

PR #7541 完整报告

verl-project/verl

[ci] chore: use Ascend recipe baselines for NPU nightly CI

合并时间: 2026-08-24 19:46

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7541>

执行摘要

- 一句话: NPU 夜间 CI 改用 Ascend recipe 基线文件
- 推荐动作: 该 PR 为运维性变更, 若关注 CI 基础设施的配置管理, 可精读工作流中环境变量和 curl 下载的使用方式。值得注意的设计决策是将基线来源外置, 便于统一维护, 但需配套监控外部仓库的可用性。未来若扩展至更多拓扑或 A3 基线, 可参考此模式。

功能与动机

PR 描述指出, 原先的 `check_npu.py` 和性能基线文件依赖 CI 运行器上预置的文件, 导致基线更新滞后或配置不一致。改为从 `verl-ascend-recipe` 仓库获取, 可统一基线来源, 减少本地文件管理成本, 并确保基线匹配当前 `recipe` 版本。

实现拆解

变更集中在 GitHub Actions 工作流文件 `.github/workflows/nightly_ascend.yml`, 具体拆解如下:

1. 新增环境变量配置: 在 workflow 顶层新增 `env` 块, 定义 `ASCEND_RECIPE_BASELINE_URL`, 指向 `https://raw.githubusercontent.com/verl-project/verl-ascend-recipe/main/baseline`, 作为后续下载的基础路径。
2. 替换本地文件引用: 在四个夜间 CI 任务 (`ppo_qwen3-8b-fsdp-vllm`、`grpo_qwen3_vl_8b-Instruct_fsdp2-vllm`、`dapo-moonlight-16b-megatron-vllm`、`gsdp-qwen3-30b-megatron-vllm`) 的检查步骤中, 将原有的 `python check_npu.py --base <local_path>` 改为先用 `curl --fail --location --retry 3` 下载 `check_npu.py` 和对应的基线文件到当前目录, 再以下载的 `baseline.txt` 作为基准执行检查。
3. 统一基线文件命名: 四个任务均使用 `baseline.txt` 作为下载后临时文件名, 避免路径冲突, 并简化命令。
4. 配套调整: 未新增测试, 但因涉及 CI 配置变更, 需在 NPU 环境实际运行夜间 CI 验证网络可达性与下载正确性。

关键文件:

- `.github/workflows/nightly_ascend.yml` (模块 CI 配置; 类别 `infra`; 类型 `infrastructure`): 唯一变更文件, 定义了基线下载逻辑, 影响所有 NPU 夜间 CI 的检查环节。

关键符号: 未识别

关键源码片段

[.github/workflows/nightly_ascend.yml](#)

唯一变更文件，定义了基线下载逻辑，影响所有 NPU 夜间 CI 的检查环节。

```
# 在 workflow 顶层新增 env，定义基线下载基础 URL
env:
  ASCEND_RECIPE_BASELINE_URL: "https://raw.githubusercontent.com/verl-project/verl-ascend-recipe/main/baseline"

jobs:
  # 以 ppo_qwen3-8b-fsdp-vllm 任务为例
  nightlyCI_ppo-qwen3-8b-fsdp-vllm_ascend:
    steps:
      - name: Running ... checking script
        run: |
          cd /root/.cache/nightly_log/
          # 下载检查脚本和基线文件，--fail 确保失败即退出，--retry 3 增强鲁棒性
          curl --fail --location --retry 3 --output check_npu.py "${ASCEND_RECIPE_BASELINE_URL}/check_npu.py"
          curl --fail --location --retry 3 --output baseline.txt "${ASCEND_RECIPE_BASELINE_URL}/a2/nightly_log/run_ppo_qwen3-8b_fsdp_npu/baseline_ppo_qwen3-8b_fsdp_npu.txt"
          # 用下载的基线文件执行性能检查
          python check_npu.py --log run_ppo_qwen3-8b_fsdp_npu/run_ppo_qwen3-8b_fsdp_npu.log --base baseline.txt
```

评论区精华

该 PR 无 review 评论和讨论线程，仅有维护者 wucong25 的 APPROVED 审核。由于缺乏讨论，无法获取决策背后的权衡细节。

- 暂无高价值评论线程

风险与影响

- 风险：主要风险在于对外部仓库 verl-ascend-recipe 的网络依赖：
 - 若该仓库不可达或路径变更，下载失败将导致 CI 检查失败；curl --fail --retry 3 可缓解瞬时故障，但无法应对持久性问题。
 - 基线文件若与本地运行日志路径不匹配（例如模型名、拓扑变化），可能导致误判性能达标或失败。
 - 当前仅下载 A2 基线（a2/nightly_log/...），未包含 A3 基线，若未来需要支持 A3 节点需补充。
 - 下载的 check_npu.py 可能随 recipe 更新而变化，存在版本兼容性风险。
- 影响：影响范围限定于 NPU 夜间 CI 的检查和验收环节，不涉及训练核心逻辑。对用户无直接影响，但提升了 CI 维护的便利性，使其能自动跟随上游基线更新。团队需确保目标环境可访问 raw.githubusercontent.com，并关注 recipe 仓库的访问权限和稳定性。

- 风险标记：外部网络依赖，基线路径耦合，缺少测试覆盖

关联脉络

- PR #7549 [ci] chore: correct step naming for Ascend ci: 同一天合并的 Ascend CI 配置修正，均修改 `.github/workflows/nightly_ascend.yml`，属于同一 CI 演进线。
- PR #7558 [ci] fix: drop stale `enable_chunked_prefill=False` from Ascend NPU scripts: 同样针对 Ascend NPU 测试脚本的清理，与本 PR 同属 NPU CI 维护主题。
- PR #7456 [ci] chore: Update ascend ci image: 调整 NPU 节点数和 CI 脚本默认值，与本 PR 一样关注 NPU 夜间 CI 运行环境。