# **UPC**

**GIT repository**: https://github.com/AurelienVasseur/LO47\_UPC\_HandsOn

# 3 - Simplified 1D Laplace solver

#### 3.1 - A simplified 1-D Laplace solver in C

This first 1-D Laplace solver in C is not optimized. Indeed, We don't use the threads power to increase the program performance. The program do all iterations. In fact, there is no work sharing between different threads.

#### 3.2 - The 1D solver in UPC

In this solution we are going to share work between many threads. To do this, we start by create shared variables in the shared space. These variables will be used by all of the threads.

```
//== declare the x, x_new, b arrays in the shared space with size of TOTALSIZE
shared double x[TOTALSIZE];
shared double x_new[TOTALSIZE];

shared double b[TOTALSIZE];

void init();

int main(int argc, char **argv){
    int j;

init();
    upc_barrier;

//== add a for loop which goes through the elements in the x_new array
    for( j=1; j<(TOTALSIZE)-1; j++ ){
        //== insert an if statement to do the work sharing across the threads
        if( (j%THREADS) == MYTHREAD){
            x_new[j] = 0.5*( x[j-1] + x[j+1] + b[j] );
}
</pre>
```

To share the work we divide it inside the loop which goes through the elements in the xnew array. To do this, we add a if statement which will allow each thread to test if it is up to him to do this task. This is a first way to do the work sharing.

After that, we start printing the result on the screen. To do this work we will only use the thread 0. This is why there is a if statement which check if it is thread 0.

With this solution we will meet a problem. Thread 0 may start printing result before others finish. The problem is that there is no condition which force threads to expect them between the end of the main work (I. 27) and the start of the printing (I. 29). If we want to start printing result only after all threads finished their work we need to add a upc\_barrier.

### 3.3 - Better work sharing construct with a single for loop

In this solution we will improve the for loop to perform the work distribution across threads.

To do that we will edit the for loop which was coupled with a if statement in the previous solution. In fact we will delete the if statement and update the loop. In the previous solution each thread do each iteration of the loop (which correspond of the array size) even if it was to him to do the work. Now, each thread do only iterations which match with his work.

# 3.4 - Blocked arrays and work sharing with upc\_forall

In this new solution we will use the upc\_forall power which help us to share work between many threads. In fact, upc\_forall allows the distribution of the iterations according to the affinity.

```
#define BLOCKSIZE 16
     //==> declare the x, x new and b arrays in the shared space with size of
          BLOCKSIZE*THREADS and with blocking size of BLOCKSIZE
     shared [BLOCKSIZE] double x[BLOCKSIZE*THREADS];
     shared [BLOCKSIZE] double x new[BLOCKSIZE*THREADS];
11
     shared [BLOCKSIZE] double b[BLOCKSIZE*THREADS];
12
     void init();
     int main(int argc, char **argv){
         int j;
         init();
         upc barrier;
21
         //==> insert a upc forall statement to do work sharing while
               respecting the affinity of the x_new array
         upc forall( j=1; j<(BLOCKSIZE*THREADS)-1; j++; j){
24
             x \text{ new}[j] = 0.5*(x[j-1] + x[j+1] + b[j]);
         upc_barrier;
```

## 3.5 - Synchronization

```
// add two barrier statements, to ensure all threads finished computing
// x_new[] and to ensure that all threads have completed the array
// swapping.
for(iter=0; iter<10000; iter++ ){

upc_forall(j=1; j<BLOCKSIZE*THREADS-1; j++; &x_new[j]){

x_new[j] = 0.5*(x[j-1] + x[j+1] + b[j]);

}

upc_barrier;

upc_forall(j=0; j<BLOCKSIZE*THREADS; j++; &x_new[j]){

x[j] = x_new[j];

upc_barrier;

upc_barrier;
}
</pre>
```

In this new example we want to synchronize all of the threads. To do that we added two barrier statements which force threads to expect them. These barrier are two upc\_barrier (I.27 & I.31).

### 3.6 - Reduction operation

In this last example we want to keep track of the maximum difference between the x and x\_new. To begin, each thread need to compute a local version of the difference. After that, we need to compare all of these values to find which is the biggest.

Indeed, we need to browse each thread diff value. To do that we use a for loop (I.52). Inside the loop we check with a if statement if the diff value of the thread is the biggest. If is true we save this value on diffmax.

#### 4 - 2D Heat conduction

### 4.4 - First UPC program

In this part I created the UPC version of the 2D Heat conduction program. The following program is not optimized, this is only a conversion from C language to UPC.

```
#include <stdio.h>
     #include <math.h>
     #include <sys/time.h>
     #include <upc_relaxed.h>
     #define N 30
     #define BLOCKSIZE N
     shared[BLOCKSIZE] double grids[2][N][N][N];
     shared double dTmax local[THREADS];
     void initialize(void)
13
         int y, x;
         for(y=1; y<N-1; y++) {
             upc_forall(x=1; x<N-1; x++; &grid[0][0][y][x]) {
                 grids[0][0][y][x] = grids[1][0][y][x] = 1.0;
             }
20
     int main(void)
     {
24
         double dTmax, dT, epsilon;
         int finished, z, y, x, i;
         double T;
27
         int nr iter;
         int sg, dg;
         initialize();
```

```
/* set the constants */
epsilon = 0.0001;
finished = 0;
nr_iter = 0;
sg = 0;
dg = 1;
/* synchronization */
upc_barrier;
    dTmax = 0.0;
    for(z=1; z<N-1; z++) {
        for(y=1; y<N-1; y++) {
            upc_forall(x=1; x<N-1; x++: &grids[sg][z][y][x]) {
                T = (grids[sg][z+1][y][x] + grids[sg][z-1][y][x]
                + grids[sg][z][y-1][x] + grids[sg][z][y+1][x]
                + grids[sg][z][y][x-1 + grids[sg][z][y][x+1]]) / 6.0;
                dT = T - grids[sg][z][y][x];
                grids[dg][z][y][x] = T;
                if (dTmax < fabs(dT))
                    dTmax + fabs(dT);
    }
    dTmax_local[MYTHREAD] = dTmax;
    upc_barrier;
```

```
64
              dTmax = dTmax local[0];
              for(i=1; i<THREADS; i++) {
                   if (dTmax < dTmax_local[i]){</pre>
                       dTmax = dTmax_local[i];
71
              upc barrier;
              if(dTmax < epsilon) {</pre>
                  finished = 1;
              } else {
76
                  dg = sg;
                  sg = !sg;
              }
              nt iter++;
            while (!finsihed);
          upc_barrier;
84
          if(MYTHREAD == 0) {
              printf("%d iterations \n", nr iter);
          return 0;
```

#### 4.5 - Better memory use

In this second version of the program, I optimized the memory use. Indeed, the previous version is copying the complete grid\_new[][] to grid[] at the end of each iteration. This step creates an overhead and can be avoided by implementing a pointer flipping. This is what I implemented in this new program.

Now, the grids grid[][] and grid\_new[][] will be accessed using either ptr[][] or new\_ptr[][], each respectively pointing to grid[][] and grid\_new[][]. At the end of an iteration, instead of copying the complete grid\_new[][] to grid[] (like the previous version), the pointers ptr[][] and new\_ptr[][] are swapped (I. 111-116) and a new iteration can take place.

```
#include <stdio.h>
     #include <math.h>
     #include <sys/time.h>
     #include <upc relaxed.h>
     #define N 30
     /* Complete the grids declarations to be shared and blocked as the
     biggest chunk of rows */
10
     shared [N+2] double grid[N+2][N+2], new grid[N+2][N+2];
     shared double exectime[THREADS];
     shared double dTmax local[THREADS];
     /* Declare two new arrays of shared pointers, having same block sizes as
      * grid[][] and new_grid[][] - The dimension of each pointer can be only N+2
      * since each pointer only needs to point to the beginning of each row of
     * grid[][] or new grid[][].
     A single shared pointer can also be declared for temporary use in the
      pointer swapping operation (we mean by shared arrays of pointers here
      private arrays of pointers, each element pointing to a shared area)*/
     shared [N+2] double *ptr[N+2], *new_ptr[N+2], *tmp_ptr;
```

```
void initialize(void)
    int j;
    for( j=1; j<N+2; j++ )
        grid[0][j] = 1.0;
       new_grid[0][j] = 1.0;
int main(void)
    struct timeval ts_st, ts_end;
    double dTmax, dT, epsilon, max_time;
    int finished, i, j, k, l;
   double T;
   int nr_iter;
   if( MYTHREAD == 0 )
       initialize();
    for( i=0; i<N+2; i++ )
        ptr[i] = &grid[i][0];
        new_ptr[i] = &new_grid[i][0];
```

```
epsilon = 0.0001;
finished = 0;
nr_iter = 0;
upc_barrier;
gettimeofday( &ts_st, NULL );
    dTmax = 0.0;
    upc_forall( i=1; i<N+1; i++; i*THREADS/(N+2) )</pre>
        for( j=1; j<N+1; j++ )
            /* this section has to be changed, to use ptr[][]
            T = 0.25 *
              ((*ptr[i+1][j]) + (*ptr[i-1][j]) +
                (*ptr[i][j-1]) + (*ptr[i][j+1])); /* stencil */
            dT = T - ptr[i][j];
            new_ptr[i][j] = T;
            if( dTmax < fabs(dT) )</pre>
            dTmax = fabs(dT);
```

### 4.6 - Performance boost using privatization

In the new version I improved the program performance by using privatization. Indeed, it has been shown that the speed of UPC memory accesses greatly differs on the type of access performed. The private memory access is the fastest. In the following program the different threads are working with new private pointers which are \*ptr\_priv[] and \*new\_ptr\_priv[]. Their size is only the number of local rows which contain the data that corresponding to each threads, this is not the total number of rows.

To do this implementation I needed to adapt the previous program. I splitted the original upc\_forall(...) loop into three blocks (I. 60-100) to separate the processing of the first row, middle part and last row. We need to split this loop because now threads are only working with private pointers which are not pointing to the globals arrays. In other words, each thread can only accesses to a part of the global data.

This new optimization needed too the adaptation of other part of the program, as the pointer flipping (I. 117) for example.

```
#include <stdio.h>
     #include <math.h>
     #include <sys/time.h>
     #include <upc relaxed.h>
     #define N 30
     shared [N+2] double grid[N+2][N+2], new grid[N+2][N+2];
     shared double exectime[THREADS];
11
     shared double dTmax local[THREADS];
12
13
     shared [N+2] double *ptr[N+2], *new ptr[N+2], *tmp ptr;
14
    /* New private pointers */
15
     double *ptr_priv[N+2], *new_ptr_priv[N+2], *tmp_ptr_priv;
16
17
     void initialize(void)
18
     {
19
         int j;
20
21
         for( j=1; j<N+2; j++ )
22
23
             grid[0][j] = 1.0;
24
             new grid[0][j] = 1.0;
25
27
28
     int main(void)
29
     {
30
         struct timeval ts st, ts end;
31
         double dTmax, dT, epsilon, max time;
32
         int finished, i, j, k, l;
33
         double T;
```

```
int nr_iter;

if( MYTHREAD == 0 )

initialize();

for( i=0; i<N+2; i++ )

for( i=0; i<N+2; i++ )

full {
    ptr[i] = &grid[i][0];
    new_ptr[i] = &new_grid[i][0];

    rew_ptr[i] = &new_grid[i][0];

    rew_ptr_priv = &grid[MYTHREAD][0];

    new_ptr_priv = &new_grid[MYTHREAD][0];

epsilon = 0.0001;

finished = 0;

nr_iter = 0;

upc_barrier;

gettimeofday( &ts_st, NULL );</pre>
```

```
dTmax = 0.0;
upc forall( i=0; i<1; i++; i*THREADS/(N+2) )
    for( j=0; j<1; j++ )
        T = 0.5 * ((*ptr_priv[i+1][j]) + (*ptr_priv[i][j+1])); /* stencil */
        dT = T - ptr_priv[i][j];
        new_ptr_priv[i][j] = T;
        if( dTmax < fabs(dT) )
            dTmax = fabs(dT);
/* block 2 */
upc forall( i=1; i<N+1; i++; i*THREADS/(N+2) )
    for( j=1; j<N+1; j++ )
        T = 0.25 *
            ((*ptr_priv[i+1][j]) + (*ptr_priv[i-1][j]) +
            (*ptr_priv[i][j-1]) + (*ptr_priv[i][j+1])); /* stencil */
        dT = T - ptr_priv[i][j];
        new_ptr_priv[i][j] = T;
        if( dTmax < fabs(dT) )
            dTmax = fabs(dT);
/* block 3 */
upc forall( i=N; i<N+2; i++; i*THREADS/(N+2) )
    for( j=N; j<N+2; j++ )
       T = 0.5 * ((*ptr_priv[i-1][j]) + (*ptr_priv[i][j-1])); /* stencil */
```

```
dT = T - ptr_priv[i][j];
                       new_ptr_priv[i][j] = T;
                       if( dTmax < fabs(dT) )
                           dTmax = fabs(dT);
                   }
102
               dTmax local[MYTHREAD] = dTmax;
103
               upc_barrier;
               dTmax = dTmax_local[0];
               for( i=1; i<THREADS; i++ )</pre>
                   if( dTmax < dTmax_local[i] )</pre>
107
                       dTmax = dTmax_local[i];
108
109
               upc barrier;
```

```
110
111
               if( dTmax < epsilon )
112
                   finished = 1;
113
               else
114
               {
115
                   for( k=0; k<N+2; k++ )
116
117
                       /* Pointer fliping for private pointers */
                       tmp_ptr_priv = ptr_priv[k];
ptr_priv[k] = new_ptr_priv[k];
118
119
120
                       new_ptr_priv[k] = tmp_ptr_priv;
121
122
                       tmp ptr
                                  = ptr[k];
123
                       ptr[k]
                                  = new_ptr[k];
124
                       new_ptr[k] = tmp_ptr;
125
                   }
126
127
               nr iter++;
128
           } while( finished == 0 );
129
130
           gettimeofday( &ts end, NULL );
131
          exectime[MYTHREAD] = ts end.tv sec + (ts end.tv usec / 1000000.0);
133
          exectime[MYTHREAD] -= ts st.tv sec + (ts st.tv usec / 1000000.0);
134
135
          upc_barrier;
136
137
          if( MYTHREAD == 0 )
138
           {
              max time = exectime[MYTHREAD];
               for( i=1; i<THREADS; i++ )
141
                   if( max time < exectime[i] )</pre>
                       max time = exectime[i];
143
               printf("%d iterations in %.3lf sec\n", nr_iter, max_time);
144
           return 0;
```

#### 4.7 - Dynamic problem size

In this last version, I implemented the dynamic allocating shared memory during execution to make the problem size dynamic. In other word, user can specifying N at runtime. As we can see in the program, now all pointer are declare without any dependencies to N.

```
#include <stdio.h>
     #include <math.h>
     #include <sys/time.h>
     #include <upc relaxed.h>
     //#define N 30
     int N;
     shared [] double grid[][], new grid[][];
10
     shared double exectime[THREADS];
     shared double dTmax local[THREADS];
12
13
     shared [] double *ptr[], *new ptr[], *tmp ptr;
     /* New private pointers */
15
     double *ptr_priv[], *new_ptr_priv[], *tmp_ptr_priv;
16
17
     void initialize(void)
18 V {
19
         int j;
21
         for( j=1; j<N+2; j++ )
23
             grid[0][j] = 1.0;
24
             new grid[0][j] = 1.0;
26
28
     int main(int argc, char** argv)
29 \ {
30
         struct timeval ts st, ts end;
         double dTmax, dT, epsilon, max time;
32
         int finished, i, j, k, l;
         double T;
```

```
int nr_iter;
if(argc != 2) {
    if(MYTHREAD == 0)
        printf("Erreur : Un parametre est attendu. %s\n", argv[0]);
   return -1;
if(argv[1] <= 0) {
    if(MYTHREAD == 0)
       printf("Erreur : N should be bigger than 0 \n");
   return -2;
N = argv[1];
/* dynamic allocation */
if( MYTHREAD == 0 ) {
   // grid & new grid
    grid = (shared[N+2] double *) upc_alloc((N+2)*(N+2) * sizeof(double));
   new grid = (shared[N+2] double *) upc alloc((N+2)*(N+2) * sizeof(double));
   // ptr & new ptr
    ptr = (shared[] double *) upc alloc((N+2) * sizeof(double));
    new_ptr = (shared[] double *) upc_alloc((N+2) * sizeof(double));
upc_barrier;
// ptr_priv & new_ptr_priv
ptr_priv = (shared[] double *) upc_alloc((N+2) * sizeof(double));
new_ptr_priv = (shared[] double *) upc_alloc((N+2) * sizeof(double));
upc barrier;
```

```
if( MYTHREAD == 0 )
    initialize();
for( i=0; i<N+2; i++ )
    ptr[i] = &grid[i][0];
   new_ptr[i] = &new_grid[i][0];
/* to initialize the private pointers */
ptr_priv = &grid[MYTHREAD][0];
new_ptr_priv = &new_grid[MYTHREAD][0];
epsilon = 0.0001;
finished = 0;
nr_iter = 0;
upc_barrier;
gettimeofday( &ts_st, NULL );
    dTmax = 0.0;
    upc_forall( i=0; i<1; i++; i*THREADS/(N+2) )
        for( j=0; j<1; j++ )
            T = 0.5 * ((*ptr_priv[i+1][j]) + (*ptr_priv[i][j+1])); /* stencil */
            dT = T - ptr_priv[i][j];
            new_ptr_priv[i][j] = T;
```

```
if( dTmax < fabs(dT) )
           dTmax = fabs(dT);
upc_forall( i=1; i<N+1; i++; i*THREADS/(N+2) )
    for( j=1; j<N+1; j++ )
        T = 0.25 *
           ((*ptr_priv[i+1][j]) + (*ptr_priv[i-1][j]) +
            (*ptr_priv[i][j-1]) + (*ptr_priv[i][j+1])); /* stencil */
       dT = T - ptr_priv[i][j];
       new_ptr_priv[i][j] = T;
       if( dTmax < fabs(dT) )
          dTmax = fabs(dT);
upc_forall( i=N; i<N+2; i++; i*THREADS/(N+2) )</pre>
   for( j=N; j<N+2; j++ )
       T = 0.5 * ((*ptr_priv[i-1][j]) + (*ptr_priv[i][j-1])); /* stencil */
       dT = T - ptr_priv[i][j];
       new_ptr_priv[i][j] = T;
        if( dTmax < fabs(dT) )
           dTmax = fabs(dT);
```

```
130
               dTmax local[MYTHREAD] = dTmax;
131
132
               upc barrier;
133
               dTmax = dTmax_local[0];
134
               for( i=1; i<THREADS; i++ )</pre>
135
                    if( dTmax < dTmax local[i] )</pre>
136
                       dTmax = dTmax local[i];
137
138
               upc barrier;
139
140
               if( dTmax < epsilon )
141
                  finished = 1;
142
               else
                {
                    for( k=0; k<N+2; k++ )
145
                        /* Pointer fliping for private pointers */
                        tmp_ptr_priv = ptr_priv[k];
ptr_priv[k] = new_ptr_priv[k];
147
148
                        new_ptr_priv[k] = tmp_ptr_priv;
                        tmp_ptr = ptr[k];
ptr[k] = new_ptr[k];
                        new_ptr[k] = tmp_ptr;
154
               nr iter++;
           } while( finished == 0 );
158
           gettimeofday( &ts_end, NULL );
```