

PR #6429 完整报告

verl-project/verl

[megatron] fix: set use_mbridge to True for some npu scripts

合并时间: 2026-05-21 15:54

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6429>

执行摘要

- 一句话: 修复 NPU 脚本中 use_mbridge 参数为 True
- 推荐动作: 该 PR 本质上是跟随上游校验的合规性修复, 无技术深度, 但值得注意: 新增的 flash attention 和重计算配置可能对性能有积极影响, 可作为 NPU 脚本的默认优化模板。

功能与动机

PR#4530 添加了对 use_mbridge 标志的校验, 要求必须为 True。现有 NPU 脚本中该标志设置为 False 会导致校验失败, 因此需要修正以保持兼容。

实现拆解

1. 修改 actor 和 ref 配置中的 use_mbridge 值: 将 actor_rollout_ref.actor.megatron.use_mbridge=False 和 actor_rollout_ref.ref.megatron.use_mbridge=False 均改为 True。
2. 新增两项 transformer 优化配置: override_transformer_config.use_flash_attn=True (启用 flash attention) 和 override_transformer_config.recompute_method=uniform (设置统一重计算策略)。
3. 涉及两个文件: run_qwen3_235b_256k_megatron.sh 和 run_qwen3_30b_a3b_megatron.sh, 改动均为配置键调整, 无核心逻辑变更。

关键文件:

- examples/ascend_extras/grpo_trainer/run_qwen3_235b_256k_megatron.sh (模块 训练脚本; 类别 other; 类型 configuration): Qwen3 235B 模型的 GRPO 训练脚本, 修复了 actor 和 ref 的 use_mbridge 配置, 并新增了 flash attention 和重计算优化参数。
- examples/ascend_extras/grpo_trainer/run_qwen3_30b_a3b_megatron.sh (模块 训练脚本; 类别 other; 类型 configuration): Qwen3 30B A3B 模型的训练脚本, 变更内容与 235B 脚本完全一致。

关键符号: 未识别

关键源码片段

[examples/ascend_extras/grpo_trainer/run_qwen3_235b_256k_megatron.sh](#)

Qwen3 235B 模型的 GRPO 训练脚本, 修复了 actor 和 ref 的 use_mbridge 配置, 并新增了 flash attention 和重计算优化参数。

```
# ... 其他配置 ...
actor_rollout_ref.actor.megatron.use_dist_checkpointing=True
# 修复：必须为 True 以通过 PR#4530 的校验
actor_rollout_ref.actor.megatron.use_mbridge=True
actor_rollout_ref.actor.megatron.dist_checkpointing_path=${MCORE_MODEL_PATH}
# 新增优化：启用 flash attention 和统一重计算
actor_rollout_ref.actor.megatron.override_transformer_config.use_flash_attn=True
actor_rollout_ref.actor.megatron.override_transformer_config.recompute_method=uniform

# ... ref 配置同理 ...
actor_rollout_ref.ref.megatron.use_mbridge=True
# 新增 ref 端优化配置
actor_rollout_ref.ref.megatron.override_transformer_config.use_flash_attn=True
```

examples/ascend_extras/grpo_trainer/run_qwen3_30b_a3b_megatron.sh

Qwen3 30B A3B 模型的训练脚本，变更内容与 235B 脚本完全一致。

```
# ... 其他配置 ...
actor_rollout_ref.actor.megatron.dist_checkpointing_path=${MCORE_MODEL_PATH}
actor_rollout_ref.actor.megatron.use_dist_checkpointing=True
# 修复：启用 mbridge 以符合新校验
actor_rollout_ref.actor.megatron.use_mbridge=True
# 新增优化：flash attention 和重计算方法
actor_rollout_ref.actor.megatron.override_transformer_config.use_flash_attn=True
actor_rollout_ref.actor.megatron.override_transformer_config.recompute_method=uniform

# ... ref 配置 ...
actor_rollout_ref.ref.megatron.use_mbridge=True
# ref 端也需要 flash attention 优化
actor_rollout_ref.ref.megatron.override_transformer_config.use_flash_attn=True
```

评论区精华

无 review 讨论。仅 Gemini Code Assist 自动生成无实质反馈的评论，并由 wucong25 直接批准。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅修改 shell 脚本中的配置参数，不会影响 Python 核心逻辑。但需确认 NPU 环境确实支持 use_mbridge=True，以及新增的 flash attention 和重计算配置不会与现有环境冲突。
- 影响：影响范围限于两个 Ascend NPU GRPO 训练脚本，使它们能在 use_mbridge 校验通过的情况下正常运行。对系统的其他模块（如 trainer、rollout 等）无影响。
- 风险标记：仅配置变更，无测试覆盖，依赖上游校验

关联脉络

- PR #4530 [megatron] add use_mbridge flag validation: 本 PR 是响应 PR#4530 新增校验的修复性变更, 该 PR 要求 use_mbridge 必须为 True。