

PR #21554 完整报告

sgl-project/sglang

[CI] Remove more redundant PCG tests

合并时间: 2026-04-01 07:25

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/21554>

PR 分析报告: 移除冗余 PCG 测试

执行摘要

本 PR 移除了三个针对 GPT-OSS、Kimi-Linear 和 Qwen3 Next 模型的 Piecewise CUDA Graph (PCG) 专用测试文件, 因为 PCG 功能现已默认启用, 这些测试变得冗余。变更简化了 CI 测试套件, 减少维护负担, 对用户无直接影响, 风险较低。

功能与动机

动机源于 PCG 已默认启用, 使得专门测试 PCG 的用例不再必要。PR 描述指出: “Similar to #21485, more PCG tests that are redundant now that PCG is enabled by default.” 在关联 Issue 的评论中, Fridge003 询问测试覆盖情况, 作者 mmangkad 确认除 Qwen3 Next 外, 其他模型已测试 PCG 默认启用, Qwen3 Next 在配置中默认禁用 PCG。

实现拆解

实现非常简单, 直接删除了以下三个测试文件:

文件路径	关键内容	状态
<code>test/registered/models/test_gpt_oss_models_pcg.py</code>	包含 <code>TestGptOssPiecewiseCudaGraph</code> 类, 测试 GPT-OSS 模型的 PCG 功能。	已删除
<code>test/registered/models/test_kimi_linear_models_pcg.py</code>	包含 <code>TestKimiLinearPiecewiseCudaGraph</code> 类, 测试 Kimi-Linear 模型的 PCG 功能。	已删除
<code>test/registered/models/test_qwen3_next_models_pcg.py</code>	包含 <code>TestQwen3NextPiecewiseCudaGraph</code> 类, 测试 Qwen3 Next 模型的 PCG 功能。	已删除

这些文件都使用 unittest 框架, 并注册到 CI 套件 (如 `stage-b-test-2-gpu-large`), 现在由于 PCG 默认启用, 其逻辑可能已集成到其他通用测试中。

评论区精华

主要讨论发生在关联 Issue 的评论中:

Fridge003: “Have you tested that these models will enable PCG by default?”
mmangkad: “Yes except Qwen3 Next, not sure why it has a test since it is disabled by default in [配置文件]”

这解决了测试覆盖的疑虑，确保移除不会影响核心功能验证。PR 的 review 中，bot 无反馈，Fridge003 直接批准，表明变更得到认可。

风险与影响

- 风险分析：测试覆盖可能略微减少，但由于 PCG 默认启用，专用测试的移除不影响功能正确性。对于 Qwen3 Next 模型，需注意配置默认禁用 PCG，但其他测试应覆盖相关逻辑。无回归风险。
- 影响分析：对用户无影响；CI 测试执行时间可能减少；团队开发效率提升，代码库更简洁；影响程度轻微，属于常规清理。

关联脉络

- 相关 PR：本 PR 直接引用 #21485（移除冗余 PCG 测试），同为测试清理工作；另可参考 #22058（删除回滚测试），反映仓库中测试基础设施的持续优化趋势。
- 演进方向：这表明 PCG 功能已成熟并默认启用，测试策略从专项验证转向集成验证，有助于简化 CI 流程并聚焦核心测试用例。