

PR #25639 完整报告

sgl-project/sclang

Delete the now-unused is_work_request from scheduler.py

合并时间: 2026-05-18 18:50

原文链接: <http://prhub.com.cn/sgl-project/sclang/pull/25639>

执行摘要

- 一句话: 删除 scheduler.py 中无用的 is_work_request 函数
- 推荐动作: 值得快速合并的清理 PR。作为重构链的一环, 建议确认前序 free-item relocation PR 已合入, 以确保上下文一致。无需精读。

功能与动机

PR body 说明: codebase-wide grep shows zero callers (dead code from a now-defunct dispatch path)。作为前序 free-item relocation 重构的尾随清理, 移除无用代码以降低维护成本。

实现拆解

1. 在 python/sclang/srt/managers/scheduler.py 中删除函数 is_work_request (约 3710-3719 行), 该函数原用于判断请求是否为工作请求 (TokenizedGenerateReqInput 等类型)。删除后, 相邻函数 dispatch_event_loop 的缩进自动调整, 无需额外修改。

关键文件:

- python/sclang/srt/managers/scheduler.py (模块 调度器; 类别 source; 类型 core-logic ; 符号 is_work_request) : 删除 is_work_request 函数的唯一变更文件, 清理死代码。

关键符号: is_work_request

关键源码片段

[python/sclang/srt/managers/scheduler.py](#)

删除 is_work_request 函数的唯一变更文件, 清理死代码。

```
# 删除前: 位于 Scheduler 类后的模块级函数, 用于判断请求类型
# def is_work_request(recv_req):
# return isinstance(
# recv_req,
# (
# TokenizedGenerateReqInput,
# TokenizedEmbeddingReqInput,
# BatchTokenizedGenerateReqInput,
# BatchTokenizedEmbeddingReqInput,
```

```
# ),  
# )
```

删除后：仅保留 `dispatch_event_loop` 等活跃代码，函数定义已移除。

```
def dispatch_event_loop(scheduler: Scheduler):  
    # 根据调度模式分发事件循环  
    server_args = scheduler.server_args  
    disaggregation_mode: DisaggregationMode = scheduler.disaggregation_mode  
    if disaggregation_mode == DisaggregationMode.NULL:  
        if scheduler.enable_pdmux:  
            scheduler.event_loop_pdmux()  
        elif server_args.pp_size > 1:  
            scheduler.event_loop_pp()  
    # ... 后续分支不变
```

评论区精华

无 review 讨论。PR 仅由作者自合并，机器人提示达到日配额限制，未参与有效讨论。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低。纯删除死代码，对现有功能无影响。但未提供代码库全局搜索的正式证据（如引用 `grep` 结果或测试），理论上存在疏忽遗漏调用方的可能性。不过变更已通过 CI（PR Test 运行中），且函数注释标记为 `placeholder`，可靠性较高。
- 影响：无影响。删除的 `is_work_request` 无人调用，因此对用户、系统性能、接口无任何可见影响。仅减少 `scheduler.py` 约 12 行代码，轻微提升可读性。
- 风险标记：无风险

关联脉络

- 暂无明显关联 PR