

PR #6273 完整报告

verl-project/verl

[ci] feat: add gspo qwen3-30b in nightly npu ci

合并时间: 2026-05-09 15:58

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6273>

执行摘要

- 一句话: 为 NPU nightly CI 添加 GSPO Qwen3-30B 测试脚本
- 推荐动作: 建议开发者仔细阅读 gemini-code-assist 的评论, 并在后续 PR 中补充 ref 的 `override_transformer_config`。本 PR 本身作为 CI 增强是合理的, 但配置一致性值得跟进。

功能与动机

为了在 nightly CI 中自动验证 GSPO 训练在 NPU 上的正确性, 覆盖 Qwen3-30B MoE 模型的关键路径, 防止核心功能回归。

实现拆解

1. 新增测试脚本 `tests/special_npu/nightly_ci_ascend/run_gspo_qwen3_30b_megatron_npu.sh`, 定义了模型路径、训练参数、actor/ref/rollout 配置。其中 actor 启用了 `override_transformer_config` (如 `moe_router_dtype=fp32`、`moe_permute_fusion`、`apply_rope_fusion` 等), 而 ref 配置缺少这些覆盖。
2. 在 `nightly_ascend.yml` 中添加 job `nightlyCI_gspo-qwen3-30b-megatron-vllm_ascend`, 设置容器镜像、资源规格和运行步骤。
3. 在 `e2e_ascend.yml` 中将 `vllm_rl_job` 的 runner 规格从 `linux-aarch64-a3-16` 改为 `linux-aarch64-a3-8`, 降低资源占用。

关键文件:

- `tests/special_npu/nightly_ci_ascend/run_gspo_qwen3_30b_megatron_npu.sh` (模块 测试脚本; 类别 test; 类型 test-coverage): 新增的测试脚本, 定义了 GSPO 训练的完整参数; 其 REF 配置缺失 `override_transformer_config` 是 review 焦点。
- `.github/workflows/nightly_ascend.yml` (模块 CI 配置; 类别 infra; 类型 infrastructure): 添加了新的 nightly CI job, 调用上述测试脚本, 是 CI 配置的核心变更。
- `.github/workflows/e2e_ascend.yml` (模块 CI 配置; 类别 infra; 类型 infrastructure): 修改了 VLM RL 测试 job 的 runner 规格, 属于资源优化调整。

关键符号: 未识别

评论区精华

gemini-code-assist[bot] 在代码审查中指出 ref 配置缺少 actor 中使用的 `override_transformer_config` (如 `moe_router_dtype=fp32`、`moe_permute_fusion`、`apply_rope_fusion`)，对于 MoE 模型，这可能导致 actor 和 ref 的 log-prob 计算不一致，即使 KL 已禁用也不够健壮。此问题未被作者回应或修复，PR 仍被批准合并。

- REF 配置缺少 `override_transformer_config` (design): PR 虽已被批准合并，但该问题未在本次 PR 中修复，需后续跟进。

风险与影响

- 风险：核心风险在于 ref 配置不完整：如果未来启用 KL 或需要精确的比值，将产生训练偏差。此外，新增加的 nightly job 会延长 CI 总时长，需要监控资源消耗。E2E CI 的 runner 规格缩减可能增加任务失败概率，需观察稳定性。
- 影响：对用户无直接影响。对团队：需要关注 ref 配置一致性问题，建议后续修复。对系统：增加了 nightly CI 的覆盖范围，但可能增加维护成本。
- 风险标记：配置一致性风险，资源占用增加

关联脉络

- PR #6253 [ci] fix: remove the redundant config: 同为 NPU nightly CI 配置修复，清理冗余参数，与本 PR 的 CI 增强在同一领域。