x	y	threads per block	blocks	speedup
17	1024	256	4	0.012 / 0.269 = 0.0446
17	4096	256	16	0.003 / 0.266 = 0.01128
2	1024	256	4	0.003 / 0.273 = 0.01099
2	4096	256	16	0.003 / 0.268 = 0.01119

1. As y increases from 1024 to 4096, the speedup decreases for x=17 (from 0.0446 to 0.0113) but remains constant for x=2 (~0.011). In all cases, speedup values are much less than 1.0, indicating the GPU is significantly slower than the CPU for these small problem sizes.
2. The GPU time remains relatively constant regardless of problem size because it is dominated by overhead. The CPU time varies based on the problem, making the speedup ratio dependent on CPU performance rather than GPU benefits.
3. The value of x significantly affects CPU performance but does not affect GPU performance, since GPU execution time is overwhelmingly dominated by the overhead rather than the actual computation time.