

PR #7345 完整报告

verl-project/verl

[ci] chore: Fix npu nightly ci

合并时间: 2026-08-12 09:43

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7345>

执行摘要

- 一句话: 修复 NPU nightly CI, 升级 Ascend 依赖版本
- 推荐动作: 建议 CI 维护者精读, 尤其是 Docker 临时 tag 清理逻辑和 veomni 安装方式, 可作为 Ascend CI 标准配置参考。同时建议后续将依赖版本 (Megatron / MindSpeed / triton-ascend) 集中到常量或锁文件管理, 避免散落在 workflow 和脚本中。值得关注的设计决策是: 用 `pip install` 替代 `uv` 以支持 `--ignore-requires-python`, 以及用 `crane` 清理临时 tag, 这两点对快速维护 CI 有借鉴意义。

功能与动机

NPU nightly CI 在最近一段时间频繁失败, 主要原因是依赖版本不匹配 (如 vLLM 0.18.0 与新版镜像不兼容)、缺少 veomni 后端、以及 Docker 构建流程中残留的临时 tag 污染镜像仓库。PR 标题明确说明是 'Fix npu nightly ci', 提交记录也反映了多次版本回退和重试, 最终通过固定依赖版本和清理临时 tag 来恢复 CI 稳定性。

实现拆解

按以下步骤完成修复:

1. 修复 nightly CI 依赖: 在 `.github/workflows/nightly_ascend.yml` 的“安装当前仓库”步骤后追加 `pip install veomni==0.1.11 --ignore-requires-python --no-deps --index-url https://pypi.org/simple/`, 因为镜像中缺少 veomni 后端; 同时在 `requirements-npu.txt` 中移除不再需要的 `triton-ascend==3.2.1`, 并更新测试脚本 `tests/special_npu/nightly_ci_ascend/run_grpo_qwen3_30b_veomni_fsdp.sh` 的 `VLLM_VERSION=0.23.0`。
2. 固定 Megatron 与 MindSpeed 版本: 在 `.github/workflows/e2e_ppo_trainer_megatron_sglang_ascend.yml` 的两个 job 中新增“Configure related dependencies”步骤, 分别从 GitHub 和 gitcode 克隆 `core_r0.18.0` 分支的 Megatron-LM 与 MindSpeed, 避免因仓库默认分支漂移导致 CI 不稳定。
3. 清理 Docker 临时 tag: 在 `.github/workflows/docker-build-ascend-a2.yml` 和 `a3.yml` 中, 于多架构 manifest 合并后增加两步“Delete temporary tags”, 使用 `crane` (`go-containerregistry v0.20.2`) 删除构建期间生成的临时 tag (如 `tmp-a2-amd64`、`tmp-a2-v080-amd64` 等), 保持镜像仓库整洁。

4. 文档补充：修改 `docs/ascend_tutorial/get_start/install_guidance.rst`，将 NPU 自动识别说明改造成 `.. note::` 格式，并补充 A3 每卡 2 die、需要设置 `n_gpus_per_node=16` 的说明，方便用户正确配置资源。

测试方面没有新增单元测试，但更新了 nightly 相关配置和脚本，属于 CI 基础设施维护。

关键文件：

- `.github/workflows/nightly_ascend.yml`（模块 CI 工作流；类别 infra；类型 infrastructure）：NPU nightly CI 的入口，修复了缺少 veomni 后端导致 CI 失败的问题，新增 pip 安装步骤。
- `.github/workflows/docker-build-ascend-a2.yml`（模块 镜像构建；类别 infra；类型 infrastructure）：Docker 镜像构建流程，新增 crane 删除临时 tag 步骤，避免镜像仓库 tag 堆积。
- `.github/workflows/docker-build-ascend-a3.yml`（模块 镜像构建；类别 infra；类型 infrastructure）：与 a2 对称的镜像构建流程，同样增加 crane 清理临时 tag 步骤。
- `.github/workflows/e2e_ppo_trainer_megatron_sglang_ascend.yml`（模块 E2E 工作流；类别 infra；类型 infrastructure）：Megatron + sglang 的 NPU E2E CI，新增依赖配置步骤固定 Megatron-LM 和 MindSpeed 版本。
- `tests/special_npu/nightly_ci_ascend/run_grpo_qwen3_30b_veomni_fsdp.sh`（模块 测试脚本；类别 test；类型 test-coverage）：Nightly 实际运行的 GRPO 训练脚本，升级 vLLM 版本以适配新镜像。
- `docs/ascend_tutorial/get_start/install_guidance.rst`（模块 文档说明；类别 docs；类型 documentation）：补充 A2/A3 设备差异说明，帮助用户正确配置 `n_gpus_per_node`。
- `requirements-npu.txt`（模块 依赖清单；类别 docs；类型 documentation）：移除不再需要的旧版 triton-ascend 依赖，保持依赖清单简洁。

关键符号：未识别

关键源码片段

`.github/workflows/nightly_ascend.yml`

NPU nightly CI 的入口，修复了缺少 veomni 后端导致 CI 失败的问题，新增 pip 安装步骤。

```
# .github/workflows/nightly_ascend.yml 中的安装步骤
- name: Install the current repository
  run: |
    # 先安装当前仓库（editable 模式），确保代码是最新的
    uv pip install -v -e .
    # veomni 是 NPU 上 FSDP 的扩展后端，镜像中缺失需要单独安装
    # uv 不支持 --ignore-requires-python，因此这里必须用 pip
    pip install veomni==0.1.11 --ignore-requires-python --no-deps --index-url https://pypi.org/simple/
```

`tests/special_npu/nightly_ci_ascend/run_grpo_qwen3_30b_veomni_fsdp.sh`

Nightly 实际运行的 GRPO 训练脚本，升级 vLLM 版本以适配新镜像。

```

# tests/special_npu/nightly_ci_ascend/run_grpo_qwen3_30b_veomni_fsdp.sh 的 NPU
环境配置分支
case "${DEVICE}" in
    gpu)
        # GPU 环境保持默认配置
        ;;
    npu)
        export TASK_QUEUE_ENABLE=1
        export HCCL_OP_EXPANSION_MODE="AIV"
        export VLLM_USE_V1=1
        # 升级到 0.23.0 以匹配新版 Ascend 镜像中的 vLLM-Ascend
        export VLLM_VERSION=0.23.0
        export VLLM_ASCEND_ENABLE_NZ=0
        export HCCL_BUFFSIZE=610
        export CKPT_DIR="./ckpt30b"
        export PYTORCH_NPU_ALLOC_CONF=max_split_size_mb:1024
        export CUDA_DEVICE_MAX_CONNECTIONS=1
        # A3 每卡 2 die, 单机 8 卡等价于 16 个 NPU 设备
        n_devices_per_node=16
        ;;
    *)
        echo "Unsupported DEVICE=${DEVICE}. Expected 'gpu' or 'npu'." >&2
        exit 1
        ;;
esac

```

评论区精华

Review 中唯一一条有效评论来自 yyy2000, 指出 [.github/workflows/nightly_ascend.yml](#) 中新增的 `uv pip install veomni==0.1.11 --ignore-requires-python --no-deps` 用法有误: `uv` 并不支持 `--ignore-requires-python` 参数。作者 LeoYao123 回复“fixed”, 最终将命令改为 `pip install veomni==0.1.11 --ignore-requires-python --no-deps --index-url https://pypi.org/simple/`, 解决了该问题。该线程已关闭, 属于工具链用法修正, 无遗留争议。

- `uv pip install` 不支持 `--ignore-requires-python (correctness)`: 作者 LeoYao123 回复 'fixed', 最终将命令从 `uv` 改为 `pip`, 问题解决。

风险与影响

- 风险:
 - [.github/workflows/docker-build-ascend-a2.yml](#) 与 [a3.yml](#) 中新增的 `crane` 删除 tag 步骤依赖 GitHub Secrets (`QUAY_USERNAME` / `QUAY_PASSWORD`), 如果密钥缺失或失效, 步骤会失败; 且 `ll true` 会掩盖真正的删除失败, 可能让临时 tag 残留。
 - [e2e_ppo_trainer_megatron_sglang_ascend.yml](#) 中克隆固定分支 `core_r0.18.0`, 如果该分支被上游删除或迁移, CI 会持续失败, 需要建立镜像缓存或版本锁定。
 - [run_grpo_qwen3_30b_veomni_fsdp.sh](#) 中 `VLLM_VERSION` 从 0.18.0 升到 0.23.0, 属于跨小版本升级, 可能引入 vLLM 侧行为变化, 需要关注 NPU 上是否出现新的兼容性问

题。

- requirements-npu.txt 移除 triton-ascend==3.2.1 后，如果其他代码路径仍依赖旧版 triton，可能导致运行时 ImportError。
- 影响：影响范围集中在 NPU (Ascend) CI 和镜像构建流程，直接影响到依赖 NPU 的 daily / nightly 验证。对普通 GPU 用户无影响，但 Ascend 相关的开发者和 CI 维护者会受益：CI 恢复正常后，Megatron + vLLM + sglang 的端到端测试可以持续运行，进而保障后续算法改动在 NPU 上的回归验证。文档补充也帮助 A3 用户正确配置设备数量。改动本身不改动任何训练或推理源码，因此运行时不产生行为变化。
- 风险标记：依赖版本固定缺少校验，crane 删除 tag 依赖 Secrets, NPU 专用路径无 GPU CI 覆盖

关联脉络

- PR #7305 [ci] chore: Update npu docker build file: 该 PR 引入了临时 tag 用于多架构构建，本 PR 新增 crane 步骤清理这些临时 tag，是直接的后续修复。
- PR #7337 [ci] chore: add three baselines for npu's nightly ci: 同为 NPU nightly CI 的扩展，本 PR 修复该扩展后出现的依赖缺失问题。
- PR #7311 [ci] chore: Fix docker image upload: 共同维护 docker-build-ascend-a2/a3.yml，上 upload 逻辑调整后，本 PR 继续完善构建后的清理流程。
- PR #7338 [ci] chore: Update npu docker image cann version: NPU 镜像版本更新的延续，本 PR 进一步同步 vLLM、Megatron 等运行时依赖。