A Portable Approach for Repeatability Testing

Quan Pham[1], Tanu Malik[2], Ian Foster[1,2]
Department of Computer Science[1,§] and Computation Institute[2,¶]
University of Chicago[§,¶] and Argonne National Laboratory[¶]

Introduction

Increasingly conference committee and journal editors are encouraging authors to submit their code, data and software for repeatability testing. Repeatability testing improves peer review

1. Allows reviewers to not only read the ideas in the paper, but validate them by running the accompanying software.
2. Allows reviewers to work the software for different kinds of data and parameters to check robustness of the ideas.
3. Allows reviewers to determine the limitations and assumptions in the ideas.

Testers often want to selectively repeat portions of an experiment that may be of interest without having to go through the time-consuming process of repeating the experiment in its entirety.

Provenance-To-Use (PTU)

PTU is a tool for reducing time for repeatability testing.
Authors can use *ptu-audit* to accomplish two tasks:

1. Build a package of their source code, data, and environment variables
2. Store process and file level details about a reference execution of their system in an accompanying database.

Reviewers can use *ptu-exec* to reduce testing time:

1. Explore the provenance graph to specify the part to re-execute or replay
2. Re-run the experiment partially

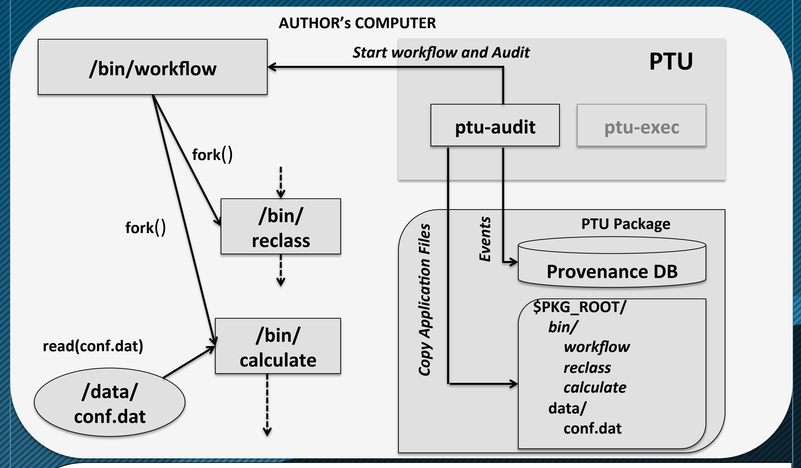
Implementation Challenge

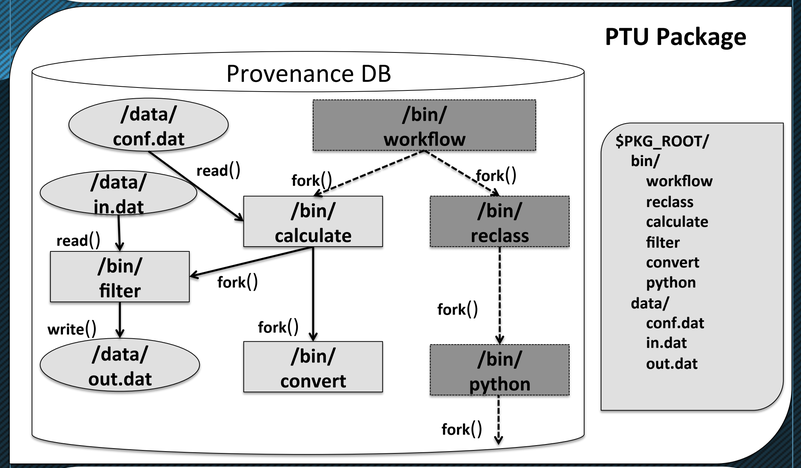
PTU uses a tracing mechanism on authors' experiment to capture codes, data files, and environment variables with detailed provenance to create a PTU package.

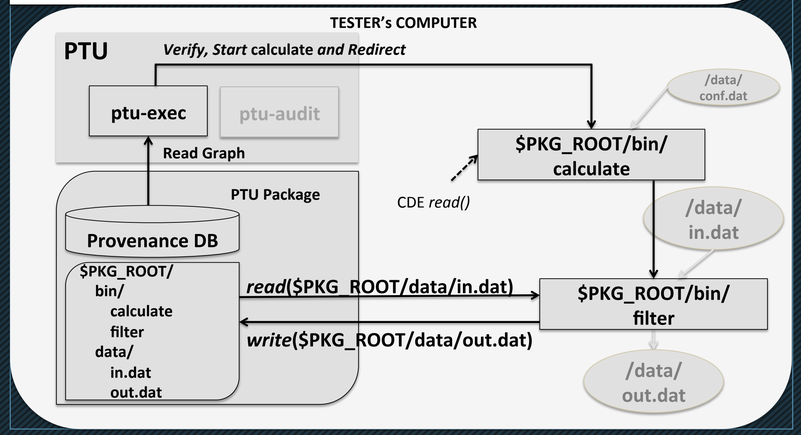
User-space	Kernel-based
• Lower performance	• Better performance
• No special privilege	• Requires root privilege
• More portable	• Restricted environment

To provide portability to PTU package, we use ptrace tracing mechanism.

PTU Tool

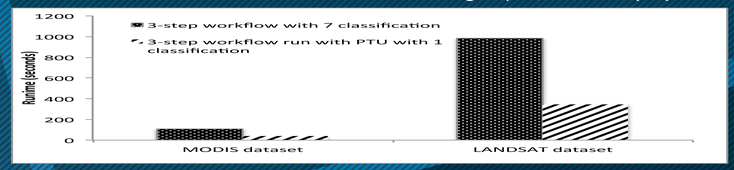




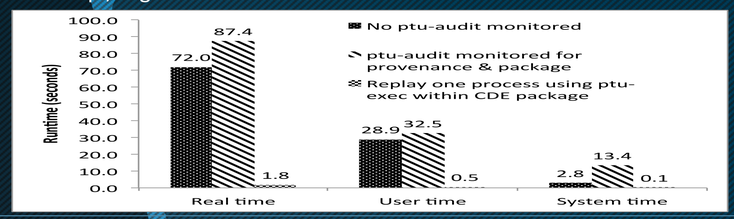


Use Cases

Best, N., et. al., *Synthesis of a Complete Land Use/Land Cover Dataset for the Conterminous United States.* RDCEP Working Paper, 2012. **12(08).**



Murphy, J., et. al., *Textual Hydraulics: Mining Online Newspapers to Detect Physical, Social, and Institutional Water Management Infrastructure,* 2013, Tech. Rep., Argonne National Lab.



Conclusions

- PTU is a step toward testing software programs that are submitted to conference proceedings and journals to conduct repeatability tests
- Easy and attractive for authors
- Fine control, efficient way for testers

Related Work

- CDE uses ptrace to help authors package the code, data, and environment for running and deploying on other machines.
- ReproZip uses kernel-based tracing SystemTap to capture detailed provenance and combine into a package.

References

- Pham, Q., et al., *SOLE: linking research papers with science objects*, in IPAW, 2012.
- Guo, P. et. al., *CDE: Using System Call Interposition to Automatically Create Portable Software Packages*, USENIX, 2011.
- Chirigati, F., et al., *ReproZip: using provenance to support computational reproducibility*, in TAPP, 2013