Lab 4 Report

Peiyu Wang Chenhui Zhou

a) The following is the performance results

```
[czg964@batman release]$ ./lab4


**===------------------------------------------------===**
Processing 16777216 elements...
Host CPU Processing time: 46.134998 (ms)
1
1
CUDA Processing time: 5.518000 (ms)
Speedup: 8.360819X
Test PASSED
[czg964@batman release]$
```

b) If the arrays are not a power of two in size, we add zero to the end to make the size become the power of two.

   We can avoid most bank conflicts by adding a variable amount of padding to each shared memory array index we compute.

c) FLOPS:
   a. CPU: The time complexity is O(n), so the FLOPS = 16777216 / (46/1000) = 364722087, the theoretical performance limits is 20G, for the bottlenecks, the cpu read and write operations takes lots of time.
   b. GPU: The time complexity is 2n, so the FLOPS = 2*16777216/(5.518/1000) = $6 * 10^9$. For the bottlenecks GPU always limited by the number of threads and blocks.