

Errata to “MinneSPEC: A New SPEC Benchmark Workload for Simulation-Based Computer Architecture Research”

AJ KleinOsowski
ajko@ece.umn.edu

David J. Lilja
lilja@ece.umn.edu

Department of Electrical and Computer Engineering
Minnesota Supercomputing Institute
University of Minnesota, Minneapolis, MN 55455

This paper¹ introduced and evaluated the MinneSPEC input sets for the SPEC CPU 2000 benchmark programs. This new workload has been developed specifically for use in simulation-based research studies.

It was recently discovered that, in the sim-profile simulator from the SimpleScalar version 3.0 suite, the counters used to track instruction mixes are explicitly declared to be unsigned ints, which makes them only 32 bits. As a result, when simulating programs with more than 4 billion (2^{32}) instructions, these counters overflow and produce incorrect outputs.

Since the benchmark programs in the SPEC CPU 2000 benchmark suite execute more than 4 billion (2^{32}) instructions with SPEC’s standard reference input sets, simulating these programs with sim-profile causes the counts of the different types of instructions executed to overflow. Consequently, the instruction mix profile results discussed in this paper and posted on the MinneSPEC project web site at <http://www.arctic.umn.edu/~lilja/minnespec/> prior to July, 2002, may contain errors.

The authors have modified sim-profile to allow it to correctly profile programs with more than 4 billion instructions and have collected corrected instruction mix profiles. In the interest of thoroughness, the authors also re-collected all of their gprof and sim-cache results, even though the gprof and sim-cache results were not affected by the sim-profile overflow errors. These updated profiles, and a corrected version of the Computer Architecture Letters paper (issued as University of Minnesota ARCTiC Labs technical report no. 02-08) are posted on the MinneSPEC project web site at <http://www.arctic.umn.edu/~lilja/minnespec/>.

The counter overflow problem caused paragraph three of the Results section and Table 1 of the June 2002 Computer Architecture Letters paper to be incorrect. Corrected versions of paragraph three and Table 1 are shown below.

Paragraph 3 of the Results section should read:

These comparisons show that the MinneSPEC function-level profiles are statistically close to the reference inputs for 18 out of the 32 benchmark configurations evaluated. The InstMix column, however, shows that the mix of dynamic instructions produced with the MinneSPEC inputs is statistically close to that produced with the reference inputs in almost all cases.

This error with the instruction mix profiles does not affect the use or distribution of the MinneSPEC workload. Only the instruction mix profiles, which are provided primarily as a convenience to MinneSPEC users, are affected. The other reported results, including the cache behavior gathered with sim-cache and the execution time profiles gathered with gprof, are unaffected by this counter overflow problem.

Furthermore, please note that the limited size of the counters in the standard distribution of sim-profile will affect all sim-profile users who run simulations longer than 4 billion instructions. If you fall into this category, the authors suggest you visit the MinneSPEC project web site to obtain the sim-profile patch which allows the correct profiling of programs with more than 4 billion instructions.

Many thanks to Hans Vandierendonck and Lieven Eeckhout of Ghent University for their detailed evaluation of the MinneSPEC workload which led to the discovery of these errors.

9 Oct 2002

¹The original version of this paper appeared in Computer Architecture Letters, June 2002.

TABLE I

BENCHMARKS INCLUDED IN THE CURRENT MINNESPEC INPUT SET DISTRIBUTION. THE *Input Reduction* COLUMN DOCUMENTS WHAT METHODS WERE USED TO DEVELOP THE MINNESPEC WORKLOADS. THE *FuncProf* AND *InstMix* COLUMNS DOCUMENT WHERE THE MINNESPEC FUNCTION-LEVEL EXECUTION PROFILE AND THE INSTRUCTION MIX ARE NEARLY IDENTICAL TO THE SPEC REFERENCE WORKLOAD, AS DETERMINED BY THE CHI-SQUARED TEST (SEE TEXT FOR DETAILS). FOR PROGRAMS WITH (NUM)[YN], THE MINNESPEC PROFILE IS COMPARED TO THE MULTIPLE SPEC REFERENCE WORKLOADS FOR THAT PROGRAM.

Benchmark	Type	Description	Input Reduction Method	FuncProf	InstMix
164.gzip.graphic	Int	Compression	modified command line, new input file	Y	Y
164.gzip.log	Int	Compression	modified command line, truncated ref file	N	Y
164.gzip.program	Int	Compression	modified command line, new input file	Y	Y
164.gzip.random	Int	Compression	modified command line, modified ref file	Y	Y
164.gzip.source	Int	Compression	modified command line, truncated ref file	N	Y
175.vpr.place	Int	FPGA Place Route	same as test	Y	Y
175.vpr.route	Int	FPGA Place Route	same as test	Y	Y
176.gcc	Int	C Compiler	same as train	5Y	5Y
177.mesa	Float	3-D Graphics	modified command line, modified test file	N	N
179.art	Float	Image Rec	modified command line	2N	2Y
181.mcf	Float	Combinatorial Opt	modified ref file	N	Y
183.quake	Float	Seismic Wave	modified train file	N	N
188.amp	Float	Chemistry	truncated ref file	N	N
197.parser	Int	Word Processing	sampled ref file	N	Y
253.perlbnk	Int	PERL Language	same as ref.makerand	1Y5N	2Y4N
255.vortex	Int	Database	modified train file	3Y	3Y
256.bzip2.graphic	Int	Compression	modified command line, modified ref file	Y	Y
256.bzip2.program	Int	Compression	modified command line, new input file	Y	Y
256.bzip2.source	Int	Compression	modified command line, sampled ref file	Y	Y
300.twolf	Int	Place and Route	sampled train file	Y	Y