

# PR #26614 完整报告

sgl-project/sglang

[CI] /rerun-test: descend into directories that a glob matched

合并时间: 2026-05-29 07:52

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/26614>

## 执行摘要

- 一句话: 修复 /rerun-test 的目录递归与多 pool 调度问题
- 推荐动作: 建议 CI 维护者和经常使用 /rerun-test 的开发者阅读此 PR, 了解 glob 递归处理和多 pool 扇出的设计; 尤其是 `_extract_runner_configs` 返回列表这一变更, 可能影响依赖该函数的其他调用方。

## 功能与动机

作为 #26422 的后续, `/rerun-test` 在实际使用中暴露了 glob 不递归和多注册池仅第一个生效的问题。具体示例: `/rerun-test registered/attention/unittest/*` 返回无匹配, 因为 glob 的 `*` 不跨越 `/`; 文件若同时注册了 `4-gpu-b200` 和 `1-gpu-large`, `_extract_runner_config` 仅返回第一个 pool, 导致所有测试只调度到第一个 pool。

## 实现拆解

1. 在 `expand_glob_spec` 函数中, 添加目录递归步骤: 当 glob 匹配到目录时, 递归搜索 `test_*.py`。
2. 将 `_extract_runner_config` 重写为 `_extract_runner_configs`, 使用 `re.finditer` 返回所有 `register_cuda_ci` 调用的 `runner_config` 和参数字符串。
3. 将 `detect_suite` 和 `_resolve_test_spec` 改为返回列表, 使一个文件可以对应多个 `dispatch` 项。
4. 在 `handle_rerun_test` 中, 迭代 `dispatch` 列表, 基于 `(mode, runs_on, test_command)` 去重, 确保每个 pool 调度一次但不丢失多 pool 扇出。
5. 在 `find_workflow_run_url` 中, 当 `display_title` 以 `...` 结尾时, 使用前缀匹配, 避免因 GitHub 截断而找不到运行 URL。

关键文件:

- `scripts/ci/utils/slash_command_handler.py` (模块 CI 脚本; 类别 `infra`; 类型 `infrastructure`; 符号 `_extract_runner_config`, `_extract_runner_configs`, `expand_glob_spec`, `detect_suite`): 该文件是整个 `/rerun-test` 命令处理的核心, 所有修复都集中于此。

关键符号: `_extract_runner_configs`, `expand_glob_spec`, `detect_suite`, `_resolve_test_spec`, `handle_rerun_test`, `find_workflow_run_url`

## 评论区精华

维护者 ch-wan 批准 PR, 并引用 #26517 中的实际运行结果确认修复生效: 'It works: <https://github.com/sgl-project/sglang/pull/26517#issuecomment-4569234604>'。

- 暂无高价值评论线程

## 风险与影响

- 风险: 1) 递归目录可能意外匹配深层目录的非预期测试文件, 但限定为 `test_*.py` 模式, 风险低。 2) 多 pool 扇出可能导致同一测试在多个 pool 上运行, 增加 CI 负载, 但这是预期行为且过去应该如此。 3) `display_title` 前缀匹配可能错误匹配不相关的运行, 但由于命令长度足够且前缀唯一, 碰撞概率极低。 4) 对现有使用明确 glob 模式 (如 `**/test_*.py`) 的用户无影响。
  - 影响: 影响范围: 所有使用 `/rerun-test` 命令的开发者。改进后, 命令能正确匹配嵌套子目录中的测试, 并自动在测试文件注册的所有 runner pool 上调度。这提高了 CI 重运行准确性和效率, 减少了人工 workaround。
  - 风险标记: 全局匹配行为变更, 调度重复风险, `display_title` 前缀匹配准确度

## 关联脉络

- PR #26422 原始 `/rerun-test` 实现: 本 PR 是基于 #26422 的后续修复。
- PR #26517 问题示例 PR: 暴露了当前 PR 所修复的 glob 不递归问题。