前　言

近年来，生成式人工智能、大模型优化与新型计算架构的快速发展，推动高性能计算与智能计算进入协同演进的新阶段. 这一趋势为复杂科学问题的求解提供了更高效的算力范式，然而，算力需求的激增与硬件能效瓶颈的凸显，使得从算法设计、架构优化到系统协同的全链条突破成为关键挑战.

为进一步推动我国学者在高性能计算与智能计算领域的融合研究，及时报道相关最新成果，我们组织了本期“高性能计算与智能计算”专题，旨在聚焦协同理论、架构创新、系统优化及前沿应用等方向. 本专题 2025 年 2 月底征稿结束，经领域专家严格评审与多轮修订，最终收录 16 篇论文，内容涵盖高性能计算系统、智能计算算法、编译调度优化、芯片架构设计等方向的最新研究进展，希望能够为同行学者带来帮助和启发.

在高性能计算系统领域，赵玉龙等人的论文“SW-IntraCC: 一种面向申威智能加速卡内部的集合通信机制”提出三级存储架构与交叉共享通信算法，通过双向算子融合策略实现通信计算重叠，有效提升了在千亿参数大模型训练中的通信效率，为国产超算平台的高效协同提供关键技术支撑. 李彪等人的论文“面向天河新一代超算系统的大规模精确对角化方法”基于无矩阵方法与异构并行算法，可以实现基于一维 Hubbard 模型的大规模精确对角化，为量子多体系统模拟提供新方案. 李仁刚等人的论文“大规模异构一致性融合计算系统的性能建模与优化”设计建模工具 HCSim，结合缓存一致性互连拓扑优化与任务调度策略，显著降低 LLAMA2-13B 大模型训练的通信开销. 李焱等人的论文“Resilio: 一种大模型弹性训练容错系统”通过检查点异步保存与资源弹性调度技术，有效缩短了千亿参数模型的端到端故障恢复时间，提升分布式训练的鲁棒性.

在智能计算领域，李国鹏等人的论文“边缘智能计算系统中加速推荐模型训练的样本调度机制”提出嵌入传输代价建模与混合调度算法（HybridDis），在非独立同分布数据场景下实现端到端 DLRM 训练加速，为边缘计算场景的资源协同管理提供新思路. 冯奕铭等人的论文“异构边缘环境下自适应分层联邦学习协同优化方法”结合深度强化学习与动态模型分割策略，通过数据贡献度量化与资源感知调度，在边缘异构环境中实现训练效率与模型精度的联合优化.

在编译与调度优化方向，徐传福等人的论文“基于缓存数据重用的稀疏矩阵向量乘序列优化”提出缓存感知的动态图划分与子矩阵调度策略，在稀疏矩阵向量乘序列计算中较 Intel MKL 库获得了较大的平均性能提升. 王昊天等人的论文“SparseMode: 用于高效 SpMV 向量化代码生成的稀疏编译框架”通过稀疏模式感知的动态数据布局选择与 SIMD 向量化代码生成，有效减少了冗余操作和内存访问，提升了计算性能. 傅游等人的论文“SW39000 处理器上顺序任务流多从核辅助并行构图算法”针对申威处理器离散访存瓶颈，设计多从核辅助构图与资源调度方法，在细粒度任务场景下超越了主核集中式并发构图. 赖远明等人的论文“SIMD-to-RVV 动态二进制翻译中的跨架构编程模型适配优化”通过消除 vtype 寄存器冗余配置与混合浮点向量化策略，在 RISC-V 平台上实现了 x86 SIMD 指令高效迁移.

在晶圆级集成芯片架构和智能计算芯片架构的创新上，董文阔等人的论文“映天湖：晶圆级通用异构多芯粒千万亿次计算机”，该研究提出标准化I/O设计与动态可重构晶上网络，基于TSMC 28nm工艺实现单晶圆千万亿次算力，为下一代异构计算平台奠定硬件基础. 徐颖等人的论文“Pipe-RLHF: 计算模式感知的RLHF并行加速框架”针对大模型对齐中的PPO算法瓶颈，提出延迟批间流水线并行与异步优化策略，实现训练效率大幅提升. 吕倩茹等人的论文“DAQ: 基于分治策略的自适应Vision Transformer低位宽量化方法”提出z-score驱动的动态分治量化框架，通过量化敏感度-计算-存储开销联合分析与硬件协同设计，实现精度与效率的联合优化，为边缘端视觉模型轻量化提供新范式.

在专用芯片架构领域，石泓博等人的论文“基于数据流架构的NTT蝶式计算加速”面向全同态加密核心算子NTT，设计RISC-V扩展指令集与数据流驱动架构，相比GPU方案取得能效的显著提升，验证了隐私计算场景的硬件加速潜力. 高睿昊等人的论文“BeeZip2: 高性能无损数据压缩领域专用加速器”针对Zstandard算法的大滑动窗口挑战，提出跨层优化的并行哈希表设计与共享处理单元架构，在压缩比不变条件下有效提升了吞吐率. 李鹏程等人的论文“一种可重构的单精度近似浮点乘法器设计”基于移位加法与非对称截断技术，设计低功耗近似浮点乘法器，为边缘计算场景提供了高能效硬件解决方案.

在当今高性能计算与智能计算协同演进的时代，亟需支持跨层协同优化的理论创新与全栈技术突破，以应对算力瓶颈与能效失衡的双重挑战. 希望本专题能够抛砖引玉，促进我国相关研究的进一步发展. 由于专题征集时间和篇幅有限，无法全面覆盖高性能计算与智能计算领域的近期研究进展，不足之处敬请各位学者谅解和批评指正. 衷心感谢广大作者对本专题组织者的信任和支持，感谢审稿专家对于稿件的认真审查以及中肯的意见和建议！特别感谢本刊编委会和编辑部对本专题的支持，感谢编辑部的辛勤工作！

钱德沛 院士（北京航空航天大学）
李肯立 教授（湖南大学）
窦 勇 教授（国防科技大学）
李华伟 研究员（中国科学院计算技术研究所）
刘 鑫 研究员（国家并行计算机工程技术研究中心）

2025年5月