

# Towards turnkey reproducibility

**Geoffrey M. Oxberry**

Lawrence Livermore National Laboratory  
Computational Engineering Division  
Energy Conversion and Storage

This work performed under the auspices of the U.S. Department of Energy  
by Lawrence Livermore National Laboratory under Contract  
DE-AC52-07NA27344.

December 13, 2012

# Problem: Reproducing someone's work can be hard

- Need to install necessary software (assume open source)
  - ▶ Takes time, expertise, patience, privileges
  - ▶ Could affect system stability
- Could wrap source in VM (virtual machine) image
  - ▶ Usually requires  $\geq 300$  MB to host; big, unwieldy
  - ▶ No separation of source code & environment means no flexibility
- High barrier means **people don't run the code**
- **Lower hosting & time barrier** by specifying environment in separate repo using configuration management software

# Solution: Specify environment with configuration management software

- Config management tools specify config in text files
  - ▶ Shell scripts (simplest, fewest prepackaged features)
  - ▶ Puppet ([puppetlabs.com](http://puppetlabs.com))
  - ▶ Chef ([wiki.opscode.com](http://wiki.opscode.com))
  - ▶ Fabric (<http://docs.fabfile.org/en/1.5/>)
  - ▶ Related: Hashdist, Blueprint, Rezip, others...
- Instantiate config using virtualization tools
  - ▶ Serial, small parallel jobs: Vagrant ([vagrantup.com](http://vagrantup.com)) + VirtualBox ([virtualbox.org](http://virtualbox.org))
  - ▶ StarCluster ([star.mit.edu/cluster](http://star.mit.edu/cluster))
  - ▶ CloudFormation ([aws.amazon.com/cloudformation](http://aws.amazon.com/cloudformation))
  - ▶ Any other virtualization software + hardware
  - ▶ Use web services instead (like Wakari, RunMyCode)
- Idea is **flexibility**: pick & choose (even none, mix)

# Example: Install Python interface for DASSL

- DASSL: differential-algebraic equation solver package in Fortran (L. Petzold)
- PyDAS: Python interface to DASSL (J. W. Allen, on GitHub)
- Example: Solve Robertson problem in IPython notebook using PyDAS
- Presentation, environment and source repos on <https://github.com/goxberry> (all labeled with ICERM-2012)
  - ▶ Requires Vagrant + VirtualBox
  - ▶ Vagrantfile to specify VM to create (here, Ubuntu 12.04)
  - ▶ Configuration in Puppet
  - ▶ README with directions for running software

# Acknowledgments

- Dr. Matt McNenly
- Dr. Dan Flowers
- Dr. David I. Ketcheson
- Dr. Aron Ahmadi
- US DOE
- DOE CSGF
- Gurpreet Singh, program manager for the DOE EERE Advanced Combustion Engine Program, for his continued support of the Advanced Combustion Numerics project at LLNL

