```
2 #include "cuda runtime.h"
3 #include "device_launch_parameters.h"
5 #include <stdio.h>
6 cudaError_t addWithCuda(int *c, const int *a, const int *b, unsigned int size);
8
   __global__ void addKernel(int *c, const int *a, const int *b)
   {
9
10
       int i = threadIdx.x;
11
       c[i] = a[i] + b[i] + 10000000;
12 }
13
14
15 int main()
16
   {
              不可变常量
17
18
        const int arraySize = 5;
        const int a[arraySize] = { 1, 2, 3, 4, 5 };
19
        const int b[arraySize] = { 10, 20, 30, 40, 50 };
20
21
        int c[arraySize] = { 0 };
22
23
       // Add vectors in parallel.
24
        cudaError_t cudaStatus = addWithCuda(c, a, b, arraySize);
25
        if (cudaStatus != cudaSuccess) {
            fprintf(stderr, "addWithCuda failed!");
26
27
            return 1;
28
29
30
        printf("\{1, 2, 3, 4, 5\} + \{10, 20, 30, 40, 50\} = \{\%d, \%d, \%d, \%d, \%d\} \setminus n",
            c[0], c[1], c[2], c[3], c[4]);
31
32
33
       // cudaDeviceReset must be called before exiting in order for profiling and
34
        // tracing tools such as Nsight and Visual Profiler to show complete traces.
        cudaStatus = cudaDeviceReset();
36
        if (cudaStatus != cudaSuccess) {
37
            fprintf(stderr, "cudaDeviceReset failed!");
38
            return 1;
39
40
41
       return 0;
42 }
43
44 // Helper function for using CUDA to add vectors in parallel.
   cudaError_t addWithCuda(int *c, const int *a, const int *b, unsigned int size)
46
                     定义设备端变量?
47
        int *dev a = 0;
48
        int *dev_b = 0;
        int *dev c = 0;
49
        cudaError t cudaStatus;
50
51
       // Choose which GPU to run on, change this on a multi-GPU system.
52
53
       cudaStatus = cudaSetDevice(0);
```

```
54
         if (cudaStatus != cudaSuccess) {
             fprintf(stderr, "cudaSetDevice failed! Do you have a CUDA-capable GPU
55
               installed?");
56
            goto Error;
57
58
        // Allocate GPU buffers for three vectors (two input, one output)
59
60
        cudaStatus = cudaMalloc((void**)&dev_c, size * sizeof(int));
         if (cudaStatus != cudaSuccess) {
61
             fprintf(stderr, "cudaMalloc failed!");
62
63
             goto Error;
64
        }
65
66
        cudaStatus = cudaMalloc((void**)&dev a, size * sizeof(int));
         if (cudaStatus != cudaSuccess) {
67
             fprintf(stderr, "cudaMalloc failed!");
68
69
             goto Error;
70
        }
71
72
        cudaStatus = cudaMalloc((void**)&dev b, size * sizeof(int));
73
         if (cudaStatus != cudaSuccess) {
74
             fprintf(stderr, "cudaMalloc failed!");
75
             goto Error;
76
        }
77
78
        // Copy input vectors from host memory to GPU buffers.
79
        cudaStatus = cudaMemcpy(dev_a, a, size * sizeof(int), cudaMemcpyHostToDevice);
80
        if (cudaStatus != cudaSuccess) {
            fprintf(stderr, "cudaMemcpy failed!");
81
82
             goto Error;
        }
83
84
85
        cudaStatus = cudaMemcpy(dev b, b, size * sizeof(int), cudaMemcpyHostToDevice);
        if (cudaStatus != cudaSuccess) {
86
             fprintf(stderr, "cudaMemcpy failed!");
87
88
             goto Error;
89
90
        // Launch a kernel on the GPU with one thread for each element.
91
92
        addKernel <<<1, size>>> (dev_c, dev_a, dev_b);
93
94
        // Check for any errors launching the kernel
95
        cudaStatus = cudaGetLastError();
96
         if (cudaStatus != cudaSuccess) {
             fprintf(stderr, "addKernel launch failed: %s\n", cudaGetErrorString
97
               (cudaStatus));
             goto Error;
98
99
        }
100
        // cudaDeviceSynchronize waits for the kernel to finish, and returns
101
102
        // any errors encountered during the launch.
        cudaStatus = cudaDeviceSynchronize();
103
104
        if (cudaStatus != cudaSuccess) {
```

```
fprintf(stderr, "cudaDeviceSynchronize returned error code %d after
               launching addKernel!\n", cudaStatus);
106
             goto Error;
107
108
109
         \ensuremath{//} Copy output vector from GPU buffer to host memory.
110
         cudaStatus = cudaMemcpy(c, dev_c, size * sizeof(int), cudaMemcpyDeviceToHost);
111
         if (cudaStatus != cudaSuccess) {
             fprintf(stderr, "cudaMemcpy failed!");
112
113
             goto Error;
114
         }
115
116 Error:
        cudaFree(dev_c);
117
118
         cudaFree(dev_a);
119
         cudaFree(dev_b);
120
121
        return cudaStatus;
122 }
123
```