

OpenMP Course

Victor Eijkhout

2024 COE 379L / CSE 392



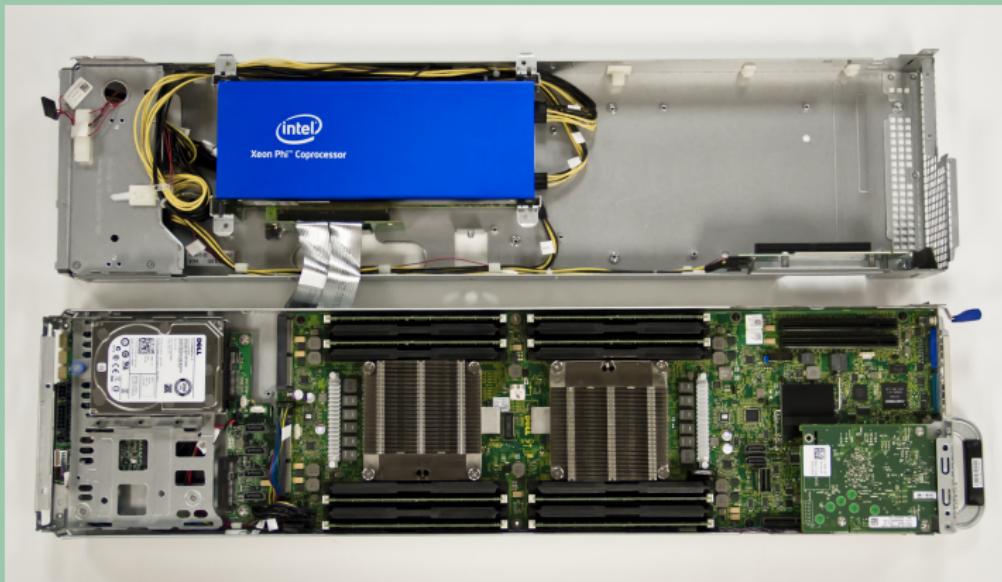
OpenMP is a flexible tool for incrementally parallelizing a shared memory-based code. This course introduces the main concepts through lecturing and exercises.



The Fork-Join model

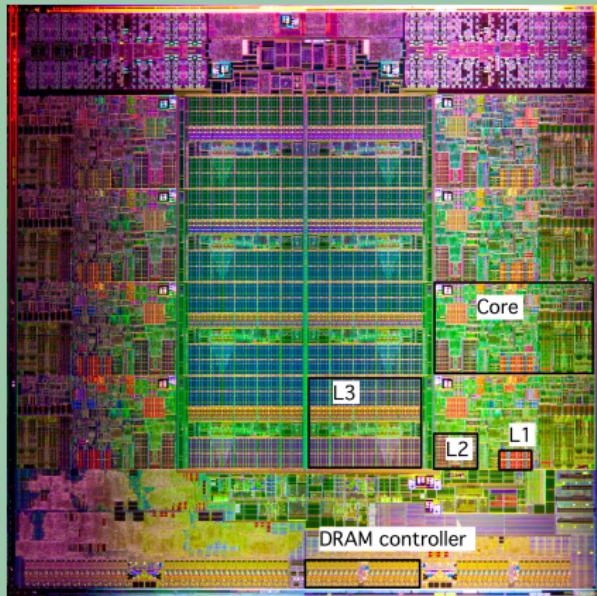


One cluster node:



A node will have 1 or 2 or (sometimes) 4 'sockets': processor chips.
There may be a co-processor attached.



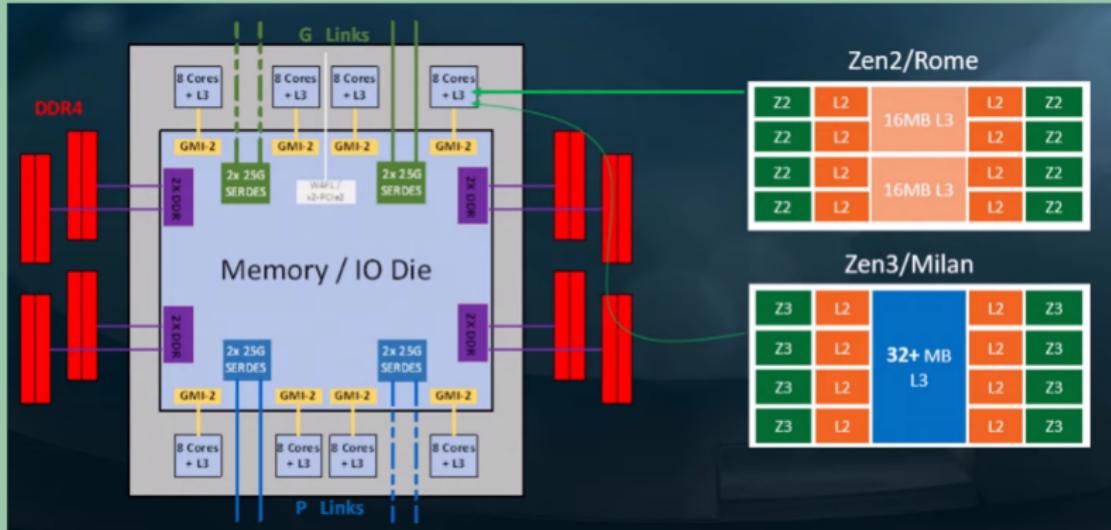


- Eight cores per socket, 16 per node.
- Access to DRAM, both this socket and other
- Shared L3 cache, private L2 and L1



Modern chip design

TACC



- Shared memory
- Shared caches
- Local caches
- 'cc-Numa': Cache-coherent Non-Uniform Memory Access

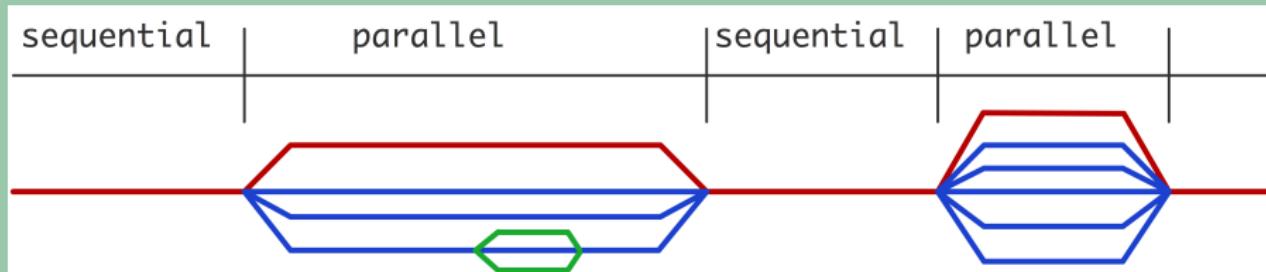


- First standard 1997, Fortran directives
- Current standard 5.2 as of 2021
- Intel Core processors: 2006;
before that: 1980s systems from Alliant, Sequent, Denelcor;
SGI Origin with up to 1024 processors
supercomputers with multi-chip nodes (Intel Paragon, Cray T3D)



Process: stream of instructions

Thread: process can duplicate itself, same code, access to same data



The OS will place threads on different cores: parallel performance.

Note: threads are software. More threads than cores or fewer is allowed.

Ignore hyperthreads



To write an OpenMP program

TACC

```
1 #include "omp.h"
```

in C, and

```
1 use omp_lib
```

or

```
1 #include "omp_lib.h"
```

for Fortran.



```
# gcc
gcc -o foo foo.c -fopenmp
# Intel compiler 23 and later
icc -o foo foo.c -fopenmp
# Intel compiler before 23
icc -o foo foo.c -qopenmp
```

Run as ordinary program.



```
1 find_package(OpenMP)
2 # depending on language:
3 target_link_libraries( ${program} PUBLIC OpenMP::OpenMP_C )
4 target_link_libraries( ${program} PUBLIC OpenMP::OpenMP_CXX )
5 target_link_libraries( ${program} PUBLIC OpenMP::OpenMP_Fortran )
```



How many cores do we have?

```
1 int omp_get_num_procs();
```

Compile and run



```
1 int omp_get_num_threads();
```

Compile and run. Observe?



To run an OpenMP program



```
export OMP_NUM_THREADS=8  
./my_omp_program
```

Quick experiment:

```
for t in 1 2 4 8 12 16; do  
    OMP_NUM_THREADS=$t ./my_omp_program  
done
```



Use a parallel region:

```
1 #pragma omp parallel  
2 {  
3     statements  
4 }
```

or

```
1 !$omp parallel  
2     statements  
3 !$omp end parallel
```

Compile and run. Observe?



What happens if you write

```
1 int t;
2 #pragma omp parallel
3 {
4     t = omp_get_thread_num();
5     printf("%d\n", t);
6 }
```

```
1 {
2     int t; t = // ...
3     printf( ... );
4 }
```

```
1 int t;
2 #pragma omp parallel private(t)
```



Write a ‘hello world’ program, where the print statement is in a parallel region. Compile and run.

Run your program with different values of the environment variable `OMP_NUM_THREADS`. If you know how many cores your machine has, can you set the value higher?



Take the hello world program of exercise 1 and insert the above functions, before, in, and after the parallel region. What are your observations?



What happens if I press that button?

TACC

Who of you has tried setting the number of threads (much) larger than the number of cores? What happened?



Data is (usually) shared in a parallel region:

Data is visible in enclosed scopes:

```
1 main() {
2     int x;
3 #pragma omp parallel
4     {
5         // you can use and set 'x' here
6     }
7     printf("x=%e\n",x); // value depends on what
8                     // happened in the parallel region
9 }
```

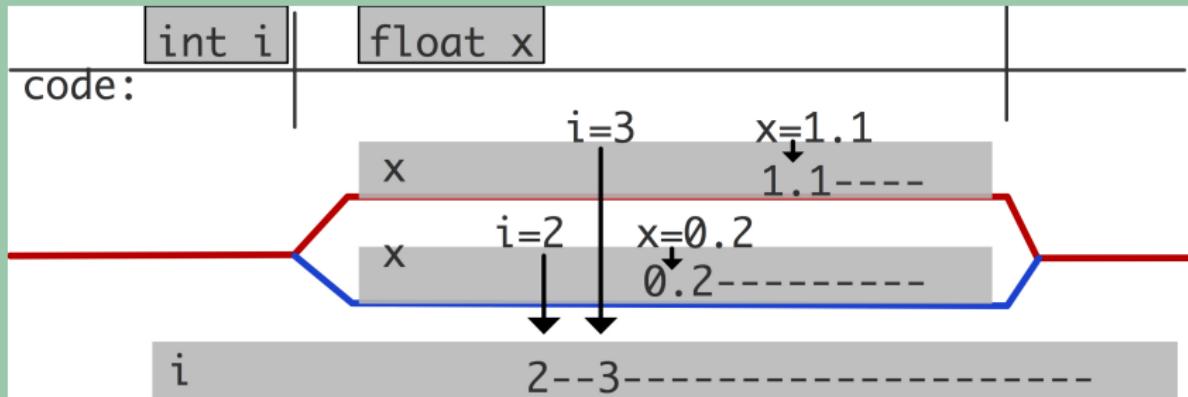
Note: this is usually the case, but strictly not the default.



```
1  {
2      int x;
3 #pragma omp parallel
4      {
5          double x;
6          // do something with x
7      }
8  }
```

- A parallel region is a scope.
- Local variables are per thread.





- Environment variable `OMP_NUM_THREADS=4`
- Clause `num_threads(4)`

```
1 #pragma omp parallel num_threads(4)
```

usually not the best solution.

- Environment variable `OMP_THREAD_LIMIT=56`
(useful for nested parallelism)



```
1 #pragma omp parallel if(N>50)
```

Can you think of an application?

I mean, what's the harm of doing one small loop inefficiently?



- Threads are software, cores are hardware.
- The OS can move threads between cores: not a good idea for performance.
- Set: `export OMP_PROC_BIND=true` and you'll be good in most of the cases.
- Look up 'affinity' in the OMP standard for all the details.



```
1 int main() {                                1 void func(...) {
2     ...                                     2 #pragma omp parallel
3 #pragma omp parallel                      3 {
4     {                                         4     ...
5     ...                                     5     }
6     func(...)                               6 }
7     ...                                     7
8 } } // end of main
```



By default, nested levels get only one thread.

Use:

```
1 void omp_set_max_active_levels(int);  
2 int omp_get_max_active_levels(void);
```

or

OMP_MAX_ACTIVE_LEVELS=3

(Deprecated: boolean switch OMP_NESTED)



Loop parallelism



Much of parallelism in scientific computing is in loops:

- Vector updates and inner products
- Matrix-vector and matrix-matrix operations
- Finite Element meshes
- Multigrid



- Suppose loop iterations are independent:
- Distribute them over the threads:
- Use `omp_get_thread_num` to determine disjoint subsets.

```
1 #pragma omp parallel
2 {
3     int threadnum = omp_get_thread_num(),
4         numthreads = omp_get_num_threads();
5     int low = N*threadnum/numthreads,
6         high = N*(threadnum+1)/numthreads;
7     for (int i=low; i<high; i++)
8         // do something with i
9 }
```

Discuss.



C: directive followed by statement or block:

```
1 #pragma omp parallel
2 {
3 #pragma omp for
4     for ( int i=0; i<N; i++)
5         ... something with i ...
6 }
```

Fortran: matching end directive

```
1 !$omp parallel
2 !$omp do
3     do i=1,n
4         ... something with i ...
5     end do
6 !$omp end do
7 !$omp end parallel
```



Here's the two-step parallelization in OpenMP:

- You use the `parallel` directive to create a team of threads;
- then you use a ‘workshare’ construct to distribute the work over the team.
- For loops that is the `for` (or `do`) construct.



Parallel and workshare together

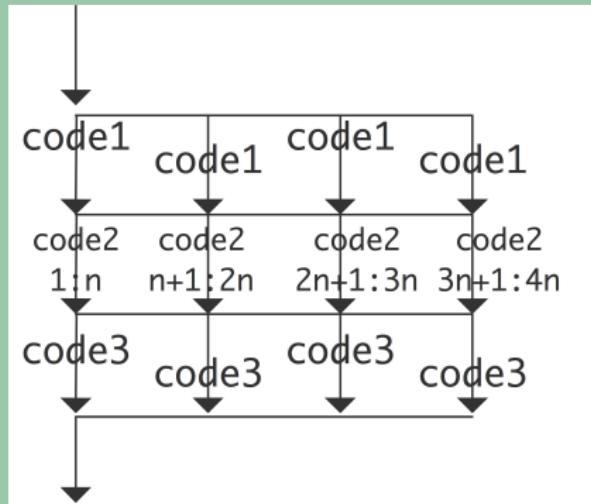
```
1 #pragma omp parallel for
2     for (i=0; i<N; i++)
3         ... something with i ...
```



Stuff inside a parallel region

TACC

```
1 #pragma omp parallel
2 {
3     code1();
4     #pragma omp for
5     for (int i=1; i<=4*N; i++) {
6         code2();
7     }
8     code3();
9 }
```



- No `break` or so.
- `continue` or `cycle` allowed.
- Fixed upper bound;
- Simple loop increment;
- ⇒ OpenMP needs to be able to calculate the number of iterations in advance.
- No `while` loops.



When is a loop parallel?

TACC

It's up to you!

This is parallelizable:

```
1   for (int i=low; i<hi; i++)
2       x[i] = // expression without x
```

Is this?

```
1   for (int i=low; i<hi; i++)
2       x[ f(i) ] = // expression
```

Histograms / binning.



Consider the code

```
1  for (int i=0; i<n; i++)  
2      x[i/2] += f(i)
```

Argue that this does not satisfy the above condition. Can you rewrite this loop to be parallelizable?



Run the following snippet

```
1 // vectorsum.c
2 for ( int iloop=0; iloop<500; ++iloop ) {
3     for ( int i=0; i<vectorsize; ++i )
4         outvec[i] += invec[i]*factor;
5     factor /= 1.1;
6 }
```

1. Sequentially
2. On one thread
3. With more than one thread.

Do a scaling study. What all can you do to improve performance? Read back through previous slides.

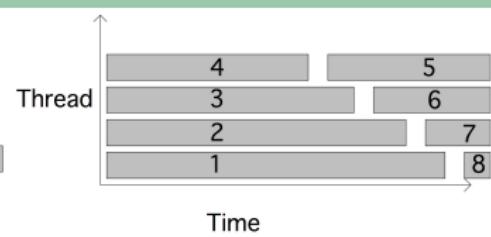
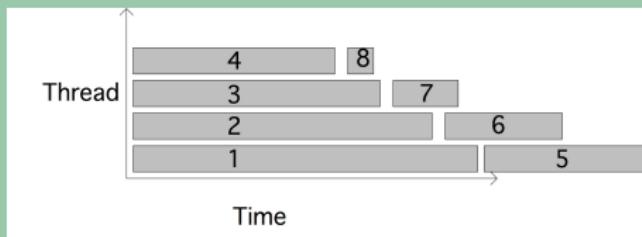


More about loops



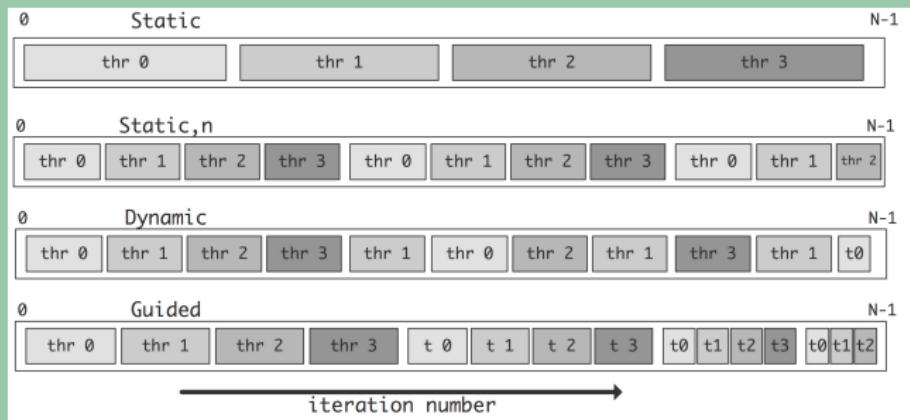
- Static scheduling of iterations.
(default in practice though not formally)
Very efficient. Good if all iterations take the same amount of time.
`schedule(static)`
- Other possibility: dynamic.
Runtime overhead; better if iterations do not take the same amount of time.
`schedule(dynamic)`

Four threads, 8 tasks of decreasing size
dynamic schedule is better:



With N iterations and t threads:

- Static: each thread gets N/t iterations.
explicit chunk size: `schedule(static,123)`
- Dynamic: each thread gets 1 iteration at a time
explicit chunk size: `schedule(dynamic,45)`



Use decreasing chunk size (with optional minimum chunk):
`schedule(guided, 6)`

More flexible than dynamic, less overhead than dynamic.



- Parallelize loop nest
- More iterations \Rightarrow better performance
- Inner loop needs to be directly nested
bounds simple function of outer bounds and var

```
1 // triangle.c
2 #pragma omp parallel for collapse(2)
3     for ( int i=0; i<vectorsize; ++i ) {
4         for ( int j=i+1; j<vectorsize; ++j ) {
5             double
6                 dx= particles[XCOORD(i)]-particles[XCOORD(j)],
7                 dy= particles[YCOORD(i)]-particles[YCOORD(j)];
8             double r = sqrt( dx*dx + dy*dy ),
9                 f = (particles[MASS(i)]*particles[MASS(j)])/(r*r);
10            }
11        }
```



- Ordered iterations: normally OpenMP can execute iterations in any sequence. You can force ordering if you absolutely have to. Bad for performance!



- Barrier: synchronize all threads
- explicit: `omp barrier`
- implicit: end of a workshare construct such as `for`
- remove implicit barriers: `nowait` clause



```
1 #pragma omp parallel for nowait
2 for ( int i=0; i<N; ++i )
3     x[i] = // something
4 #pragma omp parallel for
5 for ( int i=0; i<N; ++i )
6     y[i] = ... x[i] ...
```

Needed:

- Same number of iterations
- Same schedule



Loop data



Statements such as:

```
1 #pragma omp parallel private(x)
2 #pragma omp parallel shared(y)
3 #pragma omp parallel default(none) shared(x) private(y)
```

- Shared data: comes from outside the parallel region
- Private data: either private versions of outside, or local defined
- Default is shared. Dangerous!
- Parallel loop variables are always private.



Where is the bug?

TACC

```
1     int i,j;
2     double temp;
3 #pragma omp parallel for private(temp)
4     for(i=0;i<N;i++){
5         for (j=0;j<M;j++){
6             temp = b[i]*c[j];
7             a[i][j] = temp * temp + d[i];
8     } }
```

Is there a different way of handling `temp`?



```
1 float x=5.5;
2 #pragma omp parallel firstprivate(x)
3     // x is now private, and has value 5.5
4
5 float y;
6 #pragma omp parallel for lastprivate(y)
7 for ( /* looping */ ) {
8     // stuff
9     y = something;
10 }
11 // now y has the sequentially last value
```



Reduction



Extend the program from exercise 2. Make a complete program based on these lines:

Code:

```
1 // reduct.c
2 int tsum=0;
3 #pragma omp parallel
4 {
5     tsum += // expression
6 }
7 printf("Sum is %d\n",tsum);
```

Output:

```
1 With 4 threads, sum
2 →s/b 6
3 Sum is 6
4 Sum is 5
5 Sum is 1
6 Sum is 4
7 Sum is 6
8 Sum is 5
9 Sum is 6
10 Sum is 5
11 Sum is 3
12 Sum is 4
```

Compile and run again. (In fact, run your program a number of times.) Do you see something unexpected? Can you think of an explanation?



Race condition: simultaneous update of shared data:

process 1: $I = I + 2$

process 2: $I = I + 3$

Results can be indeterminate:

scenario 1.	scenario 2.	scenario 3.
	$I = 0$	
read $I = 0$ compute $I = 2$ write $I = 2$	read $I = 0$ compute $I = 3$ write $I = 3$	read $I = 0$ compute $I = 2$ write $I = 2$
$I = 3$	$I = 2$	$I = 5$



- Inner product loop: **missing snippet inprodsum**
- Use the `reduction(+:s)` clause.
- All the usual operations are available; you can also make your own.



Compute π by *numerical integration*. We use the fact that π is the area of the unit circle, and we approximate this by computing the area of a quarter circle using *Riemann sums*.

- Let $f(x) = \sqrt{1 - x^2}$ be the function that describes the quarter circle for $x = 0 \dots 1$;
- Then we compute

$$\pi/4 \approx \sum_{i=0}^{N-1} \Delta x f(x_i) \quad \text{where } x_i = i\Delta x \text{ and } \Delta x = 1/N$$

Write a program for this, and parallelize it using OpenMP parallel for directives.

1. Put a `parallel` directive around your loop. Does it still compute the right result? Does the time go down with the number of threads? (The answers should be no and no.)
2. Change the `parallel` to `parallel for` (or `parallel do`). Now is the result correct? Does execution speed up? (The answers should now be no and yes.)
3. Put a `critical` directive in front of the update. (Yes and very much no.)
4. Remove the `critical` and add a clause `reduction(+:quarterpi)` to the `for` directive. Now it should be correct and efficient.



We continue with exercise 6. We add 'adaptive integration' where needed, the program refines the step size¹. This means that the iterations no longer take a predictable amount of time.

1. Use the `omp parallel for` construct to parallelize the loop. As in the previous lab, you may at first see an incorrect result. Use the `reduction` clause to fix this.
2. Your code should now see a decent speedup, but possibly not for all cores. It is possible to get completely linear speedup by adjusting the schedule.
Start by using `schedule(static,n)`. Experiment with values for n . When can you get a better speedup? Explain this.
3. Since this code is somewhat dynamic, try `schedule(dynamic)`. This will actually give a fairly bad result. Why? Use `schedule(dynamic,n)` instead, and experiment with values for n .
4. Finally, use `schedule(guided)`, where OpenMP uses a heuristic. What results does that give?

¹It doesn't actually do this in a mathematically sophisticated way, so this code is more for the sake of the example.



1. Use the `omp parallel for` construct to parallelize the loop. As in the previous lab, you may at first see an incorrect result. Use the reduction clause to fix this.
2. Your code should now see a decent speedup, using up to 8 cores. However, it is possible to get completely linear speedup. For this you need to adjust the schedule.
Start by using `schedule(static,n)`. Experiment with values for n . When can you get a better speedup? Explain this.
3. Since this code is somewhat dynamic, try `schedule(dynamic)`. This will actually give a fairly bad result. Why? Use `schedule(dynamic,n)` instead, and experiment with values for n .
4. Finally, use `schedule(guided)`, where OpenMP uses a heuristic. What results does that give?
5. `schedule(auto)` : leave it up to the system.
6. `schedule(runtime)` : leave it up to environment variables; good for experimenting.



```
1 // reductarray.c
2 #pragma omp parallel for schedule(static,1) \
3     reduction(+:data[:nthreads])
4 for (int it=0; it<nthreads; it++) {
5     for (int i=0; i<nthreads; i++)
6         data[i]++;
7 }
```



```
1 int *alloced = (int*)malloc( nthreads*sizeof(int) );
2 for (int i=0; i<nthreads; i++)
3     alloced[i] = 0;
4 #pragma omp parallel for schedule(static,1) \
5     reduction(+:alloced[:nthreads])
6 for (int it=0; it<nthreads; it++) {
7     for (int i=0; i<nthreads; i++)
8         alloced[i]++;
9 }
```



Each thread gets a copy of the array on the stack
⇒ possible stack overflow
set `OMP_STACKSIZE`
also see `ulimit` on the Unix level.



Exercises



The Jacobi method for solving linear system $Ax = b$ is given by

$$x_i^{(n+1)} = a_{ii} \left(b_i - \sum_{j \neq i} a_{ij} x_j^{(n)} \right)$$

Insert OpenMP directives in the code, and check that it converges with the same precision, regardless the number of threads.

Study speedup. Does the problem size play a role?



How much performance improvement do you get from considering

- removing barriers by `nowait` clauses
- affinity
- first-touch



Experiment with scheduling options.
Do you see any effect?
In particular try small chunk sizes.



Can you put the whole iteration loop in a parallel region?

Does this give further performance improvement?



Workshare constructs



- The `omp parallel` creates a team of threads.
- Now you need to distributed work among them.
- Already seen: `for`, `do`
- Similar: `sections`
- Not obvious: `single`
- Fortran only: `workshare` (works with array notation, but compiler support seems mediocre)
- Story in itself: `task`



Independent separate calculations:

```
double fx = f(x), gx = g(x), hx = h(x);  
..... fx ... gx ... hx .....
```

```
#pragma omp sections  
{  
    #pragma omp section  
    fx = f(x)  
    #pragma omp section  
    gx = g(x)  
    #pragma omp section  
    hx = h(x)  
}
```

Adding them together:

```
s = f(x)+g(x)+h(x);
```



```
int a;  
#pragma omp parallel  
{  
    #pragma omp single  
        a = f(); // some computation  
    #pragma omp sections  
        // various different computations using 'a'  
}
```

- Is executed by a single thread.
- Has implicit barrier, so the result is available to everyone after.
- master is similar, does not have barrier.



What is the difference between this approach and how the same computation would be parallelized in MPI?



Thread data



You have already seen some of the basics:

- Data declared outside a parallel region is shared.
- Data declared in the parallel region is private.
(Fortran does not have this block scope mechanism)

```
int i;  
#pragma omp parallel  
{ double i; .... }
```

- You can change all this with clauses:

```
int i;  
#pragma omp parallel private(i)
```



```
int i; double t;  
#pragma omp parallel for  
for (i=0; i<N; i++) {  
    t = sin(i*pi*h);  
    x[i] = t*t;  
}
```

- The loop variable is automatically private.
- The temporary `t` is shared, but conceptually private to each iteration: needs to be declared private.
(What happens if you don't?)



- Private data is uninitialized

```
int i = 3;  
#pragma omp parallel private(i)  
    printf("%d\n",i); // undefined!
```

- To import a value:

```
int i = 3;  
#pragma omp parallel firstprivate(i)  
    printf("%d\n",i); // undefined!
```

- lastprivate to preserve value of last iteration.



- `default(shared)` or `default(private)`
- useful for debugging: `default(none)`
because you have to specify everything as shared/private



- Private data disappears after the parallel region.

What if you want data to persist?

- Directive `threadprivate`

```
double seed;  
#pragma omp threadprivate(seed)
```

- Standard application: random number generation.
- Tricky: has to be global or static.



- Statically allocated arrays can be made private.
- Dynamically allocated ones can not: the pointer becomes private.



Synchronization



- The loop and sections directives do not specify an ordering, sometimes you want to force an ordering.
- Barriers: global synchronization.
- Critical sections: only one process can execute a statement this prevents race conditions.
- Locks: protect data items from being accessed.



- Every workshare construct has an implicit barrier:

```
#pragma omp parallel
{
    #pragma omp for
    for ( .. i .. )
        x[i] = ...
    #pragma omp for
    for ( .. i .. )
        y[i] = ... x[i] ... x[i+1] ... x[i-1] ...
}
```

First loop is completely finished before second.

- Explicit barrier:

```
#pragma omp parallel
{
    x = f();
    #pragma omp barrier
    .... x ...
}
```



- Critical section: One update at a time.

```
#pragma omp parallel
{
    double x = f();
#pragma omp critical
    global_update(x);
}
```

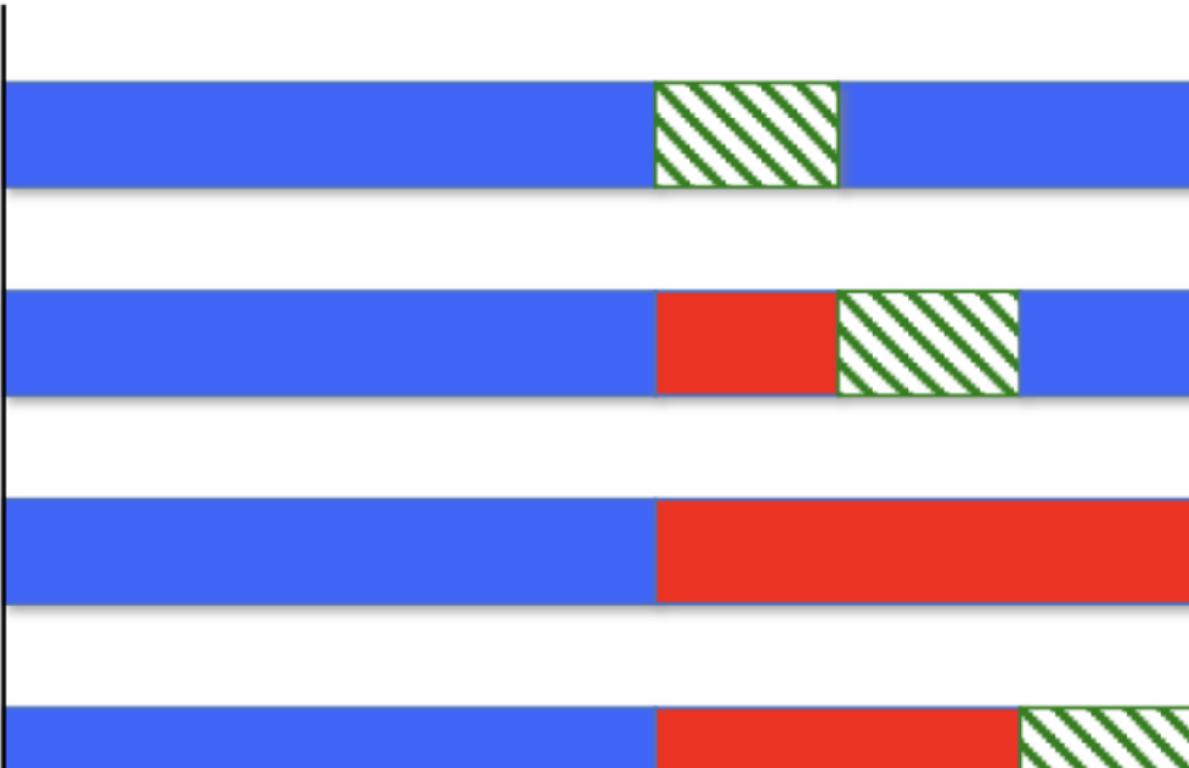
- `atomic` : special case for simple operations, possible hardware support

```
#pragma omp atomic
    t += x;
```



- Critical sections are not cheap! The operating system takes thousands of cycles to coordinate the threads.
- Use only if minor amount of work.
- Do not use if a reduction suffices.
- Name your critical sections.
- Explore locks if there may not be a data conflict.





- Critical sections are coarse:
they dictate exclusive acces to a *statement*
- Suppose you update a big table
updates to non-conflicting locations should be allowed
- Locks protect a single data item.



Tasks

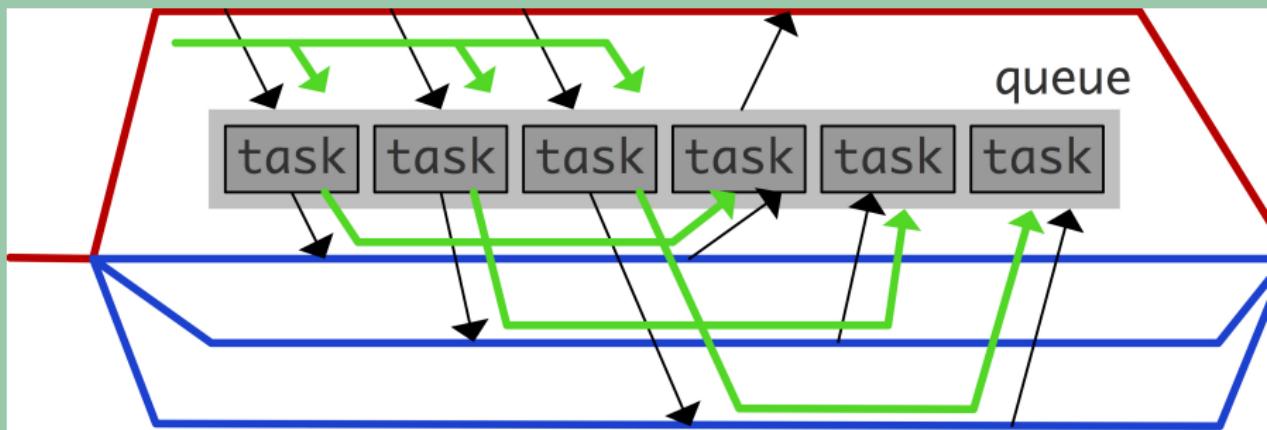


- You have seen loops and sections.
- How about linked lists or trees?
- Tasks are very flexible:
you create work, it goes on a queue, gets executed later

```
p = head_of_list();
while (!end_of_list(p)) {
#pragma omp task
    process( p );
    p = next_element(p);
}
```



- There is one queue (per team), not visible to the programmer.
- One thread starts generating tasks.
- Tasks can recursively generate tasks.
- You never know who executes what.



Use tasks to find the smallest factor of a large number (using $2999 \cdot 3001$ as test case): generate a task for each trial factor. Start with this code:

```
1  int factor=0;
2  #pragma omp parallel
3  #pragma omp single
4  for (int f=2; f<4000; f++) {
5      { // see if 'f' is a factor
6          if (N%f==0) { // found factor!
7              factor = f;
8          }
9      }
10     if (factor>0)
11         break;
12 }
13 if (factor>0)
14     printf("Found a factor: %d\n",factor);
```

- Turn the factor finding block into a task.
- Run your program a number of times:

```
for i in `seq 1 1000` ; do ./taskfactor ; done | grep -v 299
```



Mechanisms for task synchronization:

- `taskwait`: wait for all previous tasks (not nested)
- `taskgroup`: wait for all tasks, including nested
- `depend`: synchronize on data items.



```
int process( node n ) {  
    if (n.is_leaf)  
        return n.value;  
    for ( c : n.children) {  
#pragma omp task  
        process(c);  
#pragma omp taskwait  
        return sum_of_children();  
    }  
}
```



Example: Fibonacci

TACC

```
long fib(int n) {  
    if (n<2) return n;  
    else { long f1,f2;  
#pragma omp task  
    f1 = fib(n-1);  
#pragma omp task  
    f2 = fib(n-2);  
#pragma omp taskwait  
    return f1+f2;  
}  
  
#pragma omp parallel  
#pragma omp single  
printf("Fib(50)=%ld",fib(50));
```

(what is conceptually wrong with this example?)



```
long fibs[100];
void fib(n) {
    if (n>=2) {
        #pragma omp task \
            depend( in:fibs[n-2],in:fibs[n-1] ) \
            depend( out:fibs[n] )
        fibs[n] = fibs[n-2]+fibs[n-1];
    };
}

#pragma omp parallel
#pragma omp single
    for (i<50)
        fib(i);
```



Memory model



fill in



```
1  a = foo();  
2  flag = 1;
```

```
1  while (!flag) ;  
2  b = a;
```



```
1  a = foo();  
2  #pragma omp flush  
3  #pragma omp atomic write  
4  flag = 1;  
5  #pragma omp flush
```

```
1  myflag = 0;  
2  while (!myflag){  
3      #pragma omp flush  
4      #pragma omp atomic read  
5      myflag = flag;  
6  }  
7  #pragma omp flush  
8  b = a;
```



Affinity and other memory issues



- If you have two threads and two cores, how do you place the threads?
- The OS can ‘migrate’ threads. Is that good?



NUMANode P#0 (92GB)

Package P#0

L3 (39MB)

L2 (1024KB)

L1d (32KB)

L1i (32KB)

Core P#0

PU P#0

Core P#6

PU P#2

Core P#1

PU P#4

Core P#5

PU P#6

Core P#2

PU P#8

Core P#4

PU P#10

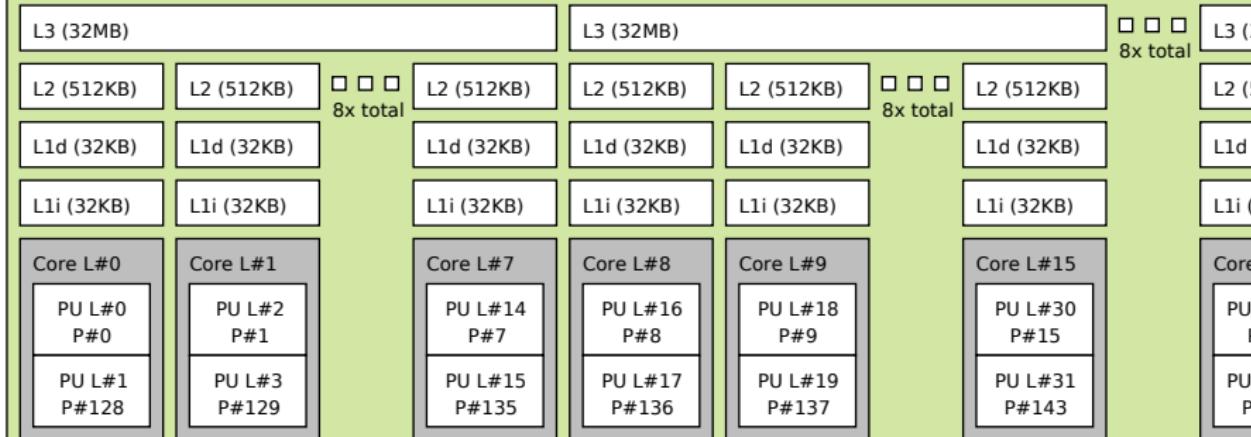
Core P#3

PU P#12



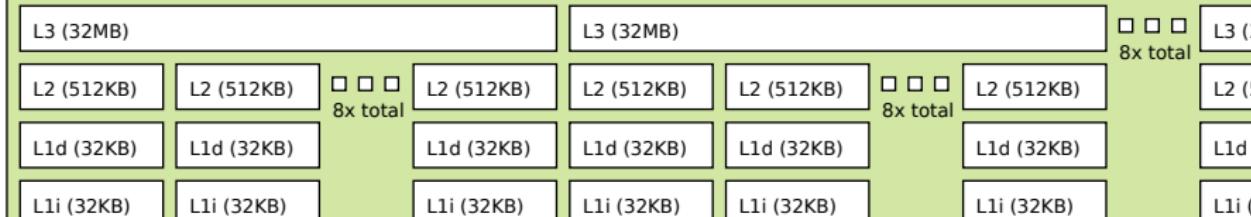
Package L#0

NUMANode L#0 P#0 (125GB)



Package L#1

NUMANode L#1 P#1 (126GB)



- `OMP_PROC_BIND`: either `true/false` or `close/spread`
- `OMP_PLACES`: `sockets/cores/threads` or explicit list.



amask



- Report for an affinity env.
 - Executable command: `amask_<type>` types:{**omp**, **mpi**, **hybrid**}.
 - Can instrument code: `amask_<type>()` argumentless function calls.

Usage: `module load amask`

```
export OMP_<affinity>= ...
export I_MPI_<affinity>= ...
export OMP_<affinity>= ... \
    I_MPI_<affinity>= ...
```

`amask_omp`
`ibrun amask_mpi`
`ibrun amask_hybrid`

-- for pure OpenMP
-- for pure MPI
-- MPI + OpenMP



Remaining topics



OpenMP 4 has mechanisms for offloading.



Processors have 4 or 8-wide SIMD.
convert OpenMP loop to SIMD vector instructions.

