Εργαστήριο Παράλληλων Συστημάτων - Εργασία 2

Χρήστος Μαργιώλης Ιανουάριος 2023

Περιεχόμενα

| 1 | Προ | ογράμματα | 2 |
|----------|-------------|----------------------------|----|
| | $1.\dot{1}$ | 'Ασκηση 2A | 2 |
| | | 1.1.1 Κώδιχας | 2 |
| | | 1.1.2 Ενδεικτικά τρεξίματα | 6 |
| | 1.2 | 'Ασκηση 2B-A | 7 |
| | | 1.2.1 Κώδιχας | 7 |
| | | 1.2.2 Ενδεικτικά τρεξίματα | 10 |
| | 1.3 | 'Ασκηση 2Β-Β | 13 |
| | | 1.3.1 Κώδιχας | 13 |
| | | 1.3.2 Ενδεικτικά τρεξίματα | 16 |
| 2 | Про | οβλήματα | 18 |

1 Προγράμματα

Οι κώδικες έχουν σχόλια μόνο στα σημεία που θεώρησα ότι μπορεί να προκύψει κάποιο «μπέρδεμα».

1.1 'Ασκηση 2Α

1.1.1 Κώδικας

```
#include <err.h>
#include <stdio.h>
#include <stdlib.h>
#include <time.h>
#include <omp.h>
static void
                pretty_print(int *, int, const char *);
static int
                cmpfunc(const void *, const void *);
                merge(int *, int *, int *, int *, int *);
static void
                multisort(int *, int *, int);
static void
 * Print the contents of an array like:
 * array = [x, y, z]
 */
static void
pretty_print(int *arr, int n, const char *name)
{
        int i;
        printf("\n%s = [", name);
        for (i = 0; i < n; i++)
                printf("%d%s", arr[i], (i == n - 1) ? "" : ", ");
        printf("]\n");
}
 * Passed to qsort(3).
 */
static int
cmpfunc(const void *a, const void *b)
{
        return (*(int *)a - *(int *)b);
}
/*
```

```
* Merge sort
 */
static void
merge(int *a, int *enda, int *b, int *endb, int *res)
        while (a <= enda && b <= endb) \{
                 if (*a < *b)
                         *res++ = *a++;
                 else
                         *res++ = *b++;
        }
        while (a <= enda)</pre>
                *res++ = *a++;
        while (b <= endb)</pre>
                *res++ = *b++;
}
static void
multisort(int *arr, int *space, int n)
        int quarter, *sta, *spa, *stb, *spb, *stc, *spc, *std, *spd;
         * Sort with qsort(3) directly if we can't split the array into 4
         * quarters.
        if ((quarter = n / 4) < 4)
                 qsort(arr, n, sizeof(int), cmpfunc);
        else {
                 /* Split the array into 4 quarters. */
                sta = arr;
                spa = space;
                 stb = sta + quarter;
                spb = spa + quarter;
                stc = stb + quarter;
                spc = spb + quarter;
                std = stc + quarter;
                spd = spc + quarter;
                 /* Sort each quarter */
#pragma omp task
                multisort(sta, spa, quarter);
#pragma omp task
                multisort(stb, spb, quarter);
#pragma omp task
                multisort(stc, spc, quarter);
#pragma omp task
```

```
multisort(std, spd, n - 3 * quarter);
                /* Wait for the tasks above to finish. */
#pragma omp taskwait
#pragma omp task
                /* Merge A and B into SpaceA */
                merge(sta, sta + quarter - 1, stb, stb + quarter - 1, spa);
#pragma omp task
                /* Merge C and D into SpaceC */
                merge(stc, stc + quarter - 1, std, arr + n - 1, spc);
#pragma omp taskwait
                /* Merge the two resulting couples (SpaceA and SpaceC). */
                merge(spa, spc - 1, spc, space + n - 1, arr);
        }
}
int.
main(int argc, char *argv[])
        int *a, *space, i, n, ntd;
        double start, end;
        if (argc < 3) {
                fprintf(stderr, "usage: %s nthreads n\n", *argv);
                return (1);
        if ((ntd = atoi(argv[1])) < 1)</pre>
                err(1, "can't use nthreads n < 1");</pre>
        if ((n = atoi(argv[2])) < 1)
                err(1, "can't use n < 1");
        srand(time(NULL));
        omp_set_num_threads(ntd);
        if ((a = malloc(n * sizeof(int))) == NULL)
                err(1, "malloc");
        if ((space = malloc(n * sizeof(int))) == NULL)
                err(1, "malloc");
        for (i = 0; i < n; i++)
                a[i] = rand() % 100;
        /* Calculate speed up */
        start = omp_get_wtime();
        pretty_print(a, n, "A_unsorted");
        multisort(a, space, n);
```

```
pretty_print(a, n, "A_multisort");
end = omp_get_wtime();
printf("Total time: %f seconds\n", end - start);
free(a);
return (0);
}
```

1.1.2 Ενδεικτικά τρεξίματα

usage: ./a.out nthreads n

 Γ ia nthreads = 2 xai n = 10:

christos@tpad\$./ex2a 2 10

A_unsorted = [82, 43, 18, 24, 3, 17, 89, 13, 63, 44]

A_multisort = [3, 13, 17, 18, 24, 43, 44, 63, 82, 89] Total time: 0.000030 seconds

 Γ ia nthreads = 8 xai n = 100:

christos@tpad\$./ex2a 8 100

A_unsorted = [39, 13, 48, 42, 60, 32, 38, 0, 13, 11, 82, 68, 25, 18, 52, 8, 45, 19, 17, 8, 47, 56, 77, 74, 58, 38, 34, 72, 63, 95, 23, 3, 9, 71, 97, 21, 55, 87, 73, 20, 51, 7, 88, 76, 25, 41, 36, 71, 60, 53, 31, 59, 61, 60, 86, 71, 98, 20, 95, 14, 67, 18, 69, 28, 42, 66, 49, 97, 54, 74, 18, 57, 81, 58, 33, 59, 99, 21, 8 2, 12, 26, 13, 23, 87, 74, 9, 10, 24, 81, 58, 38, 1, 76, 7, 29, 70, 26, 31, 68, 80]

A_multisort = [0, 1, 3, 7, 7, 8, 8, 9, 9, 10, 11, 12, 13, 13, 13, 14, 17, 18, 18, 18, 19, 20, 20, 21, 21, 2 3, 23, 24, 25, 25, 26, 26, 28, 29, 31, 31, 32, 33, 34, 36, 38, 38, 38, 39, 41, 42, 42, 45, 47, 48, 49, 51, 52, 53, 54, 55, 56, 57, 58, 58, 58, 59, 59, 60, 60, 60, 61, 63, 66, 67, 68, 68, 69, 70, 71, 71, 71, 72, 73, 74, 74, 74, 76, 76, 77, 80, 81, 81, 82, 82, 86, 87, 87, 88, 95, 95, 97, 97, 98, 99] Total time: 0.000057 seconds

Για nthreads = 16 και n = 1000000. Λόγω του αριθμού των στοιχείων, το στιγμιότυπο δείχνει μόνο τον χρόνο υπολογισμού:

Total time: 0.519716 seconds

1.2 'Ασκηση 2Β-Α

1.2.1 Κώδικας

```
#include <stdio.h>
#include <time.h>
#define N
             (1 << 2)
#define DIM
                (N * N)
 * This formula for calculating the number of blocks is mentioned at "out of
 * the blocks" section in:
 * https://developer.nvidia.com/blog/even-easier-introduction-cuda/
#define BLKSIZE (1 << 8)</pre>
#define NBLK
             ((DIM + BLKSIZE - 1) / BLKSIZE)
__global__ void
convolution(float *a, float *aconv)
{
       float c11, c12, c13, c21, c22, c23, c31, c32, c33;
       int i, j, x, stridex;
        /*
         * Each thread gets a slice of the rows to work with. Grid-stride idiom
         * mentioned at section "out of the blocks" in:
         * https://developer.nvidia.com/blog/even-easier-introduction-cuda/
       x = blockIdx.x * blockDim.x + threadIdx.x;
       stridex = blockDim.x * gridDim.x;
       /* Random weight values */
       c11 = +0.2; c21 = +0.5; c31 = -0.8;
       c12 = -0.3; c22 = +0.6; c32 = -0.9;
       c13 = +0.4; c23 = +0.7; c33 = +0.10;
       if (x < 1 | | x > N - 1)
               return;
       for (i = x; i < N - 1; i += stridex) {
                for (j = 1; j < N - 1; j++) {
                        /* Taken from the lab's example code. */
                        aconv[i * N + j] =
                            c11 * a[(i - 1)]
                                                * N + (j - 1)] +
                            c12 * a[i
                                                * N + (j - 1)] +
                            c13 * a[(i + 1)]
                                            * N + (j - 1)] +
```

```
c21 * a[(i - 1)]
                                                 * N + j] +
                             c22 * a[i
                                                 * N + j] +
                             c23 * a[(i + 1)]
                                                 * N + j] +
                             c31 * a[(i - 1)]
                                                 * N + (j + 1)] +
                                                 * N + (j + 1)] +
                             c32 * a[i
                             c33 * a[(i + 1)]
                                                 * N + (j + 1)];
                }
        }
}
__global__ void
min_diagonal(float *arr, float *min_arr)
        int x, stridex, i;
        x = blockIdx.x * blockDim.x + threadIdx.x;
        stridex = blockDim.x * gridDim.x;
        if (x >= N)
                return;
        /* Calculate local minimums */
        min_arr[x] = arr[x * N + x];
        for (i = x; i < N; i += stridex)
                if (arr[i * N + i] < min_arr[x])</pre>
                        min_arr[x] = arr[i * N + i];
}
static void
pretty_print(float *arr, const char *name)
{
        int i, j;
        printf("\n\%s = [\n", name);
        for (i = 0; i < N; i++) {
                printf("\t[");
                for (j = 0; j < N; j++) {
                         printf("%.2f%s", arr[i * N + j],
                            (j == N - 1) ? "]\n" : ", ");
                }
        printf("]\n");
}
main(int argc, char *argv[])
```

```
float *a, *aconv, *min_arr, min;
int i;
srand(time(NULL));
 * Use unified memory to avoid having additional device arrays and
 * memcpying from host to device and vice versa.
 * https://developer.nvidia.com/blog/unified-memory-cuda-beginners/
cudaMallocManaged(&a, DIM * sizeof(float));
cudaMallocManaged(&aconv, DIM * sizeof(float));
cudaMallocManaged(&min_arr, DIM * sizeof(float));
/* Initialize array */
for (i = 0; i < DIM; i++)
        a[i] = (float)(rand() % 100);
convolution<<<NBLK, BLKSIZE>>>(a, aconv);
/* Wait for all devices to finish */
cudaDeviceSynchronize();
min_diagonal<<<NBLK, BLKSIZE>>>(aconv, min_arr);
cudaDeviceSynchronize();
/*
 * Find global minimum using the local minimums calculated in
 * min_diagonal().
 */
min = min_arr[0];
for (i = 0; i < N; i++)
        if (min_arr[i] < min)</pre>
                min = min_arr[i];
pretty_print(a, "A");
pretty_print(aconv, "A_conv");
printf("Min_diagonal(A_conv): %.2f\n", min);
cudaFree(a);
cudaFree(aconv);
cudaFree(min_arr);
return (0);
```

}

1.2.2 Ενδεικτικά τρεξίματα

 Γ ia NxN = 8x8 xai blocksize = 256

```
cuda17@rncp-ubuntu:"$ nvcc ex2b_a.cu -o ex2b_a
cuda17@rncp-ubuntu:"$ ./ex2b_a
A = [
        [12.00, 20.00, 37.00, 60.00, 14.00, 43.00, 62.00, 48.00]
        [27.00, 90.00, 88.00, 73.00, 71.00, 4.00, 44.00, 84.00]
        [87.00, 89.00, 68.00, 0.00, 18.00, 89.00, 21.00, 26.00]
        [32.00, 16.00, 32.00, 10.00, 46.00, 7.00, 7.00, 58.00]
        [79.00, 44.00, 70.00, 93.00, 40.00, 84.00, 94.00, 67.00]
        [74.00, 82.00, 92.00, 98.00, 39.00, 37.00, 34.00, 78.00]
        [26.00, 2.00, 30.00, 97.00, 91.00, 51.00, 23.00, 76.00]
        [67.00, 8.00, 86.00, 13.00, 15.00, 93.00, 23.00, 94.00]
A conv = [
        [0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00]
        [0.00, 53.40, 17.80, 8.70, 23.20, -12.20, 3.70, 0.00]
        [0.00, -26.70, 47.50, -14.90, 14.50, 34.10, -68.40, 0.00]
        [0.00, 48.10, 133.10, 51.30, 29.70, 99.60, 63.50, 0.00]
        [0.00, 24.70, 63.30, 82.70, 10.10, 5.80, -24.20, 0.00]
        [0.00, -59.20, -56.70, 113.60, 39.70, 29.10, -6.60, 0.00]
        [0.00, -10.40, -21.20, 48.50, 14.10, 55.10, -45.20, 0.00]
        [0.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00]
Min_diagonal(A_conv): -45.20
```

 Γ ia NxN = 32x32 xai blocksize = 1024

 Γ ia NxN = 4x4 xai blocksize = 256

```
cuda17@rncp-ubuntu:~$ ./ex2b_a
A = [
        [83.00, 6.00, 87.00, 47.00]
        [7.00, 30.00, 38.00, 78.00]
        [27.00, 31.00, 9.00, 12.00]
        [90.00, 89.00, 53.00, 13.00]
A conv = [
        [0.00, 0.00, 0.00, 0.00]
        [0.00, -34.90, -29.40, 0.00]
        [0.00, 92.00, 21.90, 0.00]
        [0.00, 0.00, 0.00, 0.00]
Min_diagonal(A_conv): -34.90
```

1.3 'Ασκηση 2B-B

1.3.1 Κώδικας

```
#include <stdio.h>
#include <time.h>
#define N
                (1 << 3)
#define M
                (1 << 3)
#define DIM
                (N * M)
#define BLKSIZE (1 << 10)</pre>
#define NBLK
              ((DIM + BLKSIZE - 1) / BLKSIZE)
 * Calculations taken from lab's example code.
 */
__global__ void
transnorm(float *a, float *atrans, float *x, float *y)
{
        int i, j, idx, stridex;
        /* Each thread gets a slice of the rows to work with */
        idx = blockIdx.x * blockDim.x + threadIdx.x;
        stridex = blockDim.x * gridDim.x;
        if (idx >= N)
                return;
        /* First thread initializes y */
        if (threadIdx.x == 0) {
                for (i = 0; i < M; i++)
                        y[i] = 0;
        for (i = idx; i < N; i += stridex) {
                for (j = 0; j < M; j++) {
                        /* Transpose A */
                        atrans[j * N + i] = a[i * M + j];
                        y[j] = atrans[j * M + i] * a[i * M + j] * x[j];
                }
        }
}
static void
pretty_print_1d(float *arr, const char *name, int n)
{
        int i;
        printf("\n%s = [", name);
```

```
for (i = 0; i < n; i++) {
                printf("%.2f%s", arr[i],
                   (i == n - 1) ? "" : ", ");
        printf("]\n");
}
static void
pretty_print_2d(float *arr, const char *name, int w, int h)
{
        int i, j;
        printf("\n%s = [\n", name);
        for (i = 0; i < w; i++) {
                printf("\t[");
                for (j = 0; j < h; j++) {
                        printf("%.2f%s", arr[i * h + j],
                           (j == h - 1) ? "] \n" : ", ");
        }
        printf("]\n");
}
int
main(int argc, char *argv[])
        float *a, *atrans, *x, *y;
        int i, j;
        srand(time(NULL));
        /*
         * Use unified memory to avoid having additional device arrays and
         * memcpying from host to device and vice versa.
         */
        cudaMallocManaged(&a, DIM * sizeof(float));
        cudaMallocManaged(&atrans, DIM * sizeof(float));
        cudaMallocManaged(&x, M * sizeof(float));
        cudaMallocManaged(&y, M * sizeof(float));
        /* Initialize arrays */
        for (i = 0; i < N; i++) {
                x[i] = (float)(rand() % 100);
                for (j = 0; j < M; j++)
                        a[i * M + j] = (float)(rand() % 100);
        }
```

```
transnorm<<<NBLK, BLKSIZE>>>(a, atrans, x, y);
/* Wait for all devices to finish */
cudaDeviceSynchronize();

pretty_print_2d(a, "A", N, M);
pretty_print_2d(atrans, "A_trans", M, N);
pretty_print_1d(x, "X", M);
pretty_print_1d(y, "Y", M);

cudaFree(a);
cudaFree(atrans);
cudaFree(x);
cudaFree(y);

return (0);
```

}

1.3.2 Ενδεικτικά τρεξίματα

 Γ ia NxM = 4x2 xai blocksize = 256

```
cuda17@rncp-ubuntu:"$ ./ex2b_b
        [7.00, 59.00]
        [73.00, 40.00]
        [46.00, 69.00]
        [46.00, 10.00]
A_trans = [
        [7.00, 73.00, 46.00, 46.00]
        [59.00, 40.00, 69.00, 10.00]
X = [11.00, 74.00]
    [539.00, 200836.00]
```

```
cuda17@rncp-ubuntu:"$ ./ex2b b
A = [
        [67.00, 59.00, 97.00, 12.00, 98.00, 5.00, 56.00, 45.00]
        [50.00, 17.00, 87.00, 98.00, 94.00, 28.00, 54.00, 20.00]
        [73.00, 38.00, 3.00, 67.00, 71.00, 38.00, 25.00, 57.00]
        [66.00, 87.00, 85.00, 74.00, 7.00, 97.00, 23.00, 19.00]
A trans = [
        [67.00, 50.00, 73.00, 66.00]
        [59.00, 17.00, 38.00, 87.00]
        [97.00, 87.00, 3.00, 85.00]
        [12.00, 98.00, 67.00, 74.00]
        [98.00, 94.00, 71.00, 7.00]
        [5.00, 28.00, 38.00, 97.00]
        [56.00, 54.00, 25.00, 23.00]
        [45.00, 20.00, 57.00, 19.00]
X = [8.00, 47.00, 98.00, 70.00, 0.00, 0.00, 0.00, 0.00]
 = [35912.00, 0.00, 0.00, 0.00, 0.00, 0.00, 0.00]
```

 Γ ia NxM = 8x8 xai blocksize = 1024

```
cuda17@rncp-ubuntu:~$ ./ex2b_b
A = [
        [17.00, 88.00, 47.00, 98.00, 69.00, 77.00, 61.00, 29.00]
        [42.00, 59.00, 79.00, 29.00, 67.00, 31.00, 37.00, 83.00]
        [91.00, 42.00, 75.00, 67.00, 42.00, 91.00, 91.00, 74.00]
        [32.00, 67.00, 69.00, 49.00, 37.00, 57.00, 48.00, 35.00]
        [77.00, 48.00, 8.00, 4.00, 91.00, 67.00, 84.00, 72.00]
        [67.00, 61.00, 69.00, 96.00, 52.00, 63.00, 71.00, 72.00]
        [15.00, 15.00, 32.00, 36.00, 47.00, 99.00, 5.00, 48.00]
        [62.00, 96.00, 23.00, 41.00, 26.00, 72.00, 49.00, 30.00]
A_trans = [
        [17.00, 42.00, 91.00, 32.00, 77.00, 67.00, 15.00, 62.00]
        [88.00, 59.00, 42.00, 67.00, 48.00, 61.00, 15.00, 96.00]
        [47.00, 79.00, 75.00, 69.00, 8.00, 69.00, 32.00, 23.00]
        [98.00, 29.00, 67.00, 49.00, 4.00, 96.00, 36.00, 41.00]
        [69.00, 67.00, 42.00, 37.00, 91.00, 52.00, 47.00, 26.00]
        [77.00, 31.00, 91.00, 57.00, 67.00, 63.00, 99.00, 72.00]
        [61.00, 37.00, 91.00, 48.00, 84.00, 71.00, 5.00, 49.00]
        [29.00, 83.00, 74.00, 35.00, 72.00, 72.00, 48.00, 30.00]
X = [16.00, 75.00, 28.00, 69.00, 26.00, 86.00, 5.00, 88.00]
Y = [4624.00, 580800.00, 61852.00, 662676.00, 123786.00, 509894.00, 18605.00, 74008.00]
```

2 Προβλήματα

Δεν υλοποίησα την άσκηση 2Β-Γ (συνδιακύμανση).