工作日志 03-19-2018

Problem

- 今日开始研究有关“并行矩阵乘法算法在不同拓扑的计算机集群上的应用”；
- 首先研究拓扑上通信行为的代数化表示，方便根据不同拓扑、节点数据、硬件情况建立一一对应的数学模型。

Action

- 至今日上午 11 点，完成论文《并行矩阵乘法的一般构建及其分析》的初稿，并提交予邓越凡教授审查中；
- 至中午 12 点，下载整理了邓教授团队之前所研究的拓扑结构，分别有 16k2,16k3,16k4,32k2,32k3,32k4，共计 13 个特殊拓扑结构的邻接矩阵文件；
- 至下午 1 点半，开始编写拓扑节点通信的 python 模拟程序，并完成邻接矩阵读入的部分；
- 至下午四点，完成两种通信方式的模拟：
  (1) 不阻塞 + 连续线性空间： $Data_{k+1} = \left(M_{adj} + I_p\right) \cdot Data_k$
  (2) 阻塞 + 不连续条件

$$Data_{k+1} = \left(A_k + I_p\right) \cdot Data_k$$
$$A_{k+1} = A_k * \left(Ones(P,P) - (\text{int})(Data_{k+1} \geq 1) \cdot Ones(1,P)\right)$$
$$A_0 = M_{adj}$$

Keep

- 两种通信方式的模拟各有优劣：
  - “不阻塞 + 连续线性空间”的模拟方式简单且易继续进行线性规划的优化方式，通过加入链接 k 数限制构建空间，达到拓扑优化搜寻；但由于不阻塞反向传输及过量传输，对真实通信模拟往往过量计算通信总数据量。
  - “不阻塞 + 连续线性空间”很方便用于计算已有拓扑的节点间距离、平均距离、直径。
  - “阻塞 + 不连续条件”可以避免不必要的传输以及过量传输的问题，更能模拟现实的通信方式，同样可用于计算不同已有拓扑的相关性质；
  - 但“阻塞 + 不连续条件”目前的模拟函数是不连续的，难以用于拓扑的优化搜寻。
- 并行矩阵乘法的“环”：
  - 在假设 MPI 程序，send : recv = 1 : 1 的前提下，任何矩阵乘法算法中的通信都可拆分为多个“环”（节点最大连接数 $k = 2$）；
  - 并行矩阵乘法算法在拓扑上的应用可被理解为将已有拓扑拆分为多个“环”，并基于此进行优化的过程；

Future

- 计划明日继续优化通信 python 模拟程序，寻找更好的模拟模型；
- 计划在 P = 8 的拓扑上初步尝试“环”的拆分过程。