Quantifying Compiler Effects on Code Performance and Reproducibility using FLIT

Michael Bentley[1], Ian Briggs[1], Ganesh Gopalakrishnan[1], Dong H. Ahn[2], Ignacio Laguna[2], Gregory L. Lee[2], Holger Jones[2]
[1]University of Utah - Computer Science,
[2]Lawrence Livermore National Laboratory

INTRO

- HPC science requires **trust** and **reproducibility**, but also **performance**
- Compiler optimizations provide performance, but sometimes with **significant numerical differences**
- Our FLiT testing tool helps **quantify** these differences and **locate** the affected functions

METHODS

- **Differences** are measured using a **trusted** baseline compilation
- Bisect locates affected sites by **mixing files and functions** of a difference-producing compilation with the trusted baseline compilation

EXPERIMENTS

1. **MFEM:** measure performance and variability of this finite element code
2. **Laghos:** find where compiler optimizations invalidate the simulation
3. **LULESH:** show FLiT can find the function site of injected floating-point instructions

RESULTS

- One MFEM test showed **196% relative difference** from one contributing function
- FLiT Bisect found the difference-producing function in Laghos **in 40 minutes** versus 2 weeks manually.
- FLiT Bisect **found 100%** of the injections in LULESH, with **no false positives**.

DISCUSSION

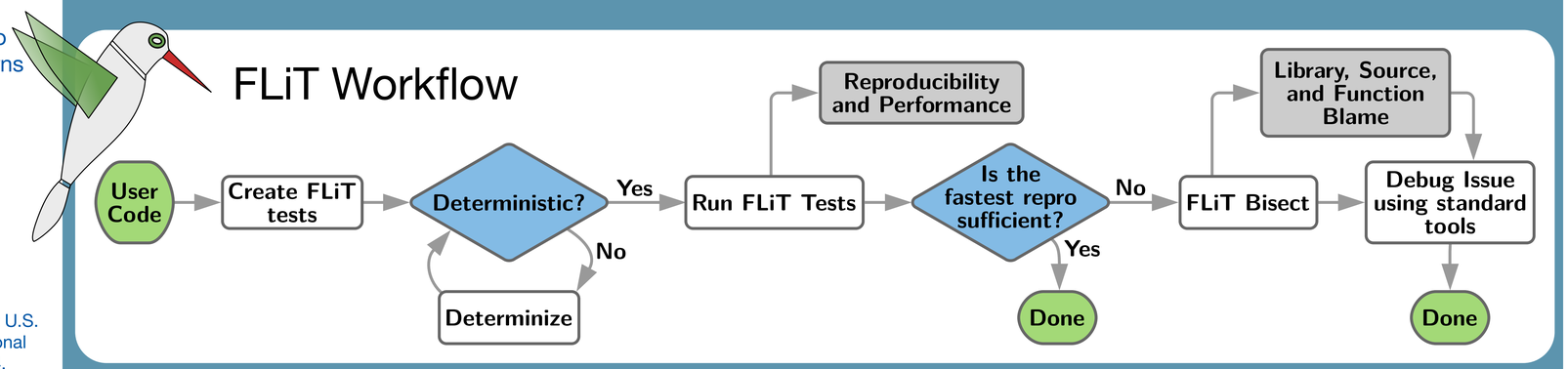
- FLiT provides testing and a workflow to resolve compiler reproducibility concerns
- Future work: find variability sites in the face of runtime nondeterminism

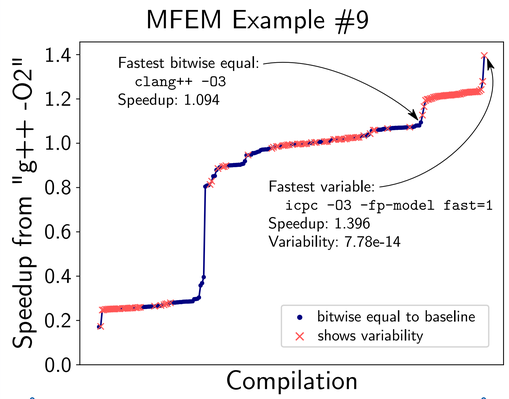
NSF OAC 1535032; Additional support for this work : U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

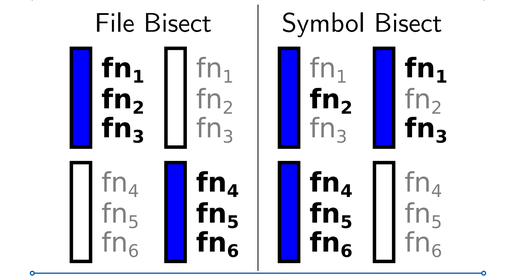
Can we find where compiler optimizations alter simulations?

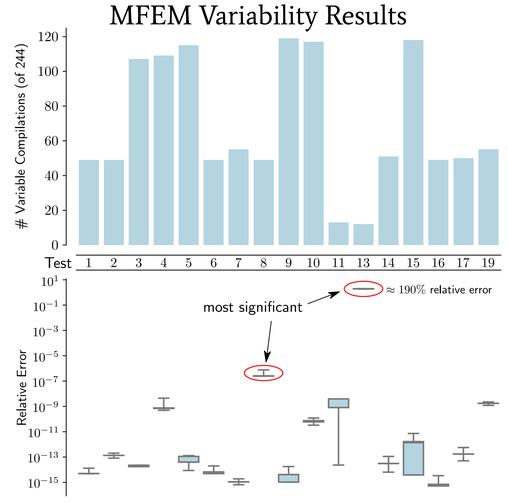
Yes.

The FLiT tool found a **196% relative error** in a finite element code and found the **one function site** where optimizations invalidated the simulation.









LULESH Injection Accuracy

Category	Count
exact finds	2,690
indirect finds	984
wrong finds	0
missed finds	0
not measurable	702
total	4,376

Laghos Bisect Results

baseline	digits k:	# files 1	# files 2	# files all	# funcs 1	# funcs 2	# funcs all	# runs 1	# runs 2	# runs all
g++ -O2	2	1	1	1	1	1	1	18	18	14
	3	1	1	1	1	1	1	18	18	14
	5	1	1	1	1	1	1	18	18	14
	all	2	3	5	1	2	7	28	37	57
xlc++ -O2	2	1	1	1	1	1	1	18	18	14
	3	1	1	1	1	1	1	18	18	14
	5	1	1	1	1	1	1	18	18	14
	all	2	3	6	1	3	7	28	37	69
xlc++ -O3 strict	2	1	1	1	1	1	1	18	18	14
	3	1	1	1	1	1	1	18	18	14
	5	1	1	1	1	1	1	18	18	14
	all	2	3	5	1	2	5	28	39	60