

PR #27803 完整报告

sgl-project/sglang

[CI] Fix stale ngram bookkeeping owner sites

合并时间: 2026-06-10 19:51

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/27803>

执行摘要

- 一句话: 修复 ngram 书签所有权失效测试
- 推荐动作: 该 PR 改动简单但有必要, 值得合并以维护测试可靠性。

功能与动机

PR #17260 删除了 ngram 推测解码的旧路径 `NgramVerifyInput._fill_requests` 和 `_free_cache`, 但测试文件中的书签所有权记录未同步删除, 导致 PR #27710 新增的所有权测试失败。具体见 CI 失败日志: <https://github.com/sgl-project/sglang/actions/runs/27271813037/job/80542982353>

实现拆解

1. 定位问题: 在 `test/registered/unit/spec/test_decode_bookkeeping_ownership.py` 中, `_OWNER_SITES` 字典包含了已废弃的 ngram 书签所有权条目, 这些条目引用了不存在的函数。
2. 清理记录: 删除了 3 条过时的所有权记录:
 - (`speculative/ngram_info.py`, `NgramVerifyInput._fill_requests`, `spec_verify_ct`): 1
 - (`speculative/ngram_info.py`, `NgramVerifyInput._free_cache`, `kv_committed_len`): 1
 - (`speculative/ngram_info.py`, `NgramVerifyInput._free_cache`, `kv_allocated_len`): 1
3. 验证: 通过 `/rerun-test test/registered/unit/spec/test_decode_bookkeeping_ownership.py` 触发 CI 重跑, 测试通过。

关键文件:

- `test/registered/unit/spec/test_decode_bookkeeping_ownership.py` (模块 书签所有权; 类别 test; 类型 test-coverage) : 删除了 3 条过时的 ngram 书签所有权记录, 修复 CI 测试失败。

关键符号: 未识别

关键源码片段

`test/registered/unit/spec/test_decode_bookkeeping_ownership.py`

删除了 3 条过时的 ngram 书签所有权记录, 修复 CI 测试失败。

```
# 删除前: 包含 NgramVerifyInput._fill_requests 和 _free_cache 的过时记录
# 删除后: 只保留有效的所有权站点
_OWNER_SITES = {
    # ... 其他有效条目 ...

    # spec v1: each verify path owns its own settlement
    ("speculative/eagle_info.py", "EagleVerifyInput.verify", "kv_committed_len"): 1,
    ("speculative/eagle_info.py", "EagleVerifyInput.verify", "kv_allocated_len"): 1,
    ("speculative/eagle_info.py", "EagleVerifyInput.verify", "spec_verify_ct"): 1,
    # 以下三条已被 PR #17260 删除的路径, 故移除
    # ("speculative/ngram_info.py", "NgramVerifyInput._fill_requests", "spec_verify_ct"): 1,
    # ("speculative/ngram_info.py", "NgramVerifyInput._free_cache", "kv_committed_len"): 1,
    # ("speculative/ngram_info.py", "NgramVerifyInput._free_cache", "kv_allocated_len"): 1,
    ("speculative/dflash_info.py", "DFlashVerifyInput.verify", "kv_committed_len"): 1,
    ("speculative/dflash_info.py", "DFlashVerifyInput.verify", "kv_allocated_len"): 1,
    ("speculative/dflash_info.py", "DFlashVerifyInput.verify", "spec_verify_ct"): 1,
    # ... 其他有效条目 ...
}
```

评论区精华

无实质性讨论, 仅由作者触发 CI 重跑和标签操作。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。仅删除测试文件中的过时记录, 不影响生产代码。测试框架在该文件新增时会扫描所有书签站点, 删除的记录已无对应源码, 不会引发回归。
- 影响: 影响范围仅限于测试文件, 修复了 ngram 所有权测试的 CI 失败, 使后续 PR 能正常通过该测试。
- 风险标记: 暂无

关联脉络

- PR #17260 [Feature] [Ngram spec] Support ngram spec v2: 该 PR 删除了旧 ngram 路径, 导致本 PR 清理测试中的过时记录。
- PR #27710 未知: 该 PR 新增了本书签所有权测试, 本 PR 修复了其 CI 失败。