

# PR #23305 完整报告

sgl-project/sglang

[misc] CI hygiene: enforce `__main__` entry, drop silent-skipped tests, fix rerun-test protoc

合并时间: 2026-04-21 12:16

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23305>

## 执行摘要

- 一句话: CI 卫生: 强制 `__main__` 入口, 删除静默跳过测试, 修复 rerun-test
- 推荐动作: 该 PR 值得精读, 尤其关注 `ci_register.py` 中 AST 验证的实现模式, 这是一种轻量级且有效的 CI 防护手段。设计决策: 使用静态分析而非运行时检查, 避免了性能开销。同时, 删除测试文件时保留原作者信息和恢复路径, 体现了良好的工程实践。

## 功能与动机

CI runner 通过 `python3 file.py -f` 调用测试文件, 但 `pytest` 风格文件若缺少 `if __name__ == "__main__":` 块, 则仅 `import` 后退出 0, 测试函数从未执行, CI 却显示绿色。此问题容易在 review 中漏掉, 导致测试覆盖率虚高。PR body 明确指出了根因和影响。

## 实现拆解

1. AST 健全性检查: 在 `python/sglang/test/ci/ci_register.py` 中添加 `_is_main_block_with_call` 函数, 使用 AST 解析每个注册文件, 检查是否存在调用 `unittest.main()` 或 `pytest.main()` 的可执行 `__main__` 块。 `collect_tests` 增加 `sanity_check=True` 参数, 在收集测试时验证所有非 `disabled` 文件, 未通过则报错。
2. 修复静默跳过文件: 为 6 个文件添加 `__main__` 块和必要的 `fixture` 修复 (如 `test_streaming_session_unit.py` 中 `_FakeReq.finished_len = None`), 使其在 CI 中实际执行。同时从 `test_dump_comparator.py` 中删除依赖 GPU 的 `TestMainBasic` 子集。
3. 删除不可修复的跳过文件: 删除 9 个因缺少 `__main__`、依赖特定硬件或性能断言而长期静默跳过的测试文件, 包括 `test_bnb.py`、`test_int4fp8_moe.py`、`test_zimage_turbo.py`、`python/sglang/jit_kernel/tests/test_dependency.py`、`test_repeat_interleave.py`、`test_glm47_moe_detector.py`、`test_model_hooks.py`、`test_mrope.py`、`test_patch_embed_perf.py`。PR body 标注了原始作者, 欢迎后续修复后重新添加。
4. 修复 rerun-test CPU job: 在 `.github/workflows/rerun-test.yml` 中添加 `Install protoc` 和 `Install Rust toolchain` 步骤, 与 `pr-test.yml` 保持一致, 避免 `sglang-grpc` 编译时因缺少 `protoc` 而失败。

关键文件:

- `python/sglang/test/ci/ci_register.py` (模块 测试注册; 类别 `test`; 类型 `test-coverage`; 符号 `_is_main_block_with_call`, `ut_parse_one_file`): 核心变更: 添加 AST 健全性检查函数 `_is_main_block_with_call`, 在 `collect_tests` 中验证每个注册文件的 `__main__` 块存在性

, 是修复静默跳过漏洞的关键。

- test/registered/unit/mem\_cache/test\_streaming\_session\_unit.py (模块 流式会话测试; 类别 test; 类型 test-coverage; 符号 test\_streaming\_release\_kv\_cache\_defers\_tail\_free) : 直接修复目标: 添加 `__main__` 块并修复 `_FakeReq.finished_len = None` (与实际 Req 默认值一致), 使 4 个测试函数在 CI 中实际运行。同时删除已被 #22897 取代的 `defers_tail_free` 测试。
- .github/workflows/rerun-test.yml (模块 CI 工作流; 类别 infra; 类型 configuration) : 修复 CPU 重跑任务: 添加 `protoc` 和 `rustup` 安装步骤, 与 `pr-test.yml` 保持一致, 避免 `sglang-grpc` 编译失败。

关键符号: `_is_main_block_with_call`, `ut_parse_one_file`,  
`test_streaming_release_kv_cache_defers_tail_free`

## 关键源码片段

### python/sglang/test/ci/ci\_register.py

核心变更: 添加 AST 健全性检查函数 `_is_main_block_with_call`, 在 `collect_tests` 中验证每个注册文件的 `__main__` 块存在性, 是修复静默跳过漏洞的关键。

```
def _is_main_block_with_call(filepath: str) -> bool:
    """检查文件是否包含调用unittest.main()或pytest.main()的可执行__main__块。"""
    with open(filepath) as f:
        tree = ast.parse(f.read())
    # 遍历顶层节点, 寻找 if __name__ == "__main__": 块
    for node in ast.walk(tree):
        if isinstance(node, ast.If) and _is_name_main(node):
            # 在 if 块内查找函数调用
            for child in ast.walk(node):
                if isinstance(child, ast.Call):
                    func_name = _get_func_name(child)
                    if func_name in {"unittest.main", "pytest.main"}:
                        return True
    return False

# 在 collect_tests 中调用
if sanity_check and registry_files:
    for filepath in registry_files:
        if not _is_main_block_with_call(filepath):
            raise RuntimeError(f"{filepath} 缺少可执行的__main__块, 将导致CI静默跳过")
```

## 评论区精华

fxmarty-amd: "Why are the tests removed? Does it not mean that some part of sglang codebase still exist, but are simply untested?" 作者回复: 这些测试文件从未在 CI 中实际运行, 删除是临时措施, 后续 PR 会恢复并正确注册。团队已知此风险, 并计划跟进修复。

- 删除测试文件是否导致功能未经测试 (question): 作者确认这些测试从未在 CI 中实际运行, 删除是临时措施, 后续 PR 会恢复并正确注册。团队已知风险。

## 风险与影响

- 风险: 测试覆盖缺失: 删除的 9 个测试文件对应功能 (如 BNB 量化、MROPE、模型钩子等) 暂时失去 CI 覆盖, 直到后续 PR 恢复。但原始代码从未在 CI 中执行, 因此实际无回归风险。CI workflow 修复: rerun-test.yml 的 protoc 修复确保 CPU 重跑任务不再因编译失败而阻塞, 但若依赖版本变化可能仍需调整。
- 影响: 对 CI 系统: 显著提升测试覆盖率真实性, 消除静默跳过漏洞。未来所有注册测试文件必须包含可执行 `__main__` 块, 否则 CI 会报错。对开发者: 需确保新增测试文件遵循该规范, review 时注意检查。对用户: 无直接影响。对团队: 降低了测试维护成本, 但需在后续 PR 中恢复已删除测试文件。
- 风险标记: 测试覆盖暂时缺失, CI workflow 修复

## 关联脉络

- PR #22897 streaming session: trim spec v2 overshoot in cache\_finished\_req: 该 PR 被删除的 defers\_tail\_free 测试意图已被此 PR supersede, 因此删除是合理的。