```
1 // header files
 2 // standard headers
 3 #include <stdio.h>
 5 // cuda headers
 6 #include <cuda.h>
 7 #include "helper_timer.h"
 8
 9 // global variables
10 //const int iNumberOfArrayElements = 5;
11 const int iNumberOfArrayElements = 11444777;
12
13 float* hostInput1 = NULL;
14 float* hostInput2 = NULL;
15 float* hostOutput = NULL;
16 float* gold = NULL;
17
18 float* deviceInput1 = NULL;
19 float* deviceInput2 = NULL;
20 float* deviceOutput = NULL;
22 float timeOnCPU = 0.0f;
23 float timeOnGPU = 0.0f;
24
25 // CUDA kernel
26
   __global__ void vecAddGPU(float* in1, float* in2, float* out, int len)
27
28
29
       int i = blockIdx.x * blockDim.x + threadIdx.x;
30
31
       if (i < len)
32
        {
33
           out[i] = in1[i] + in2[i];
        }
34
35 }
36
37 // entry-point function
38 int main(void)
39 {
        // function declarations
40
41
       void fillFloatArrayWithRandomNumbers(float*, int);
        void vecAddCPU(const float*, const float*, float*, int);
42
43
       void cleanup(void);
44
       // variable declarations
45
46
        int size = iNumberOfArrayElements * sizeof(float);
47
        cudaError_t result = cudaSuccess;
48
        // code
49
50
        // host memory allocation
51
       hostInput1 = (float*)malloc(size);
        if (hostInput1 == NULL)
52
53
        {
           printf("Host Memory allocation is failed for hostInput1 array.\n");
54
55
           cleanup();
           exit(EXIT_FAILURE);
56
```

```
57
         }
 58
 59
         hostInput2 = (float*)malloc(size);
         if (hostInput2 == NULL)
 60
 61
         {
 62
             printf("Host Memory allocation is failed for hostInput2 array.\n");
 63
             cleanup();
             exit(EXIT_FAILURE);
 64
 65
         }
 66
         hostOutput = (float*)malloc(size);
 67
         if (hostOutput == NULL)
 68
 69
         {
             printf("Host Memory allocation is failed for hostOutput array.\n");
 70
             cleanup();
 71
 72
             exit(EXIT_FAILURE);
 73
         }
 74
 75
         gold = (float*)malloc(size);
         if (gold == NULL)
 76
 77
 78
             printf("Host Memory allocation is failed for gold array.\n");
 79
             cleanup();
             exit(EXIT_FAILURE);
 80
 81
         }
 82
         // filling values into host arrays
 83
         fillFloatArrayWithRandomNumbers(hostInput1, iNumberOfArrayElements);
 84
         fillFloatArrayWithRandomNumbers(hostInput2, iNumberOfArrayElements);
 85
 86
 87
         // device memory allocation
 88
         result = cudaMalloc((void**)&deviceInput1, size);
 89
         if (result != cudaSuccess)
 90
             printf("Device Memory allocation is failed for deviceInput1 array.
 91
               \n");
 92
             cleanup();
 93
             exit(EXIT_FAILURE);
 94
         }
 95
         result = cudaMalloc((void**)&deviceInput2, size);
 96
         if (result != cudaSuccess)
 97
 98
         {
 99
             printf("Device Memory allocation is failed for deviceInput2 array.
               \n");
100
             cleanup();
             exit(EXIT_FAILURE);
101
102
         }
103
104
         result = cudaMalloc((void**)&deviceOutput, size);
         if (result != cudaSuccess)
105
106
             printf("Device Memory allocation is failed for deviceOutput array.
107
               \n");
             cleanup();
108
109
             exit(EXIT_FAILURE);
```

```
\HPP Seminar 2022\cuda\VecAdd.cu
```

```
3
110
         }
111
112
         // copy data from host arrays into device arrays
113
         result = cudaMemcpy(deviceInput1, hostInput1, size,
                                                                                      7
           cudaMemcpyHostToDevice);
114
         if (result != cudaSuccess)
115
             printf("Host to Device Data Copy is failed for deviceInput1 array.
116
117
             cleanup();
118
             exit(EXIT_FAILURE);
         }
119
120
         result = cudaMemcpy(deviceInput2, hostInput2, size,
121
                                                                                      P
           cudaMemcpyHostToDevice);
122
         if (result != cudaSuccess)
123
         {
             printf("Host to Device Data Copy is failed for deviceInput2 array.
124
               \n");
125
             cleanup();
126
             exit(EXIT_FAILURE);
         }
127
128
129
         // CUDA kernel configuration
         dim3 dimGrid = dim3((int)ceil((float)iNumberOfArrayElements / 256.0f), 1, >
130
         dim3 \ dimBlock = dim3(256, 1, 1);
131
132
133
         // CUDA kernel for Vector Addition
134
         StopWatchInterface* timer = NULL;
135
         sdkCreateTimer(&timer);
         sdkStartTimer(&timer);
136
137
         vecAddGPU << <dimGrid, dimBlock >> > (deviceInput1, deviceInput2,
138
           deviceOutput, iNumberOfArrayElements);
139
         sdkStopTimer(&timer);
140
         timeOnGPU = sdkGetTimerValue(&timer);
141
         sdkDeleteTimer(&timer);
142
143
         timer = NULL;
144
145
         // copy data from device array into host array
         result = cudaMemcpy(hostOutput, deviceOutput, size,
146
                                                                                     P
           cudaMemcpyDeviceToHost);
147
         if (result != cudaSuccess)
148
         {
             printf("Device to Host Data Copy is failed for hostOutput array.\n");
149
150
             cleanup();
151
             exit(EXIT_FAILURE);
152
         }
153
154
         // vector addition on host
155
         vecAddCPU(hostInput1, hostInput2, gold, iNumberOfArrayElements);
156
157
         // comparison
158
         const float epsilon = 0.000001f;
```

```
159
         int breakValue = -1;
160
         bool bAccuracy = true;
         for (int i = 0; i < iNumberOfArrayElements; i++)</pre>
161
162
163
             float val1 = gold[i];
164
             float val2 = hostOutput[i];
165
             if (fabs(val1 - val2) > epsilon)
166
                 bAccuracy = false;
167
                 breakValue = i;
168
169
                 break;
170
             }
171
         }
172
         char str[128];
173
174
         if (bAccuracy == false)
175
             sprintf(str, "Comparison of CPU and GPU Vector Addition is not within ➤
               accuracy of 0.000001 at array index %d", breakValue);
176
         else
             sprintf(str, "Comparison of CPU and GPU Vector Addition is within
177
               accuracy of 0.000001");
178
179
         // output
180
         printf("Array1 begins from 0th index %.6f to %dth index %.6f\n",
           hostInput1[0], iNumberOfArrayElements - 1, hostInput1
           [iNumberOfArrayElements - 1]);
         printf("Array2 begins from 0th index %.6f to %dth index %.6f\n",
181
           hostInput2[0], iNumberOfArrayElements - 1, hostInput2
           [iNumberOfArrayElements - 1]);
182
         printf("CUDA Kernel Grid dimension = %d,%d,%d and Block dimension = %d,%d, >
           %d\n", dimGrid.x, dimGrid.y, dimGrid.z, dimBlock.x, dimBlock.y,
           dimBlock.z);
183
         printf("Output Array begins from 0th index %.6f to %dth index %.6f\n",
                                                                                      0
           hostOutput[0], iNumberOfArrayElements - 1, hostOutput
                                                                                      P
           [iNumberOfArrayElements - 1]);
184
         printf("Time taken for Vector Addition on CPU = %.6f\n", timeOnCPU);
         printf("Time taken for Vector Addition on GPU = %.6f\n", timeOnGPU);
185
         printf("%s\n", str);
186
187
188
         // cleanup
189
         cleanup();
190
         return(0);
191
192 }
193
194 void fillFloatArrayWithRandomNumbers(float* arr, int len)
195 {
196
         // code
         const float fscale = 1.0f / (float)RAND_MAX;
197
198
         for (int i = 0; i < len; i++)
199
         {
200
             arr[i] = fscale * rand();
201
         }
202 }
203
204 void vecAddCPU(const float* arr1, const float* arr2, float *out, int len)
```

```
205
    {
206
         // code
207
         StopWatchInterface* timer = NULL;
         sdkCreateTimer(&timer);
208
209
         sdkStartTimer(&timer);
210
211
         for (int i = 0; i < len; i++)
212
         {
             out[i] = arr1[i] + arr2[i];
213
214
         }
215
216
         sdkStopTimer(&timer);
         timeOnCPU = sdkGetTimerValue(&timer);
217
218
         sdkDeleteTimer(&timer);
219
         timer = NULL;
220 }
221
222 void cleanup(void)
223
         // code
224
225
         if (deviceOutput)
226
             cudaFree(deviceOutput);
227
             deviceOutput = NULL;
228
229
         }
230
231
         if (deviceInput2)
232
         {
233
             cudaFree(deviceInput2);
             deviceInput2 = NULL;
234
235
         }
236
         if (deviceInput1)
237
238
             cudaFree(deviceInput1);
239
240
             deviceInput1 = NULL;
241
         }
242
         if (gold)
243
244
         {
             free(gold);
245
246
             gold = NULL;
247
         }
248
         if (hostOutput)
249
250
         {
             free(hostOutput);
251
252
             hostOutput = NULL;
253
         }
254
255
         if (hostInput2)
256
257
             free(hostInput2);
258
             hostInput2 = NULL;
259
         }
260
```

```
\HPP Seminar 2022\cuda\VecAdd.cu
```

```
6
```

```
261    if (hostInput1)
262    {
263         free(hostInput1);
264         hostInput1 = NULL;
265    }
266 }
```