

PR #27232 完整报告

sgl-project/sglang

[AMD][CI] Remove transformers pin from GLM-5.x nightly jobs

合并时间: 2026-06-04 14:34

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/27232>

执行摘要

移除 AMD 夜间 CI 中 GLM-5.x 作业的 transformers 版本锁定（固定到 commit 96f807a33b75），使这些作业使用镜像默认的 transformers，消除因旧版本导致的导入错误。

功能与动机

GLM-5.x 的 AMD 夜间测试是唯一将 transformers 固定到旧开发版的作业。由于旧版本缺少 `@strict` 装饰器，导致 `Cohere2MoeConfig` 在导入时崩溃（PR #26106）。通过移除固定，统一依赖版本，并验证 GLM-5.1 在默认 transformers 上是否正常运行。

实现拆解

1. 修改 `nightly-test-amd.yml`: 移除 4 处针对不同 GLM-5.x 作业的 `pip install git+https://github.com/huggingface/transformers.git@96f807a33b75` 命令。
2. 修改 `nightly-test-amd-rocm720.yml`: 类似移除 4 处固定命令。

变更仅涉及 CI 配置，无代码逻辑修改。

无关键源码片段，变更仅为 YAML 文件中的行删除。

评论区精华

无实质性讨论。

风险与影响

- 回归风险较低，因为移除的是降级版本，但需观察夜间 CI 是否通过。
- 影响范围局限于 AMD 夜间 CI 的 8 个 GLM-5.x 作业。
- 无测试覆盖变更，依赖回归依赖 CI 检测。

关联脉络

- PR #26106 的 `Cohere2MoeConfig` 崩溃问题由旧版 transformers 引起，本 PR 修复了根本原因。