

PR #50450 完整报告

vllm-project/vllm

[ROCm][CI] Use larger atol value for INT3 in test_quick_all_reduce.py

合并时间: 2026-07-31 10:41

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/50450>

执行摘要

- 一句话: 修复 AMD CI 上 INT3 quick allreduce 测试容差不足导致的失败
- 推荐动作: 该 PR 值得快速合并, 但建议后续关注测试容差放宽带来的敏感性下降问题, 可考虑补充注释说明理论最大误差的计算依据, 或使用更精确的量化误差模型。

功能与动机

AMD CI 上的分布式测试 ([mi355_2](#)) 持续失败, 原因在于测试使用了 `int3` 量化, 但原本固定的 `atol=2.5` 不足以容纳理论最大误差。根据 PR body 描述, 作者将 `atol` 设置为所用值范围的理论最大误差乘以 2 (因 sum reduction), 以确保测试通过。

实现拆解

1. 新增容差辅助函数: 在 `tests/distributed/test_quick_all_reduce.py` 中新增 `_quickreduce_tolerance(quant_level: QuickReduceRegime) -> dict`, 根据 `QuickReduceRegime.INT3` 分支返回 `{'atol': 5.5, 'rtol': 0.1}`, 否则返回 `{'atol': 2.5, 'rtol': 0.1}`。该函数基于理论分析: INT3 输入 `absmax` 为 23 时, 每个 rank 的误差峰值为 2.75, TP2 求和后放大至 5.5。
2. 修改 `graph_quickreduce` 路径: 将原先硬编码的 `atol=2.5, rtol=0.1` 替换为基于 `fa.qr_quant_level` 的动态容差, 使用 `_quickreduce_tolerance` 获取并解包为 `**tol`。
3. 修改 `eager_quickreduce` 路径: 同样将硬编码容差替换为动态容差, 确保两种执行模式一致。
4. 测试验证: 在 2 GPU 系统上运行指定测试用例, 修复前失败 (最大绝对差 5.0, 超过允许的 2.5), 修复后通过。

关键文件:

- `tests/distributed/test_quick_all_reduce.py` (模块 分布式测试; 类别 test; 类型 test-coverage; 符号 `_quickreduce_tolerance`): 唯一变更文件, 新增容差辅助函数并应用到两个测试路径, 是修复的核心。

关键符号: `_quickreduce_tolerance`

关键源码片段

[tests/distributed/test_quick_all_reduce.py](#)

唯一变更文件，新增容差辅助函数并应用到两个测试路径，是修复的核心。

```
# tests/distributed/test_quick_all_reduce.py

def _quickreduce_tolerance(quant_level: QuickReduceRegime) -> dict:
    """根据量化级别返回合适的容差参数。

    INT3 输入 absmax 为 23 时，单 rank 误差峰值约 2.75，
    经过 TP2 sum reduction 后误差翻倍为 5.5，因此将 atol 设为 5.5。
    其他量化级别沿用原有容差 2.5。
    """
    if quant_level == QuickReduceRegime.INT3:
        return {"atol": 5.5, "rtol": 0.1}
    return {"atol": 2.5, "rtol": 0.1}

# 在 graph_quickreduce 中使用：
tol = _quickreduce_tolerance(fa.qr_quant_level)
torch.testing.assert_close(out1, inp1, **tol)
torch.testing.assert_close(out2, inp2, **tol)

# 在 eager_quickreduce 中使用：
tol = _quickreduce_tolerance(fa.qr_quant_level)
torch.testing.assert_close(out, inp * tp_size, **tol)
```

评论区精华

无实际审核讨论，仅有 Claude 机器人提示来自 fork 分支自动审核被禁用，以及维护者 AndreasKaratzas 的 LGTM 批准。

- 暂无高价值评论线程

风险与影响

- 风险：技术风险较低，仅修改测试容差。但需注意：将 atol 从 2.5 提升至 5.5 可能掩盖真实的数值错误，降低测试敏感性。此外，该改动未覆盖所有可能的数据范围或 TP 规模，若未来输入范围扩大或 TP 增大，容差可能仍不充分。
- 影响：影响范围局限于 AMD CI 的分布式测试，对生产代码无影响。可修复 mi355_2 测试的持续失败，提升 CI 稳定性。团队无需额外适配，测试期望的数值精度有所放宽。
- 风险标记：容差放宽可能掩盖数值误差，未覆盖不同 TP 规模

关联脉络

- PR #49309 [ROCm][CI] Use explicit wvSplitKrc skinny-GEMM test tolerance for bf16 (gfx950): 同样针对 ROCm 测试容差调整，属于同类 CI 稳定性修复。