

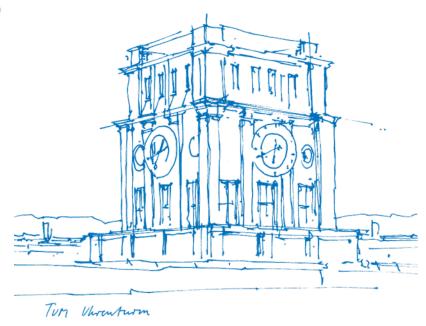
#### Parallel Programming Tutorial - OpenMP Wrap-Up

Amir Raoofy, M.Sc.

Chair for Computer Architecture and Parallel Systems (Prof. Schulz)

Technical University of Munich

16. Mai 2018





Few organizational notes



#### Organization

- Deadline for assignment 5 is extended; May 22nd
- We have an optional assignment on SIMD (will be released next week)
  - Based on guest lecture by Dr. Michael Klemm
- On June 25th we will have a lecture on Optimization of sequential programs by M.Sc. Alexis Engelke Topics: (tentative)
- Compiler optimizations
- Floating-point optimizations
- (Auto-)Vectorization
- Cache optimizations
- Eventually profiling



#### OpenMP Wrap-Up





#### Nested parallel regions revisited

```
#include <iostream>
  #include < omp.h>
  int main(){
       int num_threads=4;
       omp_set_num_threads(num_threads);
       #pragma omp parallel
10
            #pragma omp parallel for
11
           for (int i = 0; i < num_threads; i++)</pre>
12
13
                #pragma omp critical
                std::cout << "My id is: "</pre>
15
                           << omp_get_thread_num() << std::endl;
16
17
19 }
```



#### Nested parallel regions revisited

```
./example4
   #include <iostream>
   #include < omp.h>
                                                                          My id is: 0
   int main(){
                                                                          My id is: 0
                                                                          My id is: 0
        int num_threads=4;
                                                                          My id is: 0
        omp_set_num_threads(num_threads);
                                                                          My id is: 0
                                                                          My id is: 0
        #pragma omp parallel
                                                                          My id is: 0
10
            #pragma omp parallel for
11
                                                                          My id is: 0
            for (int i = 0; i < num_threads; i++)</pre>
12
                                                                          My id is: 0
13
                                                                          My id is: 0
                 #pragma omp critical
14
                 std::cout << "My id is: "
                                                                          My id is: 0
15
                             << omp_get_thread_num() << std::endl;</pre>
                                                                          My id is: 0
16
17
                                                                          My id is: 0
        }
18
                                                                          My id is: 0
19 }
                                                                          My id is: 0
                                                                          My id is: 0
```



## Nested parallel regions revisited (Cont.)

```
#include <iostream>
  #include < omp.h>
  int main(){
       int num_threads=4;
       omp_set_num_threads(num_threads);
       omp_set_nested(1);
       #pragma omp parallel
10
11
           #pragma omp parallel for
12
           for (int i = 0; i < num_threads; i++)</pre>
13
                #pragma omp critical
15
                std::cout << "My id is: "
16
                           << omp_get_thread_num() << std::endl;
17
18
20 }
```



# Nested parallel regions revisited (Cont.)

```
./example5
   #include <iostream>
   #include < omp.h>
                                                                         My id is: 1
   int main(){
                                                                         My id is: 0
                                                                         My id is: 2
        int num threads=4;
                                                                         My id is: 3
        omp_set_num_threads(num_threads);
                                                                         My id is: 1
        omp_set_nested(1);
                                                                         My id is: 2
        #pragma omp parallel
                                                                         My id is: 0
10
11
                                                                         My id is: 1
            #pragma omp parallel for
12
                                                                         My id is: 1
            for (int i = 0; i < num_threads; i++)</pre>
13
                                                                         My id is: 0
14
                                                                         My id is: 3
                 #pragma omp critical
15
                 std::cout << "My id is: "
                                                                         My id is: 2
16
                             << omp_get_thread_num() << std::endl;
17
                                                                         My id is: 3
18
                                                                         My id is: 0
                                                                         My id is: 3
20 }
                                                                         My id is: 2
```





```
#include <iostream>
#include <omp.h>

int main(){

int id;
#pragma omp parallel num_threads(4)

{

id = omp_get_thread_num();
#pragma omp critical

std::cout << "My id is: " << id << std::endl;
}

}</pre>
```



./example

## Quiz; What is the problem with this program?

```
#include <iostream>
#include <omp.h>

int main(){

int id;
#pragma omp parallel num_threads(4)

{

id = omp_get_thread_num();
#pragma omp critical

std::cout << "My id is: " << id << std::endl;
}

</pre>
```



./example

My id is: 0 My id is: 0 My id is: 3 My id is: 2



```
#include <iostream>
#include <omp.h>

My id is: 0

My id is: 0

My id is: 0

My id is: 3

My id is: 3

My id is: 3

My id is: 3

My id is: 2

#pragma omp parallel num_threads(4)

# id = omp_get_thread_num();

#pragma omp critical
#prag
```



```
./example
#include <iostream>
#include <omp.h>
                                                                                         My id is: 0
int main(){
                                                                                         My id is: 0
                                                                                         My id is: 3
    int id;
                                                                                         My id is: 2
    #pragma omp parallel num_threads(4)
         id = omp_get_thread_num();
         #pragma omp critical
                                                                                         ./example
         std::cout << "My id is: " << id << std::endl;
                                                                                         My id is: 2
                                                                                         My id is: 2
                                                                                         My id is: 0
                                                                                         My id is: 0
```



```
#include <iostream>
#include <omp.h>

int main(){

int id;
#pragma omp parallel num_threads(4) private(id)

id = omp_get_thread_num();
#pragma omp critical
std::cout << "My id is: " << id << std::endl;
}

</pre>
```



```
#include <iostream>
#include <omp.h>

int main(){

int id;
#pragma omp parallel num_threads(4) private(id)

{
    id = omp_get_thread_num();
    #pragma omp critical
    std::cout << "My id is: " << id << std::endl;
}
</pre>
```





```
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
       #pragma omp critical
            a++;
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
           printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
       #pragma omp critical
            a++; // -> shared
           b++;
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
           printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
       #pragma omp critical
           a++; // -> shared
           b++; // -> private
           c++;
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
           printf("b: %d\n", b);
           printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
       #pragma omp critical
           a++; // -> shared
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
           printf("b: %d\n", b);
           printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
       #pragma omp critical
           a++; // -> shared
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
           printf("b: %d\n", b);
           printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```

./example



```
./example
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
                                                                                 a: 5
       #pragma omp critical
            a++; // -> shared
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
./example
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
                                                                                 a: 5
       #pragma omp critical
                                                                                 b: ?
            a++; // -> shared
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
./example
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
                                                                                  a: 5
       #pragma omp critical
                                                                                  b: ?
                                                                                  c: 4
            a++; // -> shared
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
./example
int main (){
   int a =1, b =2, c =3;
   #pragma omp parallel num_threads(4) private(b) firstprivate(c)
   {
                                                                                  a: 5
       #pragma omp critical
                                                                                  b: ?
                                                                                  c: 4
            a++; // -> shared
                                                                                  a: 5
           b++; // -> private
           c++; // -> firstprivate
       #pragma omp barrier
       if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
   printf("a: %d\n", a);
   printf("b: %d\n", b);
   printf("c: %d\n", c);
   return 0;
```



```
./example
int main (){
    int a =1, b =2, c =3;
    #pragma omp parallel num_threads(4) private(b) firstprivate(c)
    {
                                                                                  a: 5
        #pragma omp critical
                                                                                  b: ?
                                                                                  c: 4
            a++; // -> shared
                                                                                  a: 5
            b++; // -> private
                                                                                  b: 2
            c++; // -> firstprivate
        #pragma omp barrier
        if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
    printf("a: %d\n", a);
    printf("b: %d\n", b);
    printf("c: %d\n", c);
    return 0;
```



```
./example
int main (){
    int a =1, b =2, c =3;
    #pragma omp parallel num_threads(4) private(b) firstprivate(c)
    {
                                                                                  a: 5
        #pragma omp critical
                                                                                  b: ?
                                                                                  c: 4
            a++; // -> shared
                                                                                  a: 5
            b++; // -> private
            c++; // -> firstprivate
                                                                                  b: 2
                                                                                  c: 3
        #pragma omp barrier
        if (omp_get_thread_num()==0){
            printf("a: %d\n", a);
            printf("b: %d\n", b);
            printf("c: %d\n", c);
    printf("a: %d\n", a);
    printf("b: %d\n", b);
   printf("c: %d\n", c);
    return 0;
```



```
int a=1;
void parallel_function()
    int b=2, c=3;
    #pragma omp parallel shared(b)
    #pragma omp parallel private(b)
        int d=4;
        #pragma omp task
            int e=5;
            a
```



```
int a=1;
void parallel_function()
    int b=2, c=3;
   #pragma omp parallel shared(b)
    #pragma omp parallel private(b)
        int d=4;
        #pragma omp task
            int e=5;
            a // shared
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared -> a=1
           b
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate
                                -> b=?
           С
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate
                                -> b=?
              // shared
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate -> b=?
             // shared
                               -> c=3
           С
           d
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                            -> a=1
              // firstprivate
                                -> b=?
             // shared
                                -> c=3
              // firstprivate
```





```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate -> b=?
             // shared
                        -> c=3
              // firstprivate
                               -> d=4
           е
```





## Quiz; Task data scoping

```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                           -> a=1
              // firstprivate -> b=?
             // shared
                       -> c=3
              // firstprivate
                               -> d=4
             // private
```





## Quiz; Task data scoping

```
int a=1;
void parallel_function()
   int b=2, c=3;
   #pragma omp parallel shared(b)
   #pragma omp parallel private(b)
       int d=4;
       #pragma omp task
           int e=5;
           a // shared
                            -> a=1
              // firstprivate
                                -> b=?
             // shared
                        -> c=3
              // firstprivate
                                \rightarrow d=4
             // private
                                -> e=5
```



# Quiz; Coarse-grained parallelization

```
1 #define N 10000
2 #define ITER 100
3 double A[N + 2][N + 2];
5 int main(int argc, char **argv)
      for (int i = 0; i < N + 2; i++)
                                       // Initialization
          for (int j = 0; j < N + 2; j++)
              A[i][i] = 0.0;
      for (int i = 0; i < N + 2; i++){ // Boundary conditions
          A[i][0] = 1.0; A[i][N + 2] = 1.0;
13
      for (int n = 0; n < 100; n++){ // Main iteration loop
17
          for (int i = 1; i < N + 1; i++)
              for (int j = 1; j < N + 1; j++)
                  A[i][j] = (A[i+1][j+1] + A[i-1][j-1] + A[i+1][j-1] + A[i-1][j+1])/4;
      return 0;
23
```



# Quiz; Coarse-grained parallelization

```
1 #define N 10000
2 #define ITER 100
3 double A[N + 2][N + 2];
5 int main(int argc, char **argv)
      for (int i = 0; i < N + 2; i++)
                                        // Initialization
          for (int j = 0; j < N + 2; j++)
              A[i][i] = 0.0;
      for (int i = 0; i < N + 2; i++){ // Boundary conditions
          A[i][0] = 1.0; A[i][N + 2] = 1.0;
13
      for (int n = 0; n < 100; n++){ // Main iteration loop
          #pragma omp parallel for
                                           // Coarse-grained parallelization
17
          for (int i = 1; i < N + 1; i++)
              for (int j = 1; j < N + 1; j++)
                  A[i][j] = (A[i+1][j+1] + A[i-1][j-1] + A[i+1][j-1] + A[i-1][j+1])/4;
      return 0;
23
```



# Quiz; Coarse-grained parallelization

```
1 #define N 10000
2 #define ITER 100
3 double A[N + 2][N + 2];
5 int main(int argc, char **argv)
      #pragma omp parallel for
                                             // First touch
      for (int i = 0; i < N + 2; i++)
                                              // Initialization
          for (int j = 0; j < N + 2; j++)
              A[i][i] = 0.0;
      for (int i = 0; i < N + 2; i++){ // Boundary conditions
          A[i][0] = 1.0; A[i][N + 2] = 1.0;
13
      for (int n = 0; n < 100; n++){ // Main iteration loop
          #pragma omp parallel for
                                            // Coarse-grained parallelization
17
          for (int i = 1; i < N + 1; i++)
              for (int j = 1; j < N + 1; j++)
                  A[i][j] = (A[i+1][j+1] + A[i-1][j-1] + A[i+1][j-1] + A[i-1][j+1])/4;
      return 0;
23
```





## Typical patterns that come up in parallel programming

- Loop parallelization (Worksharing)
  - Parallelize the for loops that are time consuming in the code
  - Make sure the loops are parallelizable (dependency analysis)
  - Put the pragmas and take care of the data attributes

#### • Example:



# Typical patterns that come up in parallel programming (Cont.)

- Divide and conquer and unstructured parallelism (Tasking)
  - Split the problem into subproblems
  - Solver the subproblems in parallel
  - Fits the Tasking in OpenMP (v3 and later)

#### • Example:

5

10

12

13

15

```
struct node
                                                           // main
                                                           #pragma omp parallel
    struct node* left;
    struct node* right;
};
                                                               #pragma omp single
                                                               traverse(root);
void traverse( struct node*p ) {
    if(p->left)
        #pragma omp task
        traverse(p->left);
    if(p->right)
        #pragma omp task
        traverse(p->right);
    process(p);
```





```
template <typename SrcView, typename DstView>
void x_gradient(const SrcView &src, const DstView &dst, int num_threads)
      int start = 0:
      int chunk_size = 16;
      std::mutex mtx;
      std::vector<std::thread> threads;
      for (int i = 0; i < num_threads; ++i)</pre>
          threads.push_back(std::thread(x_gradient_kernel < SrcView , DstView > , \
                                         std::ref(src), std::ref(dst), \
                                         std::ref(start), std::ref(chunk size), std::ref(mtx)));
      }
      for (auto &th : threads)
          th.join();
```





# Solution for Assignment 3 (Cont.)



# Solution for Assignment 3 (Cont.)

```
while (true)
           mtx.lock();
           local_start = start;
           if (src.height() - start < 1) {</pre>
               mtx.unlock(); break;
           if (src.height() - local_start < chunk_size)</pre>
               chunk size = src.height() - local start;
           start += chunk size;
           mtx.unlock();
           for (int y = local start; y < local start + chunk size; ++y) {</pre>
               typename SrcView::x_iterator src_it = src.row_begin(y);
               typename DstView::x_iterator dst_it = dst.row_begin(y);
17
               for (int x = 1; x < src.width() - 1; ++x)
                   static_transform(src_it[x - 1], src_it[x + 1], dst_it[x],
                                     halfdiff_cast_channels<dst_channel_t>());
```





```
template <typename SrcView, typename DstView>
void x_gradient(const SrcView& src, const DstView& dst, int num_threads) {
    typedef typename channel_type<DstView>::type dst_channel_t;
    omp_set_num_threads(num_threads);
    #pragma omp parallel for schedule (dynamic,5)
    for (int y=0; y<src.height(); ++y) {</pre>
        typename SrcView::x_iterator src_it = src.row_begin(y);
        typename DstView::x_iterator dst_it = dst.row_begin(y);
        for (int x=1; x<src.width()-1; ++x) {</pre>
             static_transform(src_it[x-1], src_it[x+1], dst_it[x],
                              halfdiff cast channels < dst channel t > ());
```



#### Hints for Assignment 5



#### Hints for Assignment 5

- If you use sections:
  - Pay attention to the recursive structure of traverse
  - Make sure that nesting is enabled
  - Stop nesting at a specific level to prevent creating so many parallel regions
- If you use tasks:
  - Again, pay attention to the recursive structure of traverse
  - Make sure that you create appropriate number of tasks
  - Try to restrict the number of created tasks using the appropriate clauses



Assignment 6 - Laplace 2D



## Assignment 6 - Laplace 2D

- 2d Laplace equation with fixed boundaries
- Problem domain is unit square with uniform mesh
- Finite differences are used for the discretization
- We use Jacobi iterative method to solve the equation
- Look into the code and find the bottlenecks
- Use OpenMP to parallelize the solver
- You need to get a speedup of 16 on our server with 32 logical cores
- The server has 2 NUMA nodes each with 8 cores
- Pay attention to data locality on the cores



#### Assignment 6 - Laplace 2D - Provided Files

- Makefile
  - contains rules to build executables
  - available targets: parallel, sequential, unit\_test, all (default), clean
  - 'mode=debug make [target]' to build debug version, use 'make clean' before
- main.c
  - main function argument handling + call initialization of arrays and main iteration loop
- laplace.h
  - Header and definitions for the arrays
- laplace\_seq.h
  - Sequential version of time\_step().
- student/laplace\_par.h
  - Implement the parallel version in this file
- unit\_test.c
  - The unit tests that execute both the serial and parallel version to compare results.