Maintaining Quality and Confidence in Open-Source, Evolving Software: Lessons Learned with PFLOTRAN

Jennifer M. Frederick & Glenn E. Hammond
jmfrede@sandia.gov gehammo@sandia.gov

www.pflotran.org

Sandia National Laboratories

National Nuclear Security Administration

U.S. DEPARTMENT OF ENERGY

NS41B-0012

Introduction

soft·ware ev·o·lu·tion
ˈsôf(t)wer/ /ˌevəˈlo͞oSH(ə)n/

Standard definition:

The gradual development of code, from a simple to a more complex form, due to ***repeated*** improvements and updates.

Reasons for software evolution

New domain science:
- New process models
- Increasingly mechanistic process models
- New programming paradigms
- New numerical methods for solution

New computational science:
- Change in third-party libraries
- Switching operating systems or programming language
- New or changing computer hardware

soft·ware qual·i·ty as·sur·ance
ˈsôf(t)wer/ ˈkwäləd͟ē/ əˈSHo͝orəns/

IEEE standard:

A planned and systematic pattern of all actions necessary **to provide adequate confidence** that software conforms to established technical requirements.

Software quality includes

- Correctness (validation and verification)
- Reliability
- Maintainability
- Testability
- Availability
- Portability
- Survivability
- Efficiency

When software is constantly evolving, how can we ensure software quality?

Open Source Development

Software Configuration Management

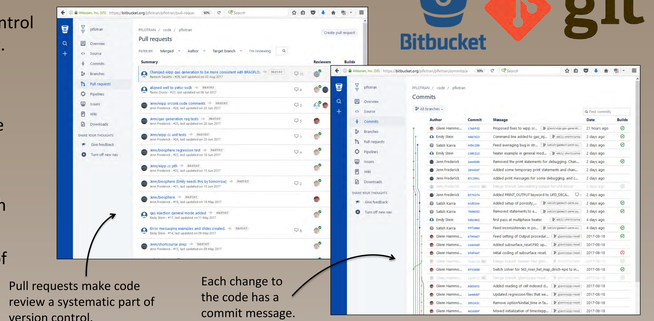
Modular Object-Oriented Design

Automated Testing Suites

Online Version-Controlled Documentation

Software Configuration Management

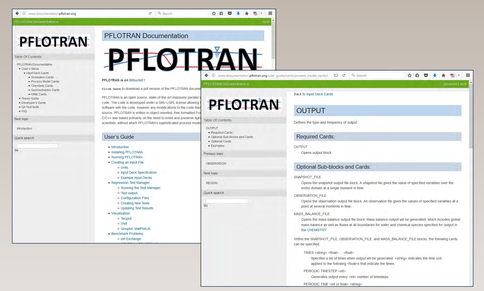
- PFLOTRAN employs the Git distributed source control management tool for configuration management.
- Git logs all changes to a code repository
  - Version control
  - Code can be rolled back if a mistake was made
- Git allows developers to:
  - Clone the base repository
  - Modify and test code in a development branch
  - Merge changes back into base repository
  - Pinpoint problematic changesets (snapshots of code versions)
- Hosted on Bitbucket.com



Pull requests make code review a systematic part of version control.

Each change to the code has a commit message.

Online Documentation



- PFLOTRAN uses a documentation generator program called Sphinx (www.sphinx-doc.org)
- The documentation consists of text files and images, written in restructured text, and organized with an index
- The documentation is version controlled and the repository is hosted on Bitbucket.com
  - When/if you roll back the code, you can roll back the documentation too!
- Sphinx creates html files as well as LaTeX -> pdf
- We host the html files on our documentation website http://documentation.pflotran.org
- The pdf User's Guide and Theory Manual can be downloaded or printed, and never falls behind the online documentation

http://www.sphinx-docs.org

Benefits of Open Source Software

- Encourages collaboration
  - Development, testing, debugging can be shared
- Transparency exposes implementation details critical to scientific reproducibility, often excluded by scientific journals
- More optimal use of funding
  - Funding can be pooled across diverse sets of projects/budgets
  - Rather than pay for licensing fees, pay for additional development
  - Infinite benefit to those unfunded (graduate students)
- Community contributions can drive the code to evolve beyond original vision

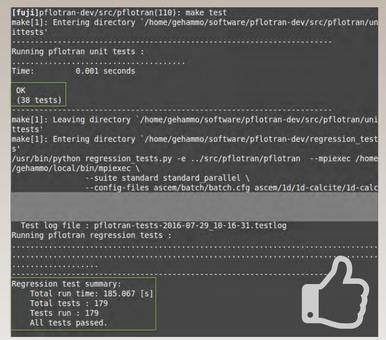
LGPL Free Software Free as in Freedom

- The original or modified source code may not be sold for profit
- Third-party software linked or wrapped around PFLOTRAN may be proprietary

Automated Testing Suites

- As open source development fosters a growing community of developers, the code should not break!
- Automated testing ensures the code works properly, but it is dependent on having good code coverage.
- **Unit tests**
  - Individual routines are executed in isolation.
  - Results are compared with a gold standard to *within a tolerance*.
- **Regression tests** – focus on changes in simulation results
  - Full simulations are executed.
  - Simulations results are sampled and compared to a gold standard to *within a tolerance*.
- *Why a tolerance?* Accommodates small variations in software and hardware (e.g., Linux vs Mac, compilers)

```
if (abs(test_value – gold_standard) > tolerance) report_error()
```

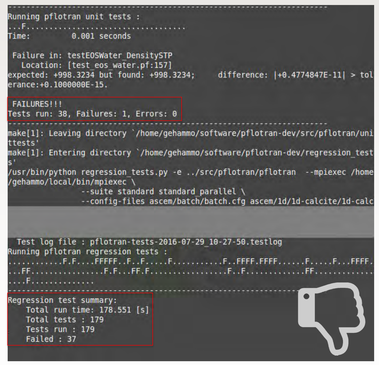


A successful run of the unit and regression tests.

Example regression test failure:

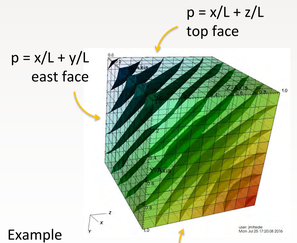
Perturb the critical pressure for the water equation of state by 10 billionths of a percent, and see what happens. . .

```
diff -r f9f01bbf557a src/pflotran/eos_water.F90
--- a/src/pflotran/eos_water.F90
+++ b/src/pflotran/eos_water.F90
@@ -893,6 +893,7 @@
     tc1 = H2O_CRITICAL_TEMPERATURE     ! K
     pc1 = H2O_CRITICAL_PRESSURE        ! Pa
+    pc1 = pc1 + 1.d-10*H2O_CRITICAL_PRESSURE
     vc1 = 0.00317d0  ! m^3/kg
     utc1 = one/tc1   ! 1/C
     upc1 = one/pc1   ! 1/Pa
```



An unsuccessful run of the unit and regression tests. Note several failing unit and regression tests! Somebody broke the code!

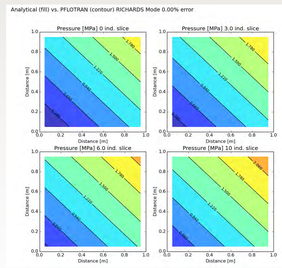
- **Verification tests**
  - Full simulations are executed, for which there is a known solution.
  - Simulation results are sampled and compared to an analytical solution *within a tolerance*.



Example Verification Test in Suite

Verification Problem Set-up:

- 3D domain (10x10x10 cells)
- Dirichlet pressure boundary conditions
- Solve for steady state pressure/flow fields



governing equation

$$\frac{\partial^2 p}{\partial x^2}+\frac{\partial^2 p}{\partial y^2}+\frac{\partial^2 p}{\partial z^2}=0$$

analytical solution

$$p(x,y,z)=p_0\left(\frac{x}{L}+\frac{y}{L}+\frac{z}{L}\right)$$

PFLOTRAN

Visit us at **http://www.pflotran.org**
Documentation at **http://www.documentation.pflotran.org**
Questions? Email us at **pflotran-users@googlegroups.com**

Bitbucket code repository **pflotran**

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-NA-0003525.

SAND2017-12231 C