that you can meet your short-term and longer-term scientific goals effectively

1. Identify your team’s “pain points” in your software development processes
 - Help: RateYourProject assessment tool:
<https://rateyourproject.org/>
 2. Set a goal for something to improve
 - Target processes and behaviors, not just tasks
 - Pick something that you can address in a few months that will give you a noticeable benefit
 3. Agree on a plan to address it, identify markers of progress and what is “done”
 - Write them down
 - Help: Progress tracking card examples:
<https://bssw-psip.github.io/ptc-catalog/catalog>
 4. Work your plan, track your progress
 5. When you are done, celebrate...
- 7 ...then pick a new pain point to address



The new process costs something to implement, but it pays off over time

Productivity and Sustainability Improvement Planning

<https://bssw.io/psip>



A goal of [BSSw.io](https://bssw.io) is to provide resources for improving your software processes. If you find useful resources that aren't on BSSw.io, consider contributing. Its easy and quick.

IDEAS
productivity

ECIP
EXASCALE
COMPUTING
PROJECT

Thanks, and Keep in Touch!

- Email comments and questions to bssw-tutorial@lists.mcs.anl.gov
- See the tutorial web site for an archive of all the materials from this tutorial
 - <https://bssw-tutorial.github.io/>
- Follow us:
 - IDEAS Productivity project
 - Website: <https://ideas-productivity.org/>
 - Announcement list: <http://eepurl.com/cQCyJ5>
 - Better Scientific Software
 - Resources: <https://bssw.io/>
 - Monthly digest: <https://bssw.io/pages/receive-our-email-digest>
 - RSS feed: <https://bssw.io/items.rss>

You may also be interested in these other software-related events at SC23:
<https://bssw.io/events/sc23-software-related-events>
(link is also on tutorial web page)

