

PR #29726 完整报告

sgl-project/sglang

[AMD] Rebalance stage-c-large-8-gpu-mi35x partitions to fix 60-min timeout

合并时间: 2026-07-01 23:17

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/29726>

执行摘要

- 一句话: 调整 AMD CI 测试估计时长以修复分区超时
- 推荐动作: 值得快速浏览以理解 CI 分区机制和实践 (基于 `est_time` 的 LPT 分区器)。对于负责 CI 维护的工程师, 这是一个良好的参考案例: 如何通过校准测试估计时间来避免分区超时, 而无需增加硬件资源或放宽超时限制。

功能与动机

`stage-c-test-large-8-gpu-amd-mi35x` 分区 0 在定时 main 分支运行时间歇性触发 60 分钟 `Run test` 步骤超时。PR body 明确指出根因是分区不平衡而非代码回归: 所有重测试均使用相同的占位 `est_time=3600`, 导致 LPT 分区器无法根据实际运行时负载进行均衡。

实现拆解

1. 分析根因: 通过检查 CI 运行日志和测试实际耗时, 发现分区 0 包含最重的三个测试 (`test_kimi_k25_mxfp4_bcg_mi35x` 约 35 分钟、`test_deepseek_r1_mxfp4_8gpu` 约 22 分钟、`test_breakable_cuda_graph` 约 3 分钟), 总计可达 60 分钟以上; 而分区 1 仅约 17 分钟。
2. 调整 `est_time`: 将 5 个测试文件的 `register_amd_ci` 或 `register_cuda_ci` 调用中的 `est_time` 参数从统一的 3600 替换为基于实测的合理值:
 - `test_kimi_k25_mxfp4_bcg_mi35x`: 3600 → 2100
 - `test_deepseek_r1_mxfp4_8gpu`: 3600 → 1300
 - `test_kimi_k25_mxfp4`: 3600 → 700
 - `test_qwen3_coder_next_8gpu`: 3600 → 450
 - `test_breakable_cuda_graph`: 120 → 200
3. 验证效果: 通过多轮 CI 运行中验证, 分区 0 包含 `kimi_k25_mxfp4_bcg` 和 `breakable_cuda_graph`, 总耗时约 34-40 分钟; 分区 1 包含其余三个测试, 总耗时约 29-39 分钟。两个分区均在 60 分钟超时限制以内, 无需调整超时阈值。

关键文件:

- `test/registered/amd/test_deepseek_r1_mxfp4_8gpu.py` (模块 AMD 测试; 类别 test; 类型 test-coverage): 修改 DeepSeek R1 MXFP4 8-GPU 测试的 `est_time` 从 3600 降至 1300, 基于实测运行时间。

- test/registered/amd/test_kimi_k25_mxfp4.py (模块 AMD 测试; 类别 test; 类型 test-coverage) : 修改 Kimi K2.5 MXFP4 测试的 est_time 从 3600 降至 700, 基于实测运行时间。
- test/registered/amd/test_kimi_k25_mxfp4_bcg_mi35x.py (模块 AMD 测试; 类别 test; 类型 test-coverage) : 修改 Kimi K2.5 MXFP4 BCG 测试的 est_time 从 3600 降至 2100, 基于实测运行时间。这是套件中最重的测试 (约 35 分钟), 正确设置其 est_time 是平衡分区的关键。
- test/registered/amd/test_qwen3_coder_next_8gpu.py (模块 AMD 测试; 类别 test; 类型 test-coverage) : 修改 Qwen3 Coder Next 8-GPU 测试的 est_time 从 3600 降至 450, 基于实测运行时间。
- test/registered/cuda_graph/breakable/test_breakable_cuda_graph.py (模块 CUDA 图测试; 类别 test; 类型 test-coverage) : 修改 breakable CUDA graph 测试的 AMD CI est_time 从 120 微调至 200, 以更准确反映其实际运行时间 (约 180 秒)。

关键符号: 未识别

评论区精华

无实质性 review 讨论。HaiShaw 直接批准, 无评论。PR 作者在 Issue 评论中提供了三次 CI 触发验证链接, 确认分区重新平衡后未超时。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。变更仅修改测试注册时的 est_time 参数, 不影响任何生产代码、模型推理逻辑或测试的行为语义。错误的 est_time 值 (例如设置过低) 可能导致未来新增测试时分区再次不平衡, 但当前值基于多次实测的上限设置, 留有足够余量。不涉及回归、性能、安全或兼容性风险。
- 影响: 影响范围仅限于 AMD MI35x 8-GPU CI 测试套件的分区调度。对用户无影响; 对系统表现为 CI 流水线取消间歇性超时失败, 提高 CI 稳定性; 对团队而言减少了因超时导致的无效重跑和排查时间。
- 风险标记: 测试配置变更, 不影响生产代码

关联脉络

- PR #27833 [Kimi/K2.5/MXFP4/BCG/MI35x] Add BCG test for MI35x: 该 PR 新增了测试 test_kimi_k25_mxfp4_bcg_mi35x.py, 其 est_time=3600 是导致分区超时的直接原因。当前 PR 将其 est_time 调整为 2100。