

PR #37108 完整报告

sgl-project/sglang

[mem_cache] Share one `ReqKvInfo` between a streaming session slot and its request

合并时间: 2026-08-31 07:36

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/37108>

执行摘要

- 一句话:
- 推荐动作:

功能与动机

实现拆解

1. 核心源码路径 - `python/sglang/srt/session/streaming_session.py` (core-logic) : 源码主路径; 涉及符号 `_free_tail`; 包含 导入关系调整、控制流调整、配置键调整; +18/-28; 关键符号: `_free_tail` - `python/sglang/srt/managers/schedule_batch.py` (core-logic) : 源码主路径; 涉及符号 `detach_kv`; 包含 符号定义调整; +5/-0; 关键符号: `detach_kv`
2. 测试与验证 - `test/registered/unit/mem_cache/test_streaming_session_unit.py` (test-coverage) : 测试配套; 涉及符号 `detach_kv`; 包含 测试覆盖调整、配置键调整、符号定义调整; +8/-14; 关键符号: `detach_kv` - `test/registered/unit/spec/test_decode_bookkeeping_ownership.py` (test-coverage) : 测试配套; 包含 测试覆盖调整; +4/-5

关键文件:

- `python/sglang/srt/session/streaming_session.py` (模块 `sglang/srt`; 类别 `source`; 类型 `core-logic`; 符号 `_free_tail`) : 源码主路径; 涉及符号 `_free_tail`; 包含 导入关系调整、控制流调整、配置键调整; +18/-28
- `python/sglang/srt/managers/schedule_batch.py` (模块 `sglang/srt`; 类别 `source`; 类型 `core-logic`; 符号 `detach_kv`) : 源码主路径; 涉及符号 `detach_kv`; 包含 符号定义调整; +5/-0
- `test/registered/unit/mem_cache/test_streaming_session_unit.py` (模块 `streaming/session/unit`; 类别 `test`; 类型 `test-coverage`; 符号 `detach_kv`) : 测试配套; 涉及符号 `detach_kv`; 包含 测试覆盖调整、配置键调整、符号定义调整; +8/-14
- `test/registered/unit/spec/test_decode_bookkeeping_ownership.py` (模块 `decode/bookkeeping/ownership`; 类别 `test`; 类型 `test-coverage`) : 测试配套; 包含 测试覆盖调整; +4/-5

关键符号: `_free_tail`, `detach_kv`

评论区精华

- 暂无高价值评论线程

风险与影响

- 风险:
- 影响:
 - 风险标记: 暂无

关联脉络

- 暂无明显关联 PR