

并行计算HW1

并行计算、分布式计算、云计算的异同

相同点：并行计算和分布式计算都是运用并行来获得更高性能，把大人物化为小任务。两者都属于高性能计算。从技术层面来看，云计算可以视为分布式计算的一种体现和发展。

不同点：并行计算强调时效性，分布式计算不强调。并行计算任务之间不相互独立，而是每个节点之间任务需要通讯和同步；分布式计算的任务之间相互独立，每个节点任务之间没有时间和先后的限制，节点间可以不通讯甚至无网络连接。从计算机用户的角度来说，并行计算是由一个用户完成，分布计算是由多个用户完成，云计算是由服务器完成，用户不参加计算。

并行计算应用调研

并行计算应用于金融模型构建和优化，例如量化交易、投资决策等领域处理因子挖掘和模型训练等任务：

(1) 该问题是针对金融领域

(2) 使用GPU进行并行计算能够极大缩短因子挖掘和模型训练时间，从而提高因子挖掘和模型开发效率。

并行计算协助在产品推荐、股票诊断等业务领域中，更加快速、准确识别用户投资偏好，并给予相应的投资建议，有效地提高客户满意度和留存率等关键业务指标。模拟结果比较物理结果真实有效。

(3) 以中信建投为例，使用了K8S (Kubernetes，一种容器编排引擎)+GPU池化的解决方案，有效的缓解了GPU资源管理难度大，利用率低的问题。也就是采用了分布式架构Kubernetes，一个分布式的集群管理系统。这个平台是用Golang编写的。

(4) 实际工作性能比较好，很有效地利用了容器编排引擎。

(5) 可拓展性比较好，一方面由于Kubernetes的设计理念是支持水平扩展，特性包括自动扩展、资源调度、服务发现和负载均衡等，可以满足不同业务需求的资源管理和调度要求，轻松地增加或减少计算节点，从而实现对GPU资源的弹性扩展。这种水平扩展性可以满足不同规模和需求的业务场景，并保持系统的稳定性和高可用性，而且Kubernetes作为开源项目，拥有庞大的社区支持和丰富的生态系统，可以获得各种扩展功能和解决方案，从而进一步提升系统的可拓展性和灵活性。另一方面GPU池化增加了GPU资源的可共享性。

按预算1000万人民币购置计算设备，请给出采购清单

(1) AMD Epyc 9654P 处理器 * 10 约85万：CPU核心数96，中央处理器线程数192，CPU速度96 x 2.4 GHz，存储器带宽460.8 千兆字节/秒

(2) 三星 旗舰SSD 990 Pro 4TB固态硬盘 * 10 约2.3元

(3) 三星32Gb DDR5内存芯片 * 120，（估价）20万

(3) 思科 (Cisco) Catalyst 9600 系列交换机 * 1 约5.4万

(4) 软件及服务，包括Linux系统许可，并行计算框架等费用若干，约10万

(5) 高速网络架构TH Express-2 估价300万

总价为：422.7万