Performance analysis of DOT Product on multicore architecture

in the context of portable code.

Context:

Computational reproducibility has been one of the core assumptions of underlying scientific computing. However, hardware developments over the past several decades have made it almost impossible to ensure computational reproducibility without incurring a severe loss of performance. This loss of reproducibility started when systems combined multigrain parallelism (SIMD, SIMT, Multicore, Cluster…) with non-determinism (out-of-order execution, execution concurrency,…). It has accelerated with recent architectural trends towards platforms with increasingly large numbers of processing elements, namely multicore CPUs and compute accelerators (GPUs, Intel Xeon Phi, FPGAs).

Programmers targeting these platforms rely on tools and libraries (ExBLAS, RARE-BLAS, MKL) to produce numerically reproducible codes or execute them efficiently. However, the incurred price is still high.

Objectives:

We have noticed that on simple operation such as the dot product of two vectors which is known to be bandwidth limited, today's reproducible implementation does not yet achieve memory-bound performance especially on GPU. The objectives of this master thesis are :

1) Reproduce and evaluate dot product performance of RARE-BLAS, Ex-BLAS, Ozaki Scheme under various execution environment (Multicore CPU, Nvidia/AMD GPU)
2) Propose and implement optimization according to the new computational capability of today's CPU/GPU for the numerically reproducible dot product (dot product with an infinitely precise accumulator)
3) Evaluate how proposed optimization achieve portable performance (use of High Level Language, Pragma, compiler's directives,…)

Contact :

- **Supervisor:**
  - David DEFOUR, (UPVD, FRANCE, david.defour@univ-perp.fr)
  - Note: This work will be done in collaboration with Daichi MUKUNOKI, (RIKEN, JAPAN, Daichi.mukunoki@riken.jp)
- **Place** : Laboratoire LAMPS, Université de Perpignan
- **Duration :** 6 months

Bibliographies :

- Chemseddine Chohra, Philippe Langlois, David Parello. Reproducible, Accurately Rounded and Efficient BLAS. *Euro-Par: Parallel Processing Workshops.*, Aug 2016, Grenoble, France. pp.609-620, ⟨10.1007/978-3-319-58943-5_49⟩. ⟨lirmm-01280324v2⟩
- Roman Iakymchuk, Stef Graillat, Caroline Collange, David Defour. ExBLAS: Reproducible and Accurate BLAS Library. *RAIM: Rencontres Arithmétiques de l'Informatique Mathématique*, Apr 2015, Rennes, France. 7ème Rencontre Arithmétique de l'Informatique Mathématique, 2015. ⟨hal-01140280⟩