

PR #2107 完整报告

THUDM/slime

Fix wrongly removed RouterArgs registry

合并时间: 2026-06-20 22:58

原文链接: <http://prhub.com.cn/THUDM/slime/pull/2107>

执行摘要

- 一句话: 修复 SGLang RouterArgs 注册遗漏
- 推荐动作: 变更微小且修复明确, 可快速合入。建议开发者检查 `sglang_router` 库版本是否为预期版本 (参考 Dockerfile 或 `sglang` 依赖), 避免因库接口变化引入新的不兼容。

功能与动机

SGLang 升级后, 重构过程中误删了 `RouterArgs.add_cli_args` 调用, 导致 SGLang 路由器的标准参数无法通过 `--sglang-router-*` 前缀自动注册, 破坏了原有配置接口的一致性。PR body 未详细说明, 但从提交信息「Fix wrongly removed RouterArgs registry」和仅 2 行新增的变更内容可明确判断这是一个回归修复。

实现拆解

1. 在函数 `add_sglang_router_arguments` 末尾新增一行 `RouterArgs.add_cli_args(parser, use_router_prefix=True, exclude_host_port=True)`, 调用 `sglang_router` 库提供的 CLI 参数注册方法, 将所有 RouterArgs 支持的参数以 `--sglang-router-` 前缀自动注入 `parser`。
2. 传参 `use_router_prefix=True` 确保参数均带有 `--sglang-router-` 前缀, 避免与全局参数冲突; `exclude_host_port=True` 排除 `host`、`port` 等已在别处管理的参数, 保持功能与 SGLang 官方用法一致。
3. 无测试、配置或部署配套改动。

关键文件:

- `slime/backends/sglang_utils/arguments.py` (模块 SGLang 参数; 类别 source; 类型 dependency-wiring): 唯一变更文件。恢复被误删的 `RouterArgs.add_cli_args` 调用, 修复 SGLang 路由器参数注册遗漏。

关键符号: `add_sglang_router_arguments`

关键源码片段

`slime/backends/sglang_utils/arguments.py`

唯一变更文件。恢复被误删的 `RouterArgs.add_cli_args` 调用, 修复 SGLang 路由器参数注册遗漏。

```
import argparse
```

```
from sglang.srt.server_args import ServerArgs
from sglang_router.launch_router import RouterArgs # 补回 RouterArgs 导入
from slime.utils.http_utils import _wrap_ipv6
```

```
def add_sglang_router_arguments(parser):
    parser.add_argument("--sglang-router-ip", type=str, default=None)
    parser.add_argument("--sglang-router-port", type=int, default=None)
    parser.add_argument("--sglang-router-request-timeout-secs", type=int, default=14400)
    # 补回该调用：自动注册所有 RouterArgs 支持的参数，均带 --sglang-router- 前缀
    # exclude_host_port=True 避免与手动管理的 host/port 冲突
    RouterArgs.add_cli_args(parser, use_router_prefix=True, exclude_host_port=True)
    return parser
```

评论区精华

该 PR 无 review 评论。

- 暂无高价值评论线程

风险与影响

- 风险：变更极微（仅 2 行新增），风险极低。但需要注意 RouterArgs.add_cli_args 的接口是否随 sglang_router 库升级发生不兼容变化（例如参数名、前缀行为），建议确保 sglang_router 版本与本次回归的上下文一致。
- 影响：仅影响 SGLang 路由器参数解析流程，影响范围局限于使用 SGLang 路由器的训练 / 推理任务。对不使用 SGLang Router 功能的用户无影响。
- 风险标记：缺少测试覆盖

关联脉络

- PR #2081 sync from internal and cleanup: 可能引入了误删 RouterArgs.add_cli_args 的变更（属于同一清理 / 重构系列）。
- PR #2072 [docker] upgrade sglang to v0.5.13: SGLang 升级后，RouterArgs 接口可能发生变化，导致原有调用在清理过程中被误删。