



Offloading Computation to a GPU with OpenMP

Kamer Kaya, Sabancı University





OpenMP (Open Multi-Processing) is a portable, scalable programming model used to develop parallel applications on shared-memory architectures. It provides a parallelization interface in C, C++, and Fortran. OpenMP

- uses pragmas (#pragma) to specify parallel regions and tasks, requiring minimal code changes.
- uses threads that share a global address space, simplifying data sharing and communication.
- supports load balancing with multiple scheduling strategies for loops and tasks.

OpenMP: Overview



Common directive:

#pragma omp parallel: Defines

 a parallel region where threads
 are created.

```
#include <omp.h>
#include <iostream>
int main() {
    #pragma omp parallel
    {
        int thread_id = omp_get_thread_num();
        std::cout << "Hello from thread " << thread_id << std::endl;
    }
    return 0;
}</pre>
```

Example: Computing Pi with OpenMP





Example: Computing Pi with OpenMP



• Let's work on this OpenMP code.





Thanks



This project has received funding from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 101101903. The JU receives support from the Digital Europe Programme and Germany, Bulgaria, Austria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, Greece, Hungary, Ireland, Italy, Lithuania, Latvia, Poland, Portugal, Romania, Slovenia, Spain, Sweden, France, Netherlands, Belgium, Luxembourg, Slovakia, Norway, Türkiye, Republic of North Macedonia, Iceland, Montenegro, Serbia