

# 并行计算 作业1

PB19050978 方谈

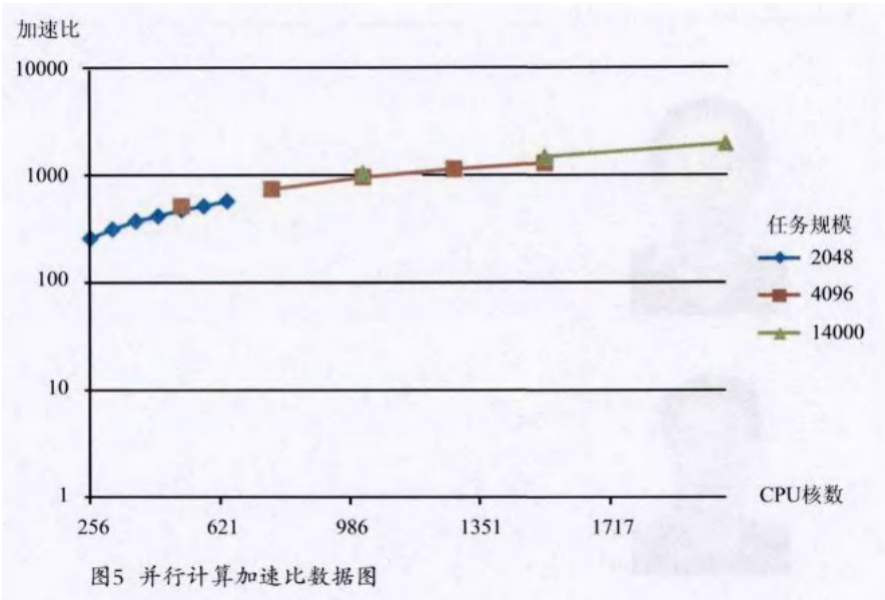
## 1.课本35页第1.11题。并行计算应用调研。

- 1.11 查阅资料,找出一个并行计算的典型应用,详细描述该应用在并行化方面成功和失败之处以及遇到的困难。试从下列方面考虑:
- ① 该应用是针对什么学科或者工程上的具体问题设计的?
  - ② 对于要解决的问题,该应用实际效果怎样? 模拟结果和物理结果进行比较的结果如何?
  - ③ 该应用的运行在什么并行计算平台上(如分布式、共享内存或向量机)? 这个应用是使用哪种开发工具开发的?
  - ④ 和所运行平台最佳性能相比较,该应用的实际工作性能怎样?
  - ⑤ 该应用的可扩展性如何? 如果不好,你认为它的扩展性的瓶颈在何处?

答: 调查的典型应用为解决指纹重卡的计算瓶颈, 见文章《应用并行计算解决指纹重卡计算瓶颈》, 吴春生, 迟学斌, 冯才刚, 王彦桐, 《科研信息化技术与应用》第2卷第6期。

针对北京200万人的指纹库, 按一般情况预估需要约42年, 在当时最好的机器上单进程跑需要两年, 而使用深腾7000-8核128结点进行并行计算用时2月。深腾7000采用Cluster架构、刀片服务器节点、Intel处理器、Linux操作系统、Ethernet或Infiniband互连等。该应用使用C++开发, 运行在分布式非共享内存的系统存储器, 访存模型为UMA。

该应用加速比如下图所示, 加速比略低于核数, 表现较好。极限性能为158Tflops, 该应用所用的所有结点的总性能与运行时的性能未知。该应用可扩展性强, 该文章发表于2011年, 查阅资料得知后续开展了1000万人的指纹重卡归并工作, 由图可知进行过更多核与更大任务规模的测试。除此之外, 在指纹相关的其他方面也有并行算法的应用, 如“基于MMX技术的指纹图像并行处理算法”、“基于CNN通用编程的指纹识别算法研究”等等。该应用可扩展性瓶颈在于并行计算时不同计算节点之间的信息交流问题与文件共享的问题等。



## 2.根据调研的应用需求，预算100万人民币购置计算机设备。请确定购买的机器配置。

答：查阅网上信息知深腾7000已停产，且无法查询到其报价，但能查询到其具体配置信息，见[深腾7000 百度百科](#)。显然100万是无法配置一个完整的深腾7000超级计算机的，仅从该应用实际需求考虑配置一台中型计算机。处理器使用12代英特尔酷睿Intel i9-12900K，16核24线程，共需要64个，每个5000元，共计32万。刀片节点IBM HS21SM，参考报价15000元，8台共12万。厚节点IBM System x3750 M4(8722A1C)，16万，需要1台。胖节点SGI Altix4700报价未知，按其配置降低标准配置一台可用的胖节点，约10万应该可以配齐。可视化结点一台，按其配置预估2万即可。存储系统使用中科曙光DS800-G35，24万一台。剩余4万用于配置网络、线路等等，若不够可适当降低处理器数量，或寻求更低价可替代的刀片节点与厚节点。软件开销、存储系统适配、人工成本、供电与散热等其他因素不在该题考虑范围内。

## 参考文献

【1】吴春生，迟学斌，冯才刚，王彦桐，《应用并行计算解决指纹重卡计算瓶颈》，《科研信息化技术与应用》第2卷第6期，2011年11月。

【2】“深腾7000”，百度百科。