ZPRIZE

ACCELERATING MSM OPERATIONS ON GPU/FPGA

Prize Sponsor

Aleo

Prize Architect

Aleo

Prize Description

Summary

Multiscalar multiplication (MSM) operations are an essential building block for zk computations. This prize will focus on computing the fastest MSM using either GPUs or FPGAs.

Optimization Objective

Compute the fixed-base point MSM for 2^26 randomly sampled scalars from the BLS 12-377 scalar field with lowest latency.

More specifically, given a fixed set of elliptic curve points from the BLS 12-377 G1 curve:

$$[P_1, P_2, \ldots, P_n]$$

and a randomly sampled input vector of finite field elements from the scalar field:

$$[k_1, k_2, \ldots, k_n]$$

Calculate elliptic curve point Q when n = 2^26

$$Q = \sum_{i=1}^{n} k_i * P_i$$

Constraints

- Sufficient documentation must be provided along with the implementation
- Only a single GPU/FPGA may be used; the problem may not be parallelized across multiple hardware instances.
- Submissions may be written in any language. The provided test harness, however, will be in Rust. So competitors submitting solutions using other languages will be required to create their own Rust bindings

Timeline

- **June 1 –** Final competitor selection
- **June 10 –** Competition begins
- **July 25 –** Mid-competition IPR
- **October 1 –** Deadline for submissions
- **October 15 –** Winners announced

Judging

Submissions will be checked for correctness and ranked by performance.   In addition, documentation (in English) must be provided along with the implementation. The documentation can be written in-line or as a separate document. It should be thorough and explanatory enough to provide an understanding of the techniques used in the submitted implementation without requiring an associated verbal explanation.

Correctness

The final correctness of the submission will be tested using randomly sampled test inputs/outputs that are not disclosed to the competitors during the competition. Submissions that fail any test cases will be judged as incorrect and lose the opportunity to win the prize.

Performance

Given input vectors consisting of 2^26 fixed elliptic curve points (bases), participants will compute a scalar multiplication of those bases with a set of four vectors of scalar elements from the associated BLS 12-377 G1 field in succession. Competitors will be provided with a set of test vectors to use while building the solution (of course, competitors can, and are encouraged to, use other vectors during the design and build process).

For scoring, solutions will be run using four randomly selected test vectors as input across ten trials in total. The winning submission will be the one with the lowest average latency across all ten trials.

In addition, all submissions will be manually reviewed by the prize committee appointed by the prize sponsor, Aleo.

Copy time for the bases to the device will not be counted (fixed base). Copy time will be counted for the 1st of the 4 sets of scalars will be counted, but not thereafter.

Hardware & Benchmarks

Competitors will be given access to one of the following:

- A dedicated instance of baseline image consisting of an AMD Epyc Milan CPU (8 cores), an A40 NVIDIA GPU and 48 GB of RAM provided by Coreweave.
- AWS credits to use an Amazon F1 instance FPGA f1.2xlarge consisting of a single FPGA and 8 CPU cores.

In addition, the prize sponsor (Aleo) will provide a test harness consisting of test vectors as well as a reference implementation that consists of:

- A starting GPU library for MSM (on NVidia cards) provided by Aleo
- A modified version of the ZCash FPGA MSM library for F1

Prize Allocation

All submissions that receive a prize must beat the benchmark solution. In the GPU case, the benchmark is given in the test harness provided to competitors. In the FPGA case, the benchmark to beat is 1M ops/second - 1M scalar multiplications and additions per second, or an MSM over a 1M points per second.

Each prize will be paid out according to how much a submission improves upon the baseline, according to the table below. Note that GPUs and FPGAs will be scored separately, and both the GPU category winner and FPGA category winner will be eligible for a total reward of up to $600,000. An additional $50,000 will be awarded to the fastest of the two (GPU vs. FPGA).

1-1.5x improvement	50%
1.5-2x improvement	+25%
2x or better improvement	+25%

Example: a competitor submits a solution using a GPU that beats the provided benchmark implementation by 5x. That competitor would earn 50% for beating the baseline from 1-1.5x, plus 25% (for beating the baseline by 1.5-2x) and another 25% for beating it by greater than 2x. So they would earn the total amount of the prize for the GPU: $600,000.

Prizes will be given out in good faith and in the sole discretion of the prize committee consisting of Jump Crypto, Aleo Systems Inc., and ZKValidator.

Notes

All submission code must be open-sourced at the time of submission. Code and documentation must be dual-licensed under both the MIT and Apache-2.0 licenses.

Questions

If there are any questions about this prize, please contact *Alex at Aleo*: zprize@aleo.org

References

[1] Scalar-multiplication algorithms. https://cryptojedi.org/peter/data/eccss-20130911b.pdf