

PR #47465 完整报告

vllm-project/vllm

[CI] Pin modelscope version to fix test breakage

合并时间: 2026-07-03 10:33

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/47465>

执行摘要

- 一句话: 锁定 modelscope 版本修复 CI 失败
- 推荐动作: PR 内容简单直接, 是典型的紧急依赖锁定修复。可快速 approve 合并, 后续应跟踪 modelscope 更新, 并在适当时候解除版本限制。

功能与动机

PR body 中提到 'We can update the code to properly tolerate more versions after this. But this is to get CI green, we should be pinning all of our dependencies regardless.' 即 modelscope 新版本破坏了测试, 需要紧急 fix 恢复 CI。

实现拆解

1. 锁定依赖版本上限: 在 requirements/test/xpu.in 中将 modelscope 改为 modelscope<1.38。
2. 更新所有 CI 配置文件中的 pip install 命令: 在 .buildkite/test-amd.yaml、.buildkite/intel_jobs/misc_intel.yaml、.buildkite/test_areas/misc.yaml 中, 将 pip install modelscope 改为 pip install 'modelscope<1.38'。
3. 更新 Dockerfile 和 NPU 脚本: 在 docker/Dockerfile 和 .buildkite/scripts/hardware_ci/run-npu-test.sh 中同样更新 modelscope 安装命令。
4. 无源码逻辑变更, 所有改动均为基础设施配置。

关键文件:

- requirements/test/xpu.in (模块 依赖配置; 类别 test; 类型 test-coverage) : 核心依赖定义文件, 将 modelscope 版本上限锁定为 <1.38。
- .buildkite/test-amd.yaml (模块 CI 配置; 类别 config; 类型 configuration) : AMD CI 配置, 修改了 regression 测试的 pip install 命令。
- .buildkite/intel_jobs/misc_intel.yaml (模块 CI 配置; 类别 config; 类型 configuration) : Intel GPU CI 配置, 同样修改了 regression 测试的安装命令。
- .buildkite/test_areas/misc.yaml (模块 CI 配置; 类别 config; 类型 configuration) : 通用测试区域配置, 修改了 regression 测试的安装命令。
- docker/Dockerfile (模块 Docker; 类别 infra; 类型 infrastructure) : Docker 构建文件, 修改了 uv pip install 中的 modelscope 版本。

- `.buildkite/scripts/hardware_ci/run-npu-test.sh` (模块 CI 脚本; 类别 `infra`; 类型 `infrastructure`) : NPU 测试脚本, 修改了 `pip install` 命令中的 `modelscope` 版本。

关键符号: 未识别

评论区精华

评论中 [AndreasKaratzas](#) 指出一个 AMD 多模态测试的回归与本次 PR 无关, 并提及可能是另一个 PR #47483 修复的问题。但未涉及技术争议。

- AMD 多模态测试回归是否与 PR 有关 (other): 确认与 PR 无关, 作者提及另一个 PR #47483 可能修复该问题。

风险与影响

- 风险: 风险极低。本次变更只是降低 `modelscope` 版本上限, 没有修改任何业务逻辑。可能的风险是如果未来需要新版本 `modelscope` 的特性, 需要重新评估并解除锁定。
- 影响: 影响范围限于 CI 测试环境, 确保所有使用 `modelscope` 的测试 (如 Regression 测试) 不会因 `modelscope` 新版本的 `breaking change` 而失败。对用户无影响。
- 风险标记: 临时锁定依赖

关联脉络

- PR #47483 [ModelRunner V2][BugFix] Free all model refs on shutdown: Issue 讨论中提及此 PR 可能修复 AMD 多模态测试回归。