

中国科学技术大学

2017 -- 2018 学年第 二 学期考试试卷

考试科目: 并行计算 得分: _____

学生所在系: _____ 姓名: _____ 学号: _____

一、2018 年 6 月 25 日,最新一期 Top500 全球超算排行榜在 ISC18 大会正式发布,落户美国橡树岭国家实验室的超级计算机 Summit,登顶全球最快超级计算机。这台计算机是由 IBM 总包设计建设,采用了 9216 颗 IBM Power9 处理器和 27648 颗 Nvidia Volta 的 GPU 加速器。

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	DOE/SC/Oak Ridge National Laboratory United States	Summit - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband IBM	2,282,544	122,300.0	187,659.3	8,806
2	National Supercomputing Center in Wuxi China	Sunway TaihuLight - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway NRCP	10,649,600	93,014.6	125,435.9	15,371
3	DOE/NNSA/LLNL United States	Sierra - IBM Power System S922LC, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband IBM	1,572,480	71,610.0	119,193.6	
4	National Super Computer Center in Guangzhou China	Tianhe-2A - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000 NUDT	4,981,760	61,444.5	100,678.7	18,482

图 1 Top500 排名前四名 (2018 年 6 月)

- (1) 请问图 1 中的 R_{max} 和 R_{peak} 各代表什么含义? 这两个值分别是如何得到的?
- (2) 排名 1 的 Summit 和排名 3 的 Sierra 都是由 IBM 生产。排名第四的 Tianhe-2A 生产单位 NUDT 的中文名称是什么?
- (3) 排名 1 的 Summit 中使用了 Infiniband。它是用来实现系统的什么功能?
- (4) 请推断 Summit 的体系结构类型是什么? 这一类的体系结构有什么基本特点?

二、2017 年 11 月 20 日《中国科学报》有题为“中国超算应用蝉联“戈登贝尔”奖”的新闻报道,其中部分内容下所示。

在 11 月 17 日（当地时间 16 日下午 1 时）于美国丹佛举行的全球超级计算大会（SC 2017）上，基于我国超级计算机“神威·太湖之光”的应用“非线性地震模拟”获得年度“戈登·贝尔”奖，这一奖项被称为超算应用领域的“诺奖”。这是我国继去年在国际上首次摘得戈登贝尔奖之后，再次揽获该奖。

据介绍，地震模拟工具可实现对地震发生过程的重现与预测模拟，是科学家理解地质构造与地震发生与传播原理的一个重要工具，对于降低与预防地震灾害所带来的巨大损失具有重要作用。据悉，该成果联合清华大学、山东大学、南方科技大学、中国科学技术大学、国家并行计算机工程技术研究中心和国家超级计算无锡中心等共同完成。项目团队基于“神威·太湖之光”设计实现了高可扩展性的非线性地震模拟工具。该工具充分发挥国产处理器在存储、计算资源等方面的优势，实现高达 18.9PFlops 的非线性地震模拟，也是国际上首次实现如此大规模的高分辨率、高频率的非线性可塑性地震模拟。该工具首次实现了对唐山大地震（M7.8，1976）发生过程的高分辨率精确模拟，使科学家可以更好地理解唐山大地震所造成的影响，并对未来地震预防预测等研究具有重要的借鉴意义。

今年（2017 年 11 月），“神威·太湖之光”实现全球超算 TOP 500 榜单的“四连冠”。就此，中国已连续 10 年领跑该榜单。与 TOP 500 着眼于高速计算硬件性能不同，戈登贝尔奖更注重于高性能计算应用水平。去年，基于“神威·太湖之光”的应用“千万核可扩展大气动力学全隐式模拟”实现了我国在此大奖上零的突破。

- (1) 结合图 1 中的数据，计算文中介绍的应用“非线性地震模拟”的机器利用率。
- (2) 文中提及了高性能计算的两个应用（地震模拟、大气模拟），请再举出 2 个高性能计算的重要应用。
- (3) “神威·太湖之光”使用的国产处理器“申威 26010”众核处理器。试简要分析使用特定构架国产处理器，在应用软件兼容性上的优势与不足。
- (4) 应用“非线性地震模拟”工作团队的主要负责人来自于清华大学地球系统科学系，其中也有我校地空学院的科研人员。请就此简要说明高性能计算应用软件开发的特点。

三、描述 Amdahl 加速比定律和 Gustafson 加速比定律。分别给出定律的推导过程，并通过对比说明两个定律的意义。

四、以下是上三角方程组回代解法的串行算法的形式化描述。

```
Begin
  (1)for i=n downto 1 do
    (1.1) $x_i = b_i / a_{ii}$ 
    (1.2)for j=1 to i-1 do
       $b_j = b_j - a_{ji}x_i$ 
       $a_{ji} = 0$ 
    endfor
  endfor
End
```

- (1) 请指出串行算法哪些部分可以并行化。
- (2) 写出并行算法的描述，并分析算法的时间复杂度（需要注明计算模型类型）。

五、阅读下面的程序，并回答问题。

```
#include "mpi.h"
void main( int argc, char** argv )
{
    int rank, size, tag=1;
    int senddata,recvdata;
    MPI_Status status;
    MPI_Init(&argc, &argv);
    MPI_Comm_rank(MPI_COMM_WORLD, &rank);
    MPI_Comm_size(MPI_COMM_WORLD, &size);
    if (rank==0){
        senddata=9999;
        MPI_Send( &senddata, 1, MPI_INT, 1, tag, MPI_COMM_WORLD);
    }
    if (rank==1)
        MPI_Recv(&recvdata, 1, MPI_INT, 0, tag, MPI_COMM_WORLD,
        &status);
    MPI_Finalize();
}
```

- (1) 请简述该程序的基本逻辑和功能。
- (2) 描述程序中使用到的 MPI 函数的功能，说明函数中参数的含义。
- (3) 请描述该程序运行的功能，并说明该程序是否可能会出现运行时错误。

六、请给出超立方（CC）的网络连接定义。写出 SIMD-CC 上的矩阵乘法的任意一种并行算法描述。分析该算法的时间复杂度。

七、有一个整数有序序列 $A[1..n]$ ，设计一个 SIMD-EREW 并行算法求出每个元素 $A[i]$ 的等值集合中的位序及等值集合的基，要求用 n 个处理器在 $O(\log n)$ 时间内完成。为帮助理解题意，给出一个计算实例如下。

A	2	2	2	5	7	7	9	9	9	9	9
位序	1	2	3	1	1	2	1	2	3	4	5
基	3	3	3	1	2	2	5	5	5	5	5