

多任务外存模式图处理系统中的I/O去重方法

Research on I/O Deduplication in Multi-Task External Memory based Graph Processing System

施展 彭亚妹 冯丹 鲍匡迪 陈俭喜

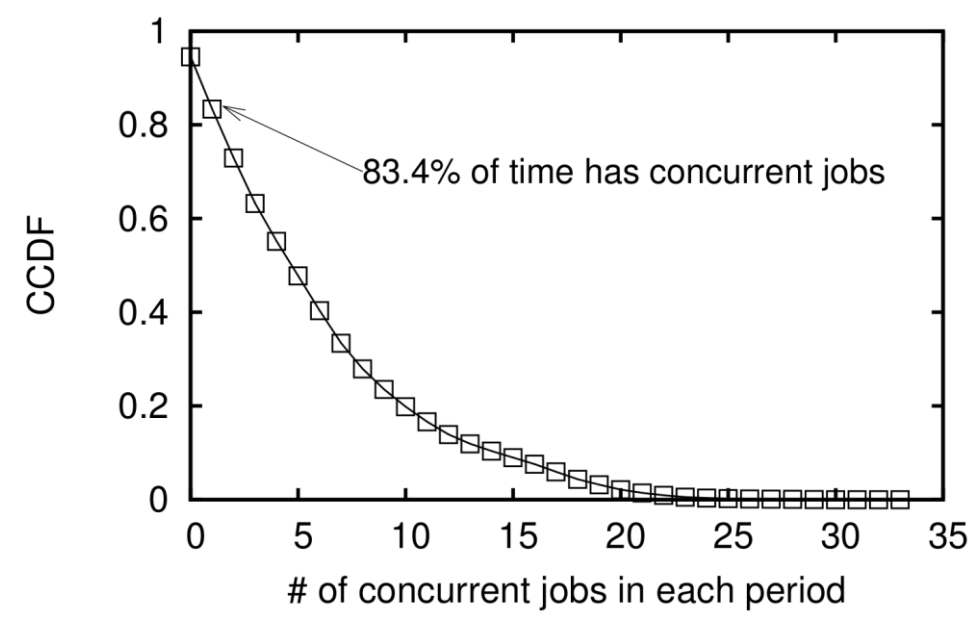
华中科技大学 武汉光电实验室, 武汉 430074

zshi@hust.edu.cn

动机与目标

图处理系统广泛应用于各个领域, 面临并行任务挑战。

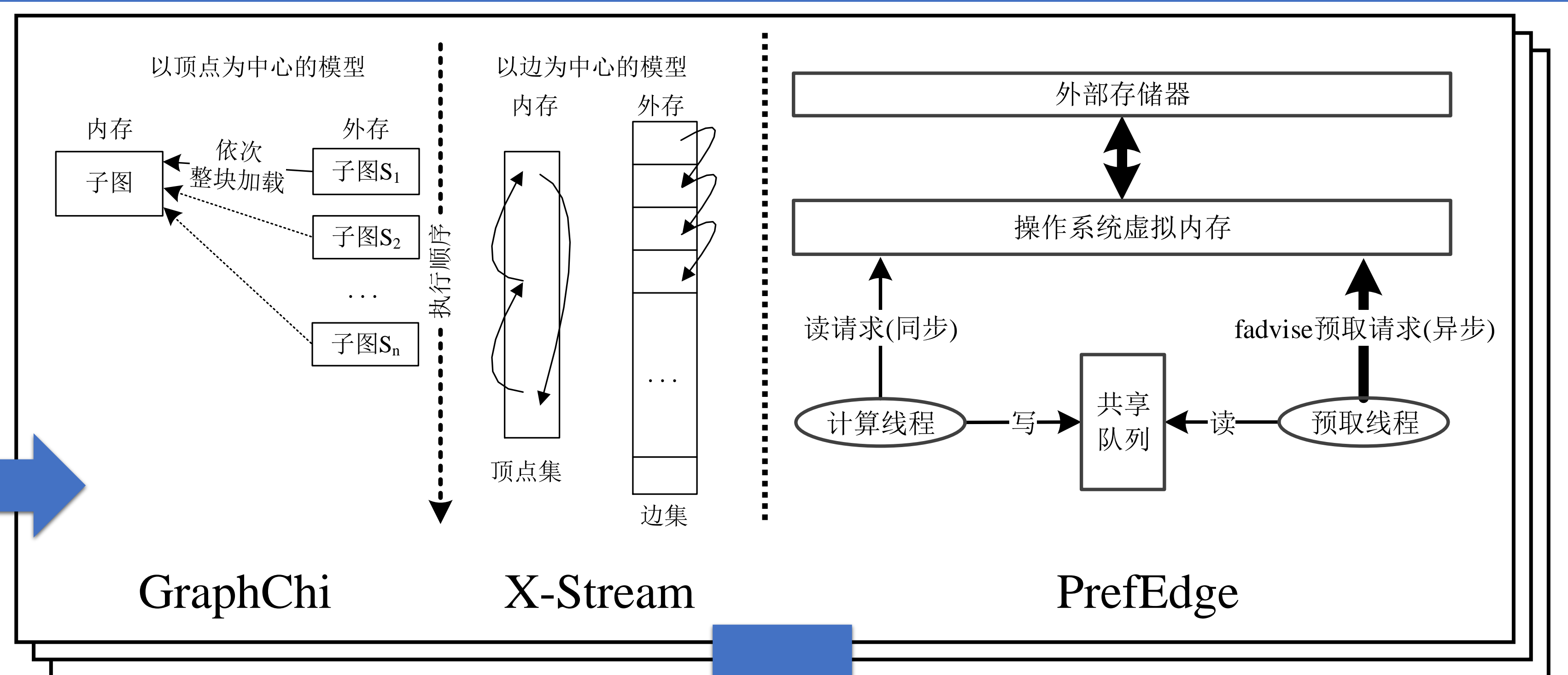
- 社交网络、文献引用、蛋白质交互、.....



	图数据	I/O
内存模式	状态: 内存 结构: 内存	状态更新: 内存+网络 结构访问: 内存
外存模式	状态: 内存+外存 结构: 外存	状态更新: 内存 结构访问: 外存

多任务I/O压力来源

- 外存模式
 - 不需要对图数据分区, 避免了网络通信开销。
 - GraphChi、X-Stream性能表现不弱于使用中小规模集群的分布式内存模式图处理系统。
- 多任务环境中
 - 内存模式的并行任务竞争网络I/O, 状态数据不能共享。
 - 外存模式的并行任务竞争外存I/O, 结构数据可以共享。
- 需要解决的问题: 重复I/O如何同步以一次完成?

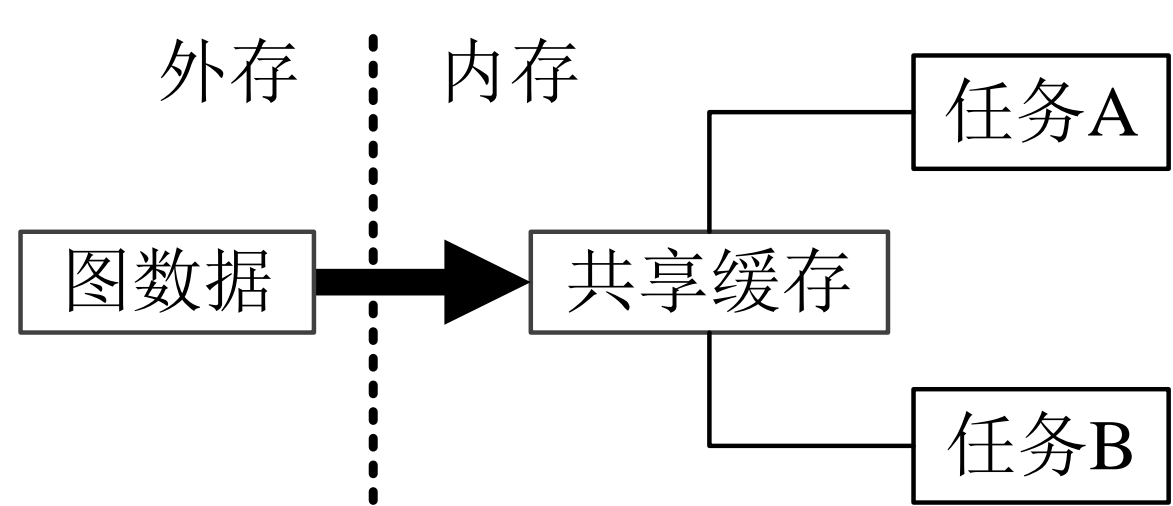
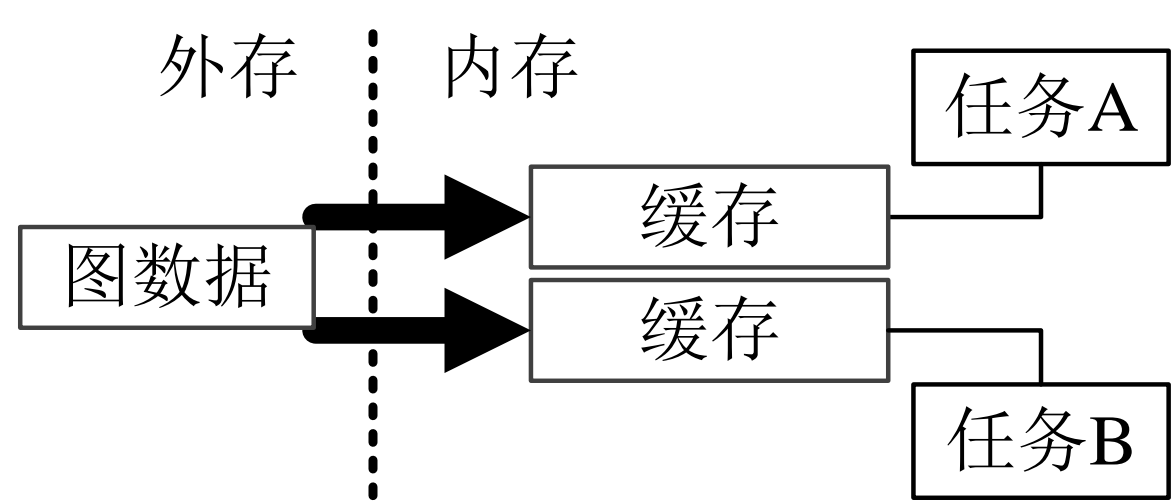


外存模式多任务并行

- 缓存内容重复将浪费内存资源;
- 并行图处理任务造成严重I/O争用;
- 部分外存I/O优化措施包括顺序化改进将被破坏。

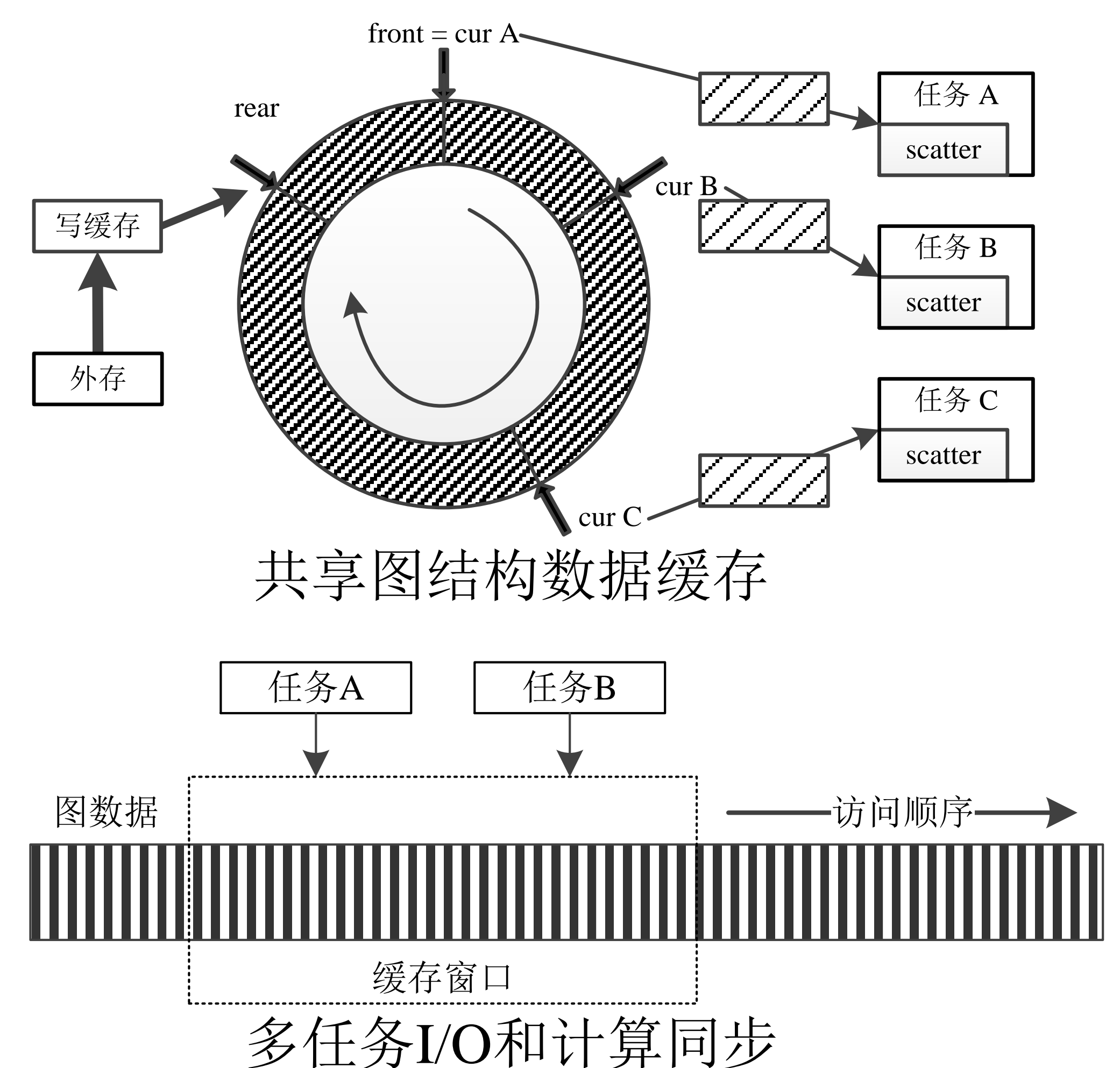
方法和设计

- GraphDeSh Graph Processing System based on I/O Deduplication and Graph Sharing



改变多任务I/O模式

- 改变多任务I/O模式
 - 使用环状共享缓存装载图结构数据。
 - 统一对外存图结构数据的访问, 尽量避免重复I/O执行。
 - 平衡调度并行任务执行, 匹配速度差异, 兼顾I/O减少与并行任务数据请求等待时间优化。
- 图处理任务之间算法的复杂度常常不一致
 - 消费图数据的速度也因此不同。
 - 在统一I/O的基础上, 共享的图结构数据缓存窗口范围内, 允许并行的图处理任务有一定程度速度差异。
 - 在一次迭代过程中, 图结构数据持续从外存加载进内存中, 缓存窗口自左向右移动, 图处理任务可以以各自的速度消费缓存窗口中的图数据。
 - 缓存窗口的宽度是并行任务在迭代进程上差距的极限。



多任务I/O和计算同步

测试与结论

名称	顶点数目	边数目	平均度	描述
Twitter	61,578,415	1,468,365,182	23.8	社交网络Twitter的用户关系
R-MAT25	33,554,432	536,870,912	16	顶点度数呈幂律分布, 差异大
SW25	33,554,432	536,870,912	16	顶点集中, 图的平均直径小
ER25	33,554,432	536,870,912	16	顶点度数呈随机分布, 比较平均

- GraphDeSh的优化效果随着图算法的计算比例下降而下降, 平均加速比为1.45倍。
- 对于不同的测试数据集, GraphDeSh相对于并行GraphChi和X-Stream的加速比最多分别为60.7%和107.1%, 最少分别为32.5%和84.5%。

