

Assignment 2 Report

Julius Putra Tanu Setiaji (A0149787E), Chen Shaowei (A0110560Y)

18 November 2018

1 Program Design

The program is implemented in Cuda. The goal is to achieve the best possible speedup. Primary design considerations are:

- Ensure maximum utilisation of computation resources.
- Minimise overhead as much as possible.
- Ensure correctness of output by ensuring that only one process writes to result buffer.

2 Program Overview

1. *In host*: Read input.
2. *In host*: Setup:
 - Copy timestamp used to compute hash from host to device.
 - Copy target value from host to device.
 - Copy previous hash from host to device.
3. *In host*: Execute the kernel with the specified blocks-per-grid and threads-per-block.
4. *In kernel*: Compute initial nonce by calculating $blockId \times threadsPerBlock + threadId$.
5. *In kernel*: Construct SHA256 input and compute hash.
6. *In kernel*: Compare first 64 bits of digest with target value.
 - If the computed hash is satisfactory, execute an `atomicExch((int*)&is_found, 1)` to determine if current thread should write to output buffer.
 - Otherwise, increment the nonce by the value of $blocksPerGrid \times threadsPerBlock$.
7. *In kernel*: Check if `is_found` is set.
 - If set, return.
 - If unset, go to 5.
8. *In host*: Wait for kernel to complete by calling `cudaDeviceSynchronize()`.
9. *In host*: Print output.

3 Points to Note / Implementation Details

- Memory transfer between host and device memory is expensive and should be minimised. In our program, the only memory transfer between host and device are:
 - Copying of input and initial parameters from host to device in the setup stage.
 - Retrieve output value from device to host.
 - * **Note:** This is an implicit transfer. The buffer for storing output values is allocated as Unified Memory.
- Input values (timestamp, target, previous hash) are stored in constant memory since they are only read, not written to.
- A test and set operation `atomicExch` is used when a valid nonce is found to ensure that only one thread writes to output buffer.
- The completion flag must to be marked as `volatile` ensure that Unified Memory optimisations do not cause the variable to be migrated into shared/local memory.
- Due to the avalanche effect in SHA256, the program searches for valid nonces incrementally starting from 0 for simplicity.

4 Performance Analysis and Discussion

4.1 Platform Overview

Various tests were run on the SoC compute cluster(XGPC) and the Jetson TX2 module in the lab. Given below is an overview of the two platforms:

- **XGPC**
 - CPU: 2× Intel Xeon 4108 (Dual socket)
 - * 8 Cores
 - * 16 Threads
 - GPU: NVIDIA Tesla V100-PCIE
 - * 80 Volta SM
 - * 64 Cuda cores per SM
- **Jetson TX2**
 - CPU: NVIDIA Denver2 + ARM Cortex A-57
 - * 2 + 4 Cores (Heterogeneous)
 - GPU: NVIDIA
 - * 2 Pascal SM
 - * 128 Cuda cores per SM
 - *Note:* System memory is shared between CPU and GPU

5 Kernel Performance

In this section, we shall discuss the performance of the Cuda kernel. All values reported are 10-run averages.

Note: The following analysis does not consider the fixed overhead of copying input parameters from host to device and subsequent retrieval of result value. This is discussed in the next section.

5.1 Kernel Performance on Jetson TX2

The program was run on the Jetson TX2 with blocks-per-grid values between 1 and 64 and threads-per-block values between 1 and 256.

TIME GRAPH
SPEEDUP GRAPH

5.2 Kernel Performance on XGPC

The program was run on the Jetson TX2 with blocks-per-grid values between 1 and 80 and threads-per-block values between 1 and 256.

TIME GRAPH
SPEEDUP GRAPH

5.3 Discussion

We observe from the results above that in general, the program becomes faster when the number of total threads increases. This trend can be attributed to the following reasons:

- Each block can only be executed by a single SM. Therefore, the program performs better when the number of blocks is at least equal to the number of available SMs.
- The number of threads per block should be at least equal to the number of Cuda cores on each SM to ensure that all Cuda cores are utilised.
- Performance continues to improve even as the number of threads per block exceeds the number of Cuda cores available on each SM. This is because the GPU can hide memory latency when there are extra schedulable warps by executing an alternative warp when the currently active warp is waiting for a memory operation. The large number of registers on the GPU allows it to maintain the execution state of multiple threads at once and thus perform very low cost "context switches". This behaviour is most noticeable on the Jetson TX2 because $\sim 50\%$ of stalls are due to memory operations.

The experiments on the Jetson TX2 demonstrate the power of the GPU's ability to hide memory latency. When considering the test cases where there are at least 2 blocks and 256 total threads, we can see that there is very little correlation between read/write misses and execution time. It is more important to ensure that there is a large number of schedulable threads so that the GPU can hide memory latencies effectively.

Due to the extremely fast memory on the Tesla GPU, tests on the XGPC node were mostly bottlenecked by execution dependencies. Memory access was the cause for only 17% of stalls. As a result, the best speedup was achieved when the program spawned exactly one thread for each Cuda core. Additional schedulable threads actually caused a degradation in performance.

6 Performance Comparison with OpenMP implementation

In this section, we shall compare the performance of our Cuda application with an alternative OpenMP implementation.

The OpenMP implementation closely follows the design of the Cuda application. The primary difference is that the kernel code is wrapped in an `#pragma omp parallel` block. The storing of the result is protected by a `#pragma omp critical` block.

The following tests were executed on a compute cluster node (XGPC6). The OpenMP implementation was executed with 32 threads, matching the number of CPU threads on the machine. The Cuda implementation was executed with 80 blocks per grid and 64 threads per block, following the result of the previous section. Both implementations were run for targets between 2^{30} and 2^{48} .

6.1 Performance Comparison

6.2 Discussion

The above chart illustrates how costly data transfer between the host and device can be. Under the Cuda implementation, execution time is almost constant for targets between 2^{37} and 2^{48} . This is because the data transfer cost incurred is orders of magnitude greater than the computation cost. We know from the previous section that the kernel itself takes approximately $0.11ms$, while the total program required $1.41s$ to complete.

Since the OpenMP implementation did not have to incur this data transfer cost, it was significantly faster when the problem size was small (target was large). However, as the problem size becomes larger, the Cuda implementation took the lead. The data transfer cost was only incurred once regardless of program size, and no data transfers are necessary during the computation itself.