OOKAMI PROJECT APPLICATION

Date: **June 15, 2021**

Project Title: Evaluation of OpenMP SIMD Transformation for ARM SVE

Usage:

☒ Testbed

☐ Production

Principal Investigator: Yonghong Yan, Associate Professor

University/Company/Institute: University of North Carolina at Charlotte

Mailing address including country: 410F Woodward Hall, Department of Computer Science, 9201 University City Blvd. Charlotte, NC 28223, USA

Phone: 704-687-8546

Email: yyan7@uncc.edu

Names & Email of initial project users:

Yonghong Yan: yyan7@uncc.edu
Patrick Flynn: pflynn5@uncc.edu
Xinyao Yi: xyi2@uncc.edu

Usage Description:

OpenMP SIMD directive provides a way for users to specify vectorization of data parallel loops so compiler can perform the directed transformation for SIMD architectures. The Ookami computing systems have A64fx processor that features ARM-based, multi-core, 512-bit SIMD-vector processor (SVE) with ultrahigh-bandwidth memory. It is designed for exhibit vector capability for future exascale computing.

We would like to experiment and evaluate our compiler for OpenMP SIMD code generation for ARM SVE architecture. We have implemented the support for code generation for ARM SVE and Intel AVX2/AVX-512 in our REX comiler for OpenMP SIMD. The evaluation will include performance studies of SVE in comparision with Intel AVX 512. We are happy to share and report our finding to the Ookami management and funding agency. Compiler is available from https://github.com/passlab/rexompiler.

Computational Resources:

Total node hours per year: ~500

Size (nodes) and duration (hours) for a typical batch job: within 10 minutes, mostly within a minute

Disk space (home, project, scratch): 40G for home, 100G for project to install compiler, test cases, 1TB for scratch for testing data.

Personnel Resources (assistance in porting/tuning, or training for your users**):**
We think the general document and support for accessing the resources such as login, batch submission should be sufficient.

Required software:
GNU compiler 7.0 and newer, GFortran, Clang/LLVM, Boost, JDK 8.0 or newer, Flex/Bison/etc standard development tools.

If your research is supported by US federal agencies:

Agency: NSF

Grant number(s): 1652732
https://www.nsf.gov/awardsearch/showAward?AWD_ID=1652732

Production projects:

Production projects should provide an additional 1-2 pages of documentation about how
(a) the code has been tuned to perform well on A64FX (ideally including benchmark data comparing performance with other architectures such as x86 or GPUs)
(b) it can make effective use of the key A64FX architectural features (notably SVE, the high-bandwidth memory, and NUMA characteristics)
(c) it can accomplish the scientific objectives within the available 32 Gbyte memory per node