Title: The Need for Reproducibility

Author(s): Robey, Robert W.

Intended for: Report

Issued: 2017-06-20 (rev.1)

The Need for Reproducibility

Bob Robey

June 18th, 2017

Background on Reproducibility

- **Reproducibility in Computational and Experimental Mathematics, ICERM Workshop[1,2]**
  - Dec 10-14, 2012, organizers Bailey, D. LBNL, Borwein, J., LeVeque, R. Univ. of Wash., Rider, B. SNL, Stein, W. Univ. of Wash., Stodden, V. Columbia Univ.
- **Supercomputing Conference Workshops**
  - Numerical Reproducibility at Exascale Workshop[3], Nov 20, 2015, organizers Keyrouz, W. and Mascagni, M., NIST
  - Obtaining Bitwise Reproducible Results – Perspectives and Latest Advances[4], Nov 22, 2013, organizer Clemons, N., Intel
- **University Working Groups**
  - Univ. of Wash. Reproducibility and Open Science, UC Berkeley, Reproducibility and Open Science, and NYU, Center for Data Science

1. **http://icerm.brown.edu/tw12-5-rcem/**

2. **Stodden, V., J. Borwein, and D. H. Bailey. "Setting the default to reproducible."** *Computational Science Research. SIAM News* **46 (2013): 4-6.**

3. **http://www.nist.gov/itl/ssd/is/numreprod2015.cfm**

4. **http://sc13.supercomputing.org/sites/default/files/WorkshopsArchive/track127.html**

Why Reproducibility?

- **Key tenet of Scientific Process**

- **Studies have shown difficulty in reproducing journal research**

  - Reproducibility Project, Center for Open Science, 2015 only 39 of 100 psychology studies

  - Other studies have had similar results, contributing to the credibility crisis for science

- **Necessary for predictive science**

- **Instill confidence in simulation results**

Purpose of this presentation

- **Start a discussion – that includes you!**
  - Is bitwise reproducible computation possible?
  - Can we improve our publication process of computational research in DOE?
  - What is peer review without reproducible results?
- **Publicize new advances in reproducibility**

Reproducibility can mean many things and cover many aspects of Computational Science

Components of Reproducibility

- **Software Process**
- **Numerical Reproducibility**
- **Data Provenance and Workflow**
- **Publication and Documentation**
- **Copyright and Licensing**

Discussion – are there other components of reproducibility?

Software Process for Reproducibility

- **Version Control**
  - What version of the code produced the results?
- **Scripts and automated workflows**
  - What were the parameters used for reported results?
  - How were plots generated?
- **Memory correctness tools for production code quality**
- **Portability and repeatability**
  - Can the results be obtained on a different system, compiler or number of processors?

Numerical Reproducibility

- **Growing issue at Exascale problem sizes**

- **Parallelism has exacerbated the problem**

- **New Exascale architectures with vectors and threads add additional difficulties**

- **Era of 64 bit numerical precision everywhere is ending**

Error in plot is shown relative to analytic answer

Global Sum Error for the Leblanc Problem
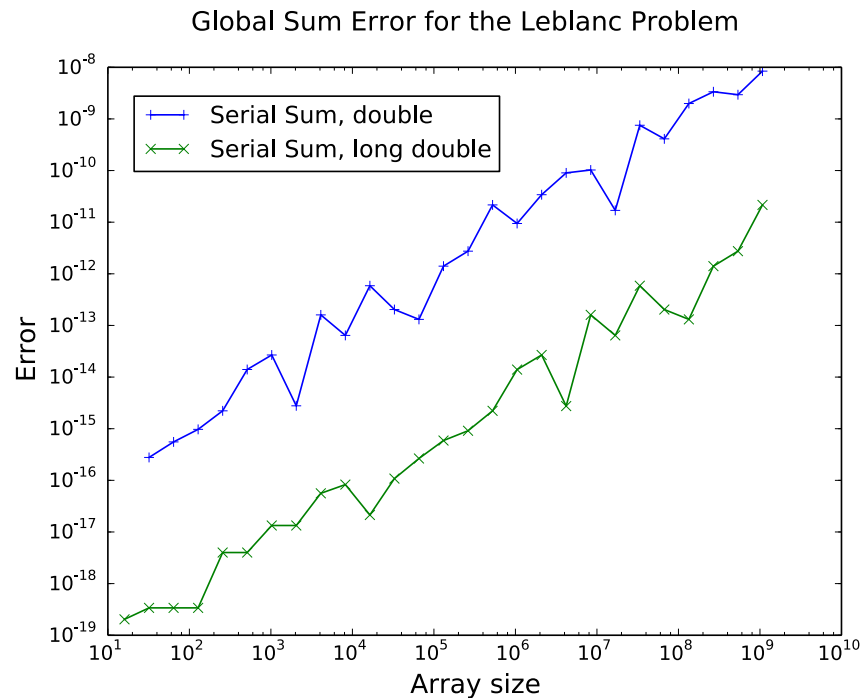
Bitwise Reproducibility

Full-application bit-for-bit reproducibility is nearly feasible at the present time. – David Bailey LBNL, SC15 Reproducibility workshop

- Majority of parallel bitwise reproducibility issue is with the global sums

- Tools are becoming available to assess where the issues are in the code

- Libraries available such as MKL, ReproBLAS, ExBLAS

The Insight

- **Reproducible global sums thought to require summation in a fixed order, but**

- **It can also be addressed by enhancing precision because regular addition is associative**

=> Can use both enhanced precision and order to reduce precision loss

Techniques to Improve the Reproducibility of the Global Sum

- **Pair-wise reduction**
- **Enhanced precision**
  - Kahan sum – accumulates error on one term
  - Knuth sum – accumulates error on both terms
  - Quadtype
- **Rounding/Truncation**
- **MPI enhanced precision sum**

How much can these improve reproducibility?

The Results

Error relative to analytic calculation
http://www.github.com/losalamos/GlobalSums

Method	Error	Run-time (msecs)
Double	-1.99e-09	0.116
Double w/truncation	0.0	0.120
Long Double	-1.31e-13	0.118
Long Double w/truncation	0.0	0.116
Kahan Sum	0.0	0.406
Knuth Sum	0.0	0.704
Pair-wise Sum	0.0	0.402
Quad Double	5.55e-17	3.010
Full Quad Double	-4.81e-27	2.454
OpenMP double	2.465e-10	0.048
OpenMP Kahan	1.39e-16	0.063

Array size ~ 134,000,000

Data Provenance and Workflow

- **How was data obtained? Who should get credit and what are the use license terms?**

- **Origin of all diagrams, pictures, plots, clip art and graphics known and credited?**

- **What modifications to the data have been performed?**

- **Where can someone access the data set?**

- **Statistical analysis calculations available?**

- **Plotting procedures available?**

Publication and Documentation

- Computational publications have not always been the best examples of reproducibility

- Often, almost a year goes by between submitting a journal paper and final revisions for publication. It really pays to automate and document the generation of the figures and computational results.

- Several journals now accept "supplemental" materials for published articles. This could be source code, data for plots, or generating the graphics and plots. (SIAM, ACM TOMS)

Supercomputing Conference 2016
Nov 13-16, 2016

Student Cluster Competition: SCC Reproducibility Initiative Winner

Replication and reproducibility of experimental computer science results in peer-reviewed paper is gaining relevance in the HPC community.

Technical Papers: Reproducibility Initiative:

As part of a major initiative that aims to increase the level of reproducibility and replicability of results, SC16 invites authors of technical papers to volunteer to publish their methodology, code and data with the paper, if their paper is accepted to SC16.

Copyright/Licensing

- **Reproducibility is easiest with publicly released code and data**

- **All products should be copyrighted and clearly licensed – shared when possible**

- **When public release is not possible, a protected repository or "lockbox" sharing system should be setup to allow appropriate reproducibility audits**

Thoughts for the HPC Community

- **Recent efforts in improving reproducibility have had an impact in a short time**
- **There is room for improvement in some areas in the HPC community**
- **Must consider upfront how to handle restricted information categories**
- **What new reproducibility techniques and procedures are most needed and could have the most impact?**
- **Who should take the lead in our community – scientists, labs, or agencies?**
- **Let us not forget cost/benefit considerations in new and on-going efforts**

Discussion

- **What aspects of reproducibility do you want to consider for your projects this summer?**
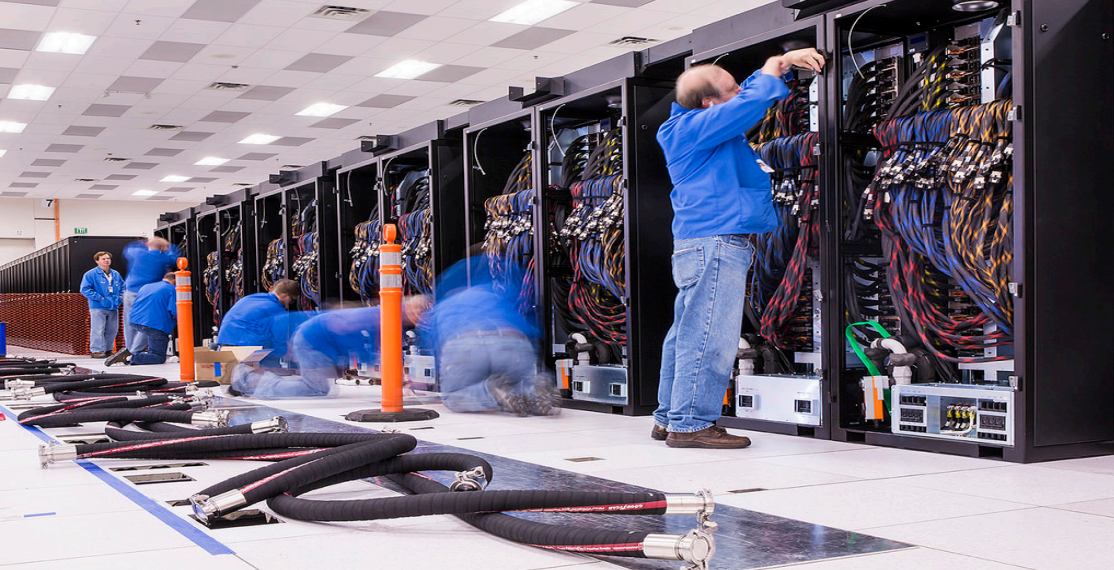
Los Alamos
NATIONAL LABORATORY
— EST.1943 —

Delivering science and technology
to protect our nation
and promote world stability