

PR #7066 完整报告

verl-project/verl

[ci] chore: fix npu and gpu ci

合并时间: 2026-07-16 19:48

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7066>

执行摘要

- 一句话: 修复 GPU 与 Ascend CI 工作流依赖配置
- 推荐动作: 该 PR 是典型的 CI 维护操作, 价值在于使 CI 流水线更加健壮。建议通读变更以了解工作流依赖管理实践, 可作为相似 CI 修复的参考。

功能与动机

CI 工作流中某些 job 缺少显式依赖, 导致 Megatron 策略测试可能在 FSDP2 测试尚未完成时启动或跳过, 影响 CI 正确性和资源利用效率。PR body 中明确提到“fix ci npu”, 表明主要目的是修复 NPU CI 问题。

实现拆解

1. 调整 GPU 工作流依赖: 在 `e2e_fully_async_policy.yml` 和 `e2e_one_step_off_policy.yml` 中, 修改 Megatron 策略测试 job 的 `needs` 字段, 从仅依赖 `setup` 改为依赖 `[setup, e2e_fully_async_policy_fsdp2]` 和 `[setup, e2e_one_step_off_policy_fsdp2]`, 确保执行顺序。
2. 调整 Ascend 工作流依赖: 在 `e2e_fully_async_policy_ascend.yml` 和 `e2e_one_step_off_policy_ascend.yml` 中, 为 Megatron 策略测试 job 新增 `needs: e2e_fully_async_policy_fsdp2_ascend` 和 `needs: e2e_one_step_off_policy_fsdp2_ascend`, 确保对应 FSDP2 测试先行完成。
3. 更新 NPU 单元测试排除列表: 在 `npu_unit_tests.yml` 中, 在 `pytest -k` 参数中追加 `and not test_sender_accepts_strided_tensor`, 跳过该已知失败的测试用例。

关键文件:

- `.github/workflows/e2e_fully_async_policy.yml` (模块 CI; 类别 infra; 类型 infrastructure) : 修改 Megatron 策略 job 的 `needs` 依赖, 增加 FSDP2 测试前置要求
- `.github/workflows/e2e_one_step_off_policy.yml` (模块 CI; 类别 infra; 类型 infrastructure) : 同步修改 `one_step_off` 策略的 Megatron job 依赖
- `.github/workflows/e2e_fully_async_policy_ascend.yml` (模块 CI; 类别 infra; 类型 infrastructure) : 为 Ascend Megatron job 新增 FSDP2 依赖 (之前无 `needs`)
- `.github/workflows/e2e_one_step_off_policy_ascend.yml` (模块 CI; 类别 infra; 类型 infrastructure) : 为 Ascend `one_step_off` Megatron job 新增 FSDP2 依赖

- `.github/workflows/npu_unit_tests.yml` (模块 CI; 类别 infra; 类型 infrastructure) : 在 NPU 单元测试中排除另一个已知失败的测试用例

关键符号: 未识别

评论区精华

该 PR 没有显著的 review 讨论。发起者 `wucong25` 提交后由 `wuxibin89` 直接批准, 无评论互动。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。变更仅涉及 CI 配置 (`.github/workflows/*.yml`), 不涉及任何源码逻辑。唯一的风险是若新的依赖关系导致循环依赖或阻塞, 但此处新增的依赖均为已知存在的 job, 不会引入循环。
- 影响: 影响范围仅限于 CI 流水线。GPU 和 Ascend 上的 Megatron 策略测试将不再并行于 FSDP2 测试启动, 而是等待其完成, 保证了测试顺序。同时 NPU 单元测试跳过了一个已知失败的用例, 减少 CI 噪音。
- 风险标记: 暂无

关联脉络

- PR #7035 [ci] test: update ascend ci docker tag: 同为 Ascend CI 维护操作, 可能涉及相同的工作流文件
- PR #7040 [ci, trainer] fix: Fix ci AssertionError for the environment variable `CUDA_DEVICE_MAX_CONNECTIONS`: 同为 CI 修复, 展示了团队持续维护 CI 稳定性的行为模式