陈子忠 老师的讲座内容

蟒山讲堂 2023.5.20 Sat

Reliable Matrix Computations via Algorithm-Based Fault Tolerance

基于算法容错的可靠矩阵运算

总体呈现了陈老师团队所发表的8-10篇高性能计算领域顶刊内容

1. 引言

计算中的错误（Error）普遍存在

1. 软错误（Error-continue）：计算错误、静态的数据错误
2. 硬错误（Error-Stop）：错误导致了程序运行终止，也就是导致了宕机

陈老师从以上问题展开，展示了他们团队的思考：（1）能否在运行时就判断计算是否正确

（2）能否在运行时修正错误（3）能否不作备份的情况下找回数据

1. 工作展示1:ABFT
2. 由于矩阵丞的复杂度为O（n^3），但检验方法上的复杂度只有O（n^2），于是检验的代价相对来说就可以忽略了。但是从idea上来说需要在其中找到求和关系即Checksum，计算中就依据此来进行更正。从而展开了陈老师团队的主要工作：FT-BLAS（一个支持算法容错的先行代数库，发表在ICS’21）
3. 对于这个工作，还有提到的应用点有：FFT、CNN、排序、地震预测等。基于此产生了对浮点数存储方面的错误分析、可扩展性分析、通过伪代码展示的共轭梯度
4. 工作展示2:Crash---主要关于无备份恢复
5. 稠密矩阵与稀疏矩阵分别有处理方法
6. 主要基于线性变化的关系：r = b – A \* X；这样的关系在线性运算就天然存在使用代价很小，无需备份出错了仍然可以找回数据，总体上大量省了开销
7. 关于程序执行时间在不同错误率的对比：随着错误率的提升，陈老师团队的方法优势会越来越大，产生量级上的提升

Q&A环节：

1. 计算机系2017级校友 蒋佳祺：关于8比特检验的问题，得到了陈老师的充分肯定，说这是个很好的问题，也在他们的研究工作中其实有考虑到。陈老师做了关于检验机制原理上的解释：即首先发现错误，然后定位错误，最后修正错误。
2. 其余Q&A ：研二王赫萌、研二付旭、研一赵雨轩等人提了问
3. 刘院长的提问：Q1从中科院的工作TensorAlpha引入，这是一个人工智能辅助研究的里程碑，问陈老师如何看待AI取代科研工作者的问题

陈老师A1 人本身是进步的，主要还是在运用种种工具和人工智能；AI做的再好终究还是为人服务的、受人控制的，始终是作为人类的工具和助手而存在的，不会取代工作者。