ReproZip
Packing Experiments for Sharing and Publication

Fernando Chirigati, Juliana Freire | NYU-Poly
Dennis Shasha | NYU

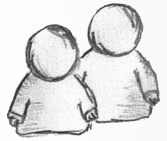
Motivation

- Published articles are not made reproducible
- Computational reproducibility may be **difficult** to achieve

Author

How to encapsulate my experiment?
Too many dependencies…
Too many files to keep track…
Sigh.

How to compile this program?
How to execute it?
How to *explore* it?
Sigh.



Reviewers
Collaborators

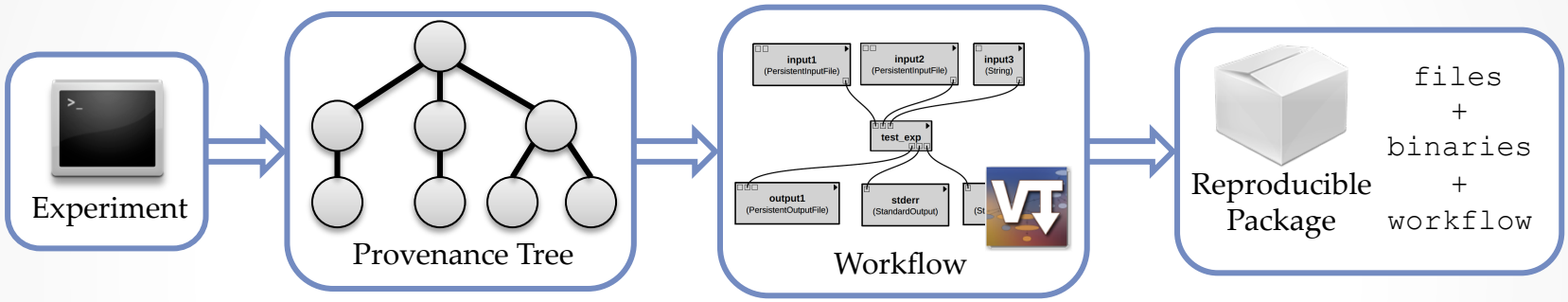
- Some current solutions require the user to adopt a system
  - GenePattern [1], Madagascar [2], Scientific Workflow Systems [3]
- Other solutions rely on capturing information about the computational environment
  - Virtual Machines
  - CDE [4]

ReproZip

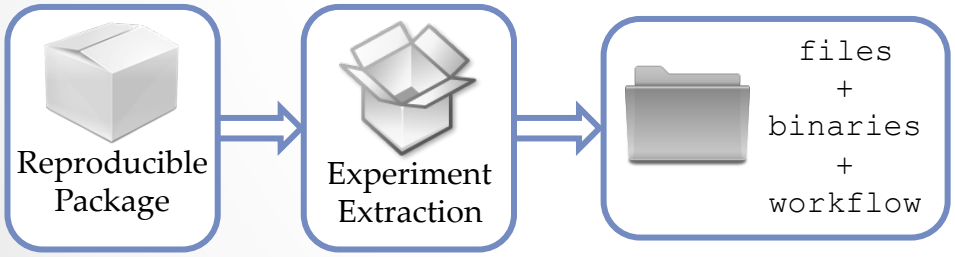
- ReproZip is a packaging solution
  - It makes it easier for *authors* to pack experiments and for *reviewers* to verify computational results
- It creates *reproducible packages* from existing experiments on computational environment *E*
  - No need to port experiments to other system
  - Leverages *provenance* of computational results
- It unpacks an experiment on computational environment *E'*
- It generates a *workflow specification* that encapsulates the execution of the experiment
  - Eases the verification process
  - Allows users to explore the experiment, while keeping track of provenance

Overview

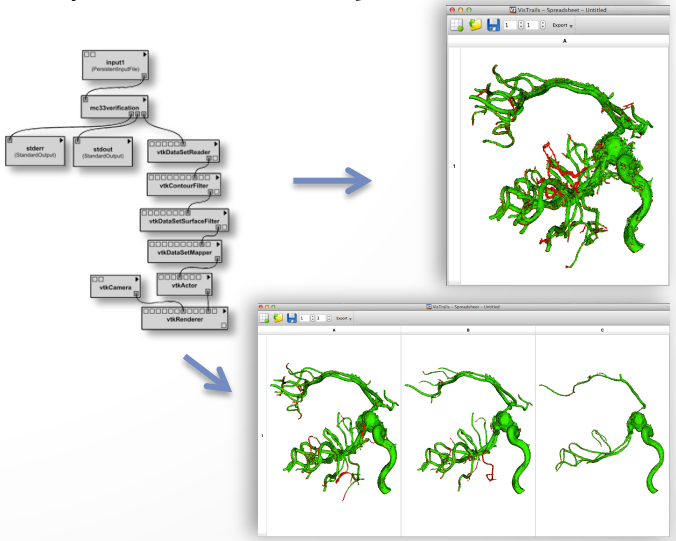
packing (on environment E)



unpacking (on environment E′)



verification and exploration



References

1. GenePattern. *http://www.broadinstitute.org/cancer/software/genepattern/*
2. Madagascar. *http://www.ahay.org/wiki/Main_Page*
3. S. B. Davidson and J. Freire. Provenance and scientific workflows: challenges and opportunities. In SIGMOD, pages 1345-1350, 2008
4. P. Guo. CDE: A Tool for Creating Portable Experimental Software Packages. Computing in Science and Engineering, 14(4):32-35, 2012
5. SystemTap. http://sourceware.org/systemtap/
6. MongoDB. http://www.mongodb.org/

Thank You!

Fernando Chirigati
fchirigati@nyu.edu
http://vgc.poly.edu/~fchirigati