

PR #6291 完整报告

verl-project/verl

[doc] chore: update vllm and vllm ascend from 0.13.0 to 0.18.0 in docs and dockerfile

合并时间: 2026-05-11 14:19

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6291>

执行摘要

该 PR 将 Ascend NPU 场景的 Dockerfile 和文档中引用的 vllm 和 vllm-ascend 版本从 v0.13.0 更新到 v0.18.0。但在 update 过程中，引入了两个严重问题：移除了 `torch_npu` 关键依赖（导致环境不可用）和禁用了 Git SSL 验证（引入安全风险）。虽然 PR 已合并，但建议后续尽快修复。

功能与动机

PR 描述为“Update vllm/vllm-ascend from v0.13.0 to v0.18.0 in docs and dockerfile”，旨在让 Ascend 平台用户可以使用新版 vllm 及 vllm-ascend 的功能或修复。

实现拆解

1. Dockerfile 版本更新：在两个 Dockerfile (Dockerfile.ascend_8.5.0_a3 和 Dockerfile.ascend_8.5.0_a2) 中，将 git clone 命令中 vllm 的分支从 v0.13.0 改为 v0.18.0，vllm-ascend 的分支从 releases/v0.13.0 改为 releases/v0.18.0。
2. Dockerfile 额外变更：仅在 Dockerfile.ascend_8.5.0_a3 中添加了华为云 pip 镜像配置和 ENV `GIT_SSL_NO_VERIFY=true`（用于绕过 Git SSL 验证）。两个 Dockerfile 中都移除了 `pip install mbridge` 命令中的 `torch_npu==2.8.0` 部分，改为只安装 mbridge。
3. 文档同步：在 docs/ascend_tutorial/quick_start/ascend_quick_start.rst 中，更新了文档示例中 git clone 的版本号，与 Dockerfile 保持一致。

无显著源码变更，均为配置和文档行修改。

评论区精华

gemini-code-assist[bot]: [high] 移除 `torch_npu==2.8.0` 是回归。`torch_npu` 是 Ascend NPU 的关键依赖，未在 `requirements-npu.txt` 中列出。此处移除将导致 NPU 特定操作失败。文档仍明确列出 `torch_npu == 2.8.0` 作为要求。请恢复显式安装。

gemini-code-assist[bot]: [security-high] 通过 `GIT_SSL_NO_VERIFY=true` 禁用 Git SSL 验证会暴露构建过程于中间人攻击。建议安装 CA 证书或使用安全代理。

审核者 wucong25 直接批准了 PR，未对上述评论进行回复或解决。

风险与影响

- 回归风险：移除 torch_npu 依赖后，Ascend NPU 环境下使用该 Dockerfile 构建的镜像将无法正常运行任何 NPU 运算，属于破坏性变更。
- 安全风险：GIT_SSL_NO_VERIFY=true 使所有 git clone 操作跳过证书验证，在构建过程中容易遭受中间人攻击。
- 兼容性风险：vllm 从 0.13.0 直接跳至 0.18.0，涉及多个大版本更新，可能存在 API 不兼容，但 PR 未包含任何适配代码。

影响范围：Ascend NPU 用户，尤其是使用这两个 Dockerfile 构建环境的开发者。

关联脉络

与该 PR 类似的版本升级有 #6222 (aarch64 vllm 0.17→0.18)，但 #6222 同时更新了测试和适配代码，本 PR 仅更新版本号。