Reflections on Building Portable Benchmarking Frameworks

ICS 2026 Workshop on Reproducible and Sustainable HPC Benchmarking

Tuomas Koskela
UCL ARC

2026-07-06

Who am I?

- I am a research software engineer at UCL Advanced Research Computing Centra (ARC)
- I lead a team of HPC Research Software Engineer
  - We develop HPC research applications and tools
  - We work on research grants with researchers
  - We usually are not the main developers of the applications
- I don’t run an HPC system

Motivation

We all Want Automatic Benchmarks

- Developers
- Researchers
- Sysadmins
- Procurements

Ingredients for Automated Benchmarking

- Figures of Merit
- Wrapping around build systems
- Templating for scheduler scripts
- Output parsing
- Logging

Disclosure: I did this too

Excalibur-tests project

- Driven by ExCALIBUR H&ES and research at Bristol and Met Office
- https://ukri-excalibur.github.io/excalibur-tests/
- Glue code for Spack and ReFrame
- Clear abstraction of platforms and benchmarks
- Support wide range of platforms and benchmarks
- Engage with local admins to maintain environments

Principles for Automated and Reproducible Benchmarking

Feedback from the UK Community

- We gave a number of talks, workshops and hackthons about benchmarking
- Received positive feedback from workshop attendees and collaborators
- Users:
  - 1 team of RSE's (UCL)
  - 1 research group (Bristol)
  - 1 team of HPC admins (DiRAC)

What are we Doing Now?

UKRI Living Benchmarks Project

Other Use Cases

Reflections

- Portability is hard and takes continuous maintenance
  - Funding software maintenance is challenging
- Institutional and individual buy-in are both important
  - Portability across institutions is also challengin
- The tools you use don’t matter, as long they enable reproducibility

Thank you for listening!