

PR #7028 完整报告

verl-project/verl

[env] fix: Sync Qwen3.5 Dockerfile from 0.8.0 Branch to Main and Pin PyArrow Version

合并时间: 2026-07-13 19:44

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7028>

执行摘要

此 PR 将 Qwen3.5 所需的 Ascend 相关 Dockerfile 从 0.8.0 分支同步到 main 分支，同时固定 PyArrow 版本上限以规避 25.0.0 的兼容问题。变更集中在 Ascend Dockerfile 和 NPU requirements 文件，不涉及核心训练 / 推理逻辑，风险较低。

功能与动机

为了在 main 分支上支持 Qwen3.5 模型的 Ascend 环境训练，需要将 0.8.0 分支中已验证的 Dockerfile 同步过来。此外，PyArrow 25.0.0 存在不兼容问题，因此需要将版本上限固定为 24.0.0。PR body 明确说明: "pin the PyArrow version to <=24.0.0 to avoid issues with version 25.0.0"。

实现拆解

1. Dockerfile 统一更新 (`docker/ascend/Dockerfile.ascend_9.0.0_a2` 和 `Dockerfile.ascend_9.0.0_a3`) :
 - 添加克隆 Megatron-Bridge 仓库并 checkout 到指定 commit `de93536e9028ecf1e4dc28608dc80f336dcdfe59`。
 - 添加克隆 transformers 仓库并 checkout 到指定 commit `cc7ab9be508ce6ed3637bba9e50367b29b742dc6`。
 - 在 x86_64 架构下额外安装 `torchvision==0.24.0+cpu`。
 - 将 transformers 的安装方式从 `pip install transformers==5.3.0` 改为从本地克隆的源码进行编辑安装 (`cd transformers && pip install -e .`)。
 - 安装额外依赖: `nvidia-modelopt[torch]>=0.37.0`、`flash-linear-attention==0.5.0`、`qwen_vl_utils`、`viztracer`、`mathruler`。
 - 在 Dockerfile 末尾设置环境变量 `PYTHONPATH=/Megatron-Bridge/src:$PYTHONPATH`。
2. 依赖版本固定 (`requirements-npu.txt`) :
 - 将 pyarrow 的版本约束从 `>=15.0.0` 改为 `>=15.0.0,<=24.0.0`，阻止 25.0.0 的安装。
3. 提交与合并: 单个 commit，由 wucong25 审核并批准。

(无可用关键源码片段)

评论区精华

- `gemini-code-assist[bot]` (high 优先级评论)：指出 Dockerfile 中使用 `git clone --depth 1` 后再 `git fetch --depth 1 origin <sha>` 的方式效率低且不可靠，推荐改用 `--filter=blob:none` 进行 blobless clone，以便完整 clone 元数据并稳定获取指定 commit。该建议未被采纳。

风险与影响

- 风险：Dockerfile 中的 `git clone` 策略若网络不稳定或远程服务不支持浅表 `fetch` 指定 SHA，可能导致构建失败。固定 PyArrow 版本上限可能在未来需要更新，但当前是合理的安全措施。
- 影响：仅影响 Ascend 平台使用 Qwen3.5 的用户。这些用户可获得正确的环境配置。对 NVIDIA 等其他平台无影响。

关联脉络

此 PR 是简单的分支同步，未发现与其他 PR 的直接关联。但值得注意的是，近期已有多个涉及 Ascend Dockerfile 的 PR（如 #7008、#6998、#6983、#6981、#7002），表明 Ascend 支持正在快速迭代中。