

Subspace GPU Performance Study

Introduction	1
GPU Performance Foundations	1
Sloth Arithmetic Foundations	2
Core Finite Field Int32 Costs	2
Square Root Int32 Costs	2
4KB Sloth Encoding Costs	3
Sloth Encoding Theoretical Performance	3
Engineering Estimates for GPU Sloth Encoding	3
Summary	3

Introduction

The goal of this performance study is to understand the potential performance of GPU-based Sloth encoding for the Subspace network. Sloth encoding is a key function in the Subspace protocol and the cost and performance of this function may impact the security of the protocol. The fundamental operations in Sloth encoding are finite field operations, and this analysis will focus on these operations to determine the theoretical encoding performance of a GPU. After understanding the theoretical encoding performance, we will look at other factors to arrive at an engineering estimate of the practical performance. For the sake of this analysis we will focus on the Nvidia 2080Ti GPU as the target platform. Other GPU models may provide better economics through lower power consumption or lower capital costs per unit of compute capability.

GPU Performance Foundations

To understand the theoretical encoding performance on a GPU, we must first understand the computational capabilities of the target GPU. GPUs feature a large number of cores with a variety of processing units within each core. These cores contain processing units for floating point arithmetic, integer arithmetic, Tensor computations, and more. For the purposes of finite field operations we are most interested in the integer arithmetic capabilities. The Nvidia 2080 Ti is based on the [Turing architecture](#), and as a result each core contains one 32-bit multiplier. To determine the total integer arithmetic capabilities of the 2080Ti we can calculate the number of 32-bit integer (Int32) operations per second as seen below.

Nvidia 2080Ti Integer Performance

Frequency: 1.545 Ghz (1.545 billion cycles per second)
of Cores: 4352 (68 SMs with 64 CUDA Cores per SM)
Int32 Units/Core: 1

Int32 Operations/Cycle: 1

Int32 Operations/Second: **6,723,840,000,000**

Sloth Arithmetic Foundations

After understanding the theoretical integer arithmetic capabilities of the GPU we must then evaluate the theoretical costs of performing the core operations required to perform Sloth encoding. The core operation of Sloth encoding is taking a square root in a finite field. Each square root involves a number of finite field operations like modular multiplication and squaring.

Core Finite Field Int32 Costs

We will begin by understanding the number of int32 operations required to compute the two finite field primitives required for Sloth encoding, multiplication and squaring.

In the case of the Subspace construction we are working in a field size of 256 bits. This means that we must break each 256-bit integer into eight 32-bit limbs to match the underlying architecture of the GPU. The cost and complexity of each finite field operation can be found below, where n is the number of limbs:

of Int32 Operations per 256-bit (modulo-specific) finite field operation

Modular multiplication: **74 limb multiplication ops** $(n^2)+n+2$

Modular squaring: **46 limb multiplication ops** $(n*(n+1)/2)+n+2$

Square Root Int32 Costs

With an understanding of the costs of the fundamental finite field arithmetic, we will then look at the cost of completing a single Square Root iteration in terms of finite field operations.

of Finite Field Operations per Square Root

Multiply: **13**

Square: **253**

When combined with the number of Int32 operations calculated above, we can arrive at the total number of Int32 limb multiplication operations required to perform a Sloth encoding of a single 256-bit block.

Total # of Int32 limb multiplication operations per Square Root

Multiply: $13 * 74 = \mathbf{962}$

Square: $253 * 46 = \mathbf{11,638}$

Total: **12,600**

4KB Sloth Encoding Costs

Finally, after understanding the costs to perform a single round of sloth, we can calculate the total number of Int32 limb multiplication operations to encode a 4KB block.

Total # of Int32 Operations per 4KB Sloth Encoding

Total: 4096 Bytes / 256 bits * 12,600 = **1,612,800 Int32 limb multiplication operations**

Sloth Encoding Theoretical Performance

Based on the theoretical capabilities and cost estimates above, the total theoretical performance of a GPU-based Sloth encoder is as follows:

$(6,723,840,000,000 \text{ Int32 Ops/S}) \div (1,612,800 \text{ Int32 Ops/4KB}) = 16,676,190 \text{ KB/s} = \mathbf{16.68GB/s}$

Engineering Estimates for GPU Sloth Encoding

With an understanding of the theoretical performance of a GPU-based Sloth encoding we set out to determine better estimates on the potential actual performance. There are three key factors that can limit the potential performance of a GPU-based Sloth encoder:

1. **Additional Instructions:** The above calculations assume that the only instructions that will be executed will be Int32 limb multiplication instructions. In reality there are a variety of other instructions that will need to be executed to both process and move the data. In our experience for heavy cryptographic workloads, the number of total instructions is approximately two to three times the number of Int32 instructions, resulting in an approximate 60% reduction in performance or **~7 GB/s**.
2. **Bandwidth and Memory Bottlenecks:** Honest farmers will need to move data both to and from the GPU before and after encoding. Modern GPUs predominantly use PCIe v4 which supports up to 32GB/s in each direction. Given the available bandwidth and the reduced performance estimate above, it is unlikely that the workload will be bandwidth bound, however, careful orchestration of the data movement and memory access will be required to prevent further degradation in the Sloth encoding performance.
3. **Power Limitations:** Modern GPUs operate in a wide range of power envelopes, generally from ~150W to ~450W. Lower end GPUs may have insufficient power to achieve the estimated performance, resulting in degraded performance.

Summary

Based on the above analysis we expect that a 2080Ti GPU may be able to perform Sloth encoding up to ~7 GB/s. In comparison, a high end desktop like the 32-core AMD Threadripper will likely perform with approximately 10-20% of the projected throughput of the 2080Ti GPU.