**中山大学本科生期末考试**

**考试科目：《并行程序设计》（C卷）**

学年学期：**2022**学年第**2**学期 姓 名：

开课单位：中山大学计算机学院 学 号：

考试方式：闭卷 年 级：

考试时长：**120**分钟 院 系：

警示 《中山大学授予学士学位工作细则》第八条：“考试作弊者，不授予学士学位。”

------------以下为试题区域，共###道大题，总分100分,考生请在答题纸上作答------------

1. **简答题（共 8 小题，每小题 5 分，共 40 分）**
2. [5分] 简要说明单指令多数据流（SIMD）和单程序多数据流（SPMD）的区别与联系。

硬件 软件 实现

同时执行 指令 程序（但每个处理单元可能执行不同的指令流）

数据并行任务 任务并行任务

1. [5分] 什么是阿姆达尔定律（Amdahl’s law）？它告诉了我们什么？

如果一个程序中只有一部分可以并行化（即 PPP 部分），而其余部分必须串行执行，则整体性能提升受到 PPP 的限制。即使我们增加处理器数目 NNN，最终的加速比也受到 PPP 的影响。

1. [5分] 如何理解多线程编程中的线程安全性？如何解决线程安全性问题？

线程安全性指的是在多个线程并发执行时，程序仍然能够正确地工作而不会产生不可预测的结果或者破坏数据结构的完整性。

1. [5分] 简要说明线程间同步的方式（两种即可）。

路障

互斥锁

条件变量（忙等待

1. [5分] CUDA编程中，可用于同步线程的机制有哪些？试举例说明其中三种。

\_\_syncthreads()用于在一个线程块内部同步所有线程

cudaDeviceSynchronize()在宿主机（Host）代码中调用的函数，用于同步所有在当前设备上的CUDA核函数执行

grid.sync() 是 Cooperative Groups 提供的线程组同步操作，确保所有线程在继续执行后续操作之前完成了乘法运算

1. [5分] 试说明原子操作的概念，并以一个CUDA中的原子操作函数为例，说明其在实际应用中的作用。

原子操作（Atomic Operation）是一种不可中断的操作，它可以在多个线程或者多个处理器之间提供对共享内存或者全局内存的安全并发访问。

不可分割性：不会被中断 原子性：要么全部执行要么不执行

对共享计数器或者共享变量进行并发更新时

1. [5分] 简述在CUDA编程中，主机内存、设备内存、及统一内存寻址的概念，并说明统一内存寻址的作用。

主机内存：存储CPU执行的程序代码、数据以及操作系统运行所需的数据。

设备内存：存储GPU执行的程序代码和数据。设备内存通常是显卡上的全局内存或者共享内存，与主机内存是物理上分离的。

统一内存：将主机内存和设备内存统一管理起来，使得程序员可以使用统一的地址空间访问数据。1程序员无需手动管理主机内存和设备内存的数据迁移2更简洁地编写

1. [5分] 简述存储体冲突的概念，结合具体例子说明如何消除存储体冲突。

存储体冲突指的是多个线程同时访问共享内存中的不同存储体单元（Bank），但这些存储体单元映射到了同一个物理存储体（Bank），从而导致需要额外的时间来调度这些访问，从而影响并行访问的效率。

1重新组织数据结构 2填充

1. **应用题（共 3 小题，每小题 5 分，共 15 分）**

根据以下代码片段，回答以下有关于循环依赖与OpenMP调度的问题：

|  |
| --- |
| #pragma omp parallel for num\_threads(4) schedule(static, 2)  for (i = 0; i < N; i++) {  A[i] = A[i] + B[i];  B[i + 1] = C[i] + D[i];  } |

1.1 [ 5分] 由于循环依赖的存在，该并行代码无法产生期望的结果（与去掉OpenMP编译指导语句的串行版本一致的运行结果），请解释原因。

1.2 [5分] 若N = 16，请问各个线程分别计算数组A的哪些元素？

1.3 [ 5分]若N的值很大，且事先未知。实验发现，将schedule(static, 2)中的第二个参数trunk\_size由2改为一个较大的值后（如10000），程序的性能有所提升。此时，还应修改原先的static调度方式以实现负载均衡。请问应将调度方式改为dynamic或guided中的哪一种？为什么？

1. **编程题（共 1 小题，每小题 15 分，共 15 分）**

1.5CM

现在有一个单精度浮点类型的数组A[N]，已知该数组的数据都分布范围为[0,10)。为了对这些数据的分布有一个更为直观的感受，可以画一个直方图：首先把这些数据的范围划分成10个同等大小的区间，也就是10个桶，然后统计每个桶中元素的数目。现在需要你编写一个MPI并行程序，实现函数histogram\_gen，该函数的功能是完成上述过程，并将每个桶中元素的数目通过printf函数输出到标准输出。请注意，数组A保存在Rank 0中，你需要想办法把它分配到各个进程。

需要实现的函数原型：void histogram\_gen(float\* A, int N, MPI\_Comm comm)

MPI参考函数接口



1. **CUDA代码分析（共 2 小题，每小题 15 分，共 30 分）**

1.5CM

1. 线程执行分析：

\_\_global\_\_ void func(float\* input, float\* output, int size) {

int tid = threadIdx.x + blockIdx.x \* blockDim.x;

if (tid \* 2 > size) return;

if (tid %2){

output[tid] = input[tid\*2]\*input[tid\*2]+input[tid\*2+1]\*input[tid\*2+1];

} else {

output[tid] = input[tid\*2]\*input[tid\*2]-input[tid\*2+1]\*input[tid\*2+1];

}

}

* 1. 说明以上代码功能。（4分）
  2. 说明其效率存在的问题，如何改进，写出改进后的源码。（11分）

2. 代码分析：

1. 说明以下代码的功能；（5分）
2. 说明其效率存在的问题，如何改进，写出改进后的源码；（10分）

\_\_global\_\_ void func(int \* out, int\* in, int n, int m){

int x = blockIdx.x\*blockDim.x+threadIdx.x;

int y = blockIdx.y\*blockDim.y+threadIdx.y;

out[x\*n+y] = in[y\*m+x];

}

**排版要求**：试卷一级标题须标注分值，至少占两行，字体加粗。行距适当留疏，格式整齐一致，保持卷面美观。各题之间不留答题区域，考生一律将答案写在专用答题纸（本）上。**试卷头版式不可改变,内容可根据考试科目实际需要增减**。

**说明：**此份仅为试卷模板，教师需根据实际内容将红色字迹替换。