

PR #25032 完整报告

sgl-project/sglang

[CI] Temporarily disable multimodal-gen test_update_weights_from_disk (flaky)

合并时间: 2026-05-13 03:28

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/25032>

执行摘要

- 一句话: 临时禁用 flaky 测试 test_update_weights_from_disk
- 推荐动作: 该 PR 为临时缓解措施, 技术含量低, 但管理上重要。可快速合并; 后续应单独开 Issue 跟踪根因修复, 避免测试永久禁用。

功能与动机

CI 中的 multimodal-gen-test-1-gpu 套件因 test_update_weights_from_disk.py 频繁超时 (24 次), 导致每次提交的 CI 信号被阻塞, 影响了其他无关的多模态生成变更的合并。PR body 明确指出“blocking per-commit signal on unrelated multimodal-gen changes”。

实现拆解

1. 在 python/sglang/multimodal_gen/test/run_suite.py 的 STANDALONE_FILES["1-gpu"] 列表中, 将 "test_update_weights_from_disk.py" 注释掉, 并添加注释说明原因。
2. 在对应的 STANDALONE_FILE_EST_TIMES["1-gpu"] 字典中, 将 "test_update_weights_from_disk.py": 480.0 注释掉, 确保 CI 不会为该测试分配时间预算。
3. 测试文件本身未被删除, 后续只需取消两处注释即可重新启用。

关键文件:

- python/sglang/multimodal_gen/test/run_suite.py (模块 测试套件; 类别 test; 类型 test-coverage; 符号 STANDALONE_FILES, STANDALONE_FILE_EST_TIMES): 唯一变更文件: 注释掉 flaky 测试的引用和时间估算, 是 PR 的核心改动。

关键符号: 未识别

关键源码片段

[python/sglang/multimodal_gen/test/run_suite.py](#)

唯一变更文件: 注释掉 flaky 测试的引用和时间估算, 是 PR 的核心改动。

```
# python/sglang/multimodal_gen/test/run_suite.py (部分)
```

```
STANDALONE_FILES = {  
    "1-gpu": [  
        "../cli/test_generate_t2i_perf.py",
```

```
# Temporarily disabled: 24 timeout failures since 2026-04-09 across
# multimodal-gen-test-1-gpu. Re-enable after the flakiness is fixed.
# "test_update_weights_from_disk.py",
],
"2-gpu": [
    "test_disagg_server.py",
],
}

STANDALONE_FILE_EST_TIMES = {
    "1-gpu": {
        "../cli/test_generate_t2i_perf.py": 240.0,
        # See STANDALONE_FILES note above — temporarily disabled.
        # "test_update_weights_from_disk.py": 480.0,
    },
    "2-gpu": {
        "test_disagg_server.py": 600.0,
    },
}
```

评论区精华

该 PR 由 hnyls2002 直接批准，无 review 评论或讨论。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅注释掉一条测试引用和时间估算，不影响其他测试和功能代码。若 root cause 长时间未修复，可能会导致该测试的回归缺失；但测试文件仍保留，且 PR 明确预期后续修复后重新启用。
- 影响：对用户无影响。对系统而言，multimodal-gen-test-1-gpu CI 套件将不再执行此测试，CI 信号更可靠，不再被 flaky 测试阻塞。团队后续需要修复超时 root cause 并重新启用。
- 风险标记：极少风险

关联脉络

- PR #14021 Disable flaky test_update_weights_from_disk (RL variant): PR body 提到相同名称的 RL 测试已在 #14021 中独立禁用，两者不互相影响，但均属于同一类 flaky 测试处理。