

融合分布式计算框架 Ray 实现高通量 XPCS 数据分析

Monday, 25 August 2025 16:40 (20 minutes)

X 射线光子相关谱技术因其在探测纳米尺度动力学行为方面的独特优势，在材料科学、软凝聚态物理和生物物理等领域应用日益广泛。然而，随着同步辐射光源亮度的提升和探测器技术的发展，XPCS 实验产生的海量时间序列数据对后续计算分析，特别是自相关函数的计算，提出了严峻挑战。传统单机或简单并行化方法在处理大规模、高帧频 XPCS 数据时，面临着计算效率低下、扩展性不足和资源利用率不高等瓶颈问题。为应对这一挑战，本研究提出并实现了一种基于分布式计算架构的创新解决方案，其核心是采用 ray 框架将 XPCS 的大量非相关计算进行分布式处理。利用 Ray 的 Actor 模型封装核心计算任务（如 mutiltau 相关算法），通过其分布式任务调度能力，实现了计算任务在集群资源（包括 CPU 和 GPU）上的动态分配与负载均衡。实际应用表明，基于 Ray 构建的分布式 XPCS 处理系统能够线性扩展计算能力，大幅缩短数据处理时间（相比单机处理速度提升数十倍至百倍量级），并显著提升计算资源的整体利用率。

Summary

Primary authors: 刘, 建利; 孙, 浩凯 (高能所); 胡, 誉 (高能所)

Presenter: 刘, 建利

Session Classification: 数据处理软件与分析方法

Track Classification: 数据处理软件与分析方法