

PR #27815 完整报告

sgl-project/sglang

[AMD] Update test_aiter_allgather_amd.py data types alignment between benchmark aiter and custom all-reduce kernel

合并时间: 2026-06-17 15:48

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/27815>

执行摘要

- 一句话: AMD 测试删除了不支持的整型列
- 推荐动作: 简单 bugfix, 无需精读, 但展示了测试应基于底层能力边界设计的原则。

功能与动机

test_aiter_allgather_amd.py 测试在 AMD CI 中因指定了底层 all-reduce kernel 不支持的整型 dtype 而失败, 这些 dtype 本就不应该被测试覆盖。

实现拆解

1. 修改 test/registered/ops/test_aiter_allgather_amd.py 中的 dtype_names 列表: 删除 7 个整型 dtype (uint64_t, int64_t, uint32_t, int32_t, int16_t, uint8_t, int8_t), 只保留 float32, float16, bfloat16。
2. 同步更新了注释和说明。

关键文件:

- test/registered/ops/test_aiter_allgather_amd.py (模块 AMD AllGather; 类别 test; 类型 test-coverage): 唯一的变更文件, 删除了 7 个不支持的整型 dtype。

关键符号: 未识别

评论区精华

- 暂无高价值评论线程

风险与影响

- 风险: 无风险: 仅测试层面移除不支持的 dtype, 不影响任何 kernel 或模型逻辑。
- 影响: 仅影响 AMD CI 中该测试的通过情况, 消除假阳性失败。
- 风险标记: 测试覆盖调整

关联脉络

- PR #28226 fix allreduce-fusion residual-accuracy test: 同属 AMD CI 修复, PR body 提及该 PR 会处理相关的 allreduce-fusion residual-accuracy test 失败。