PR #6066 完整报告

verl-project/verl

[recipe, cfg] fix: update NPU script parameters for Qwen3 GSPO and DAPO math recipes

合并时间：2026-04-20 17:17

原文链接：http://prhub.com.cn/verl-project/verl/pull/6066

执行摘要

- 一句话：修复 NPU 训练脚本参数，移除冗余环境变量并修正专家并行配置。
- 推荐动作：该 PR 值得快速浏览以了解 NPU 训练脚本的配置细节，但不需要深入研读。重点关注 review 中关于 enable_expert_parallel 参数无效的讨论，这揭示了 vLLM 引擎配置参数的实际生效机制。对于需要在 NPU 上启用专家并行的用户，应该参考 review 建议设置 expert_parallel_size 参数。

功能与动机

根据 PR 描述，在验证过程中发现了 NPU 配方脚本中的参数设置不正确或不一致，需要修复以确保脚本在 NPU 训练运行时的正确性和稳定性。具体来说，需要更新 run_qwen3_32b_gspo_npu.sh 和 dapo_30b_a3b_math_fsdp_npu.sh 两个脚本的参数设置。

实现拆解

1. 移除 GSPO 脚本冗余环境变量：在 examples/gspo_trainer/run_qwen3_32b_gspo_npu.sh 中删除了 export VLLM_ATTENTION_BACKEND=XFORMERS 这一行，该变量在 NPU 环境中可能已过时或不需要。
2. 尝试启用 DAPO 脚本专家并行：在 verl/experimental/fully_async_policy/shell/dapo_30b_a3b_math_fsdp_npu.sh 中添加了 actor_rollout_ref.rollout.engine_kwargs.vllm.enable_expert_parallel=True 参数，意图启用 vLLM 引擎的专家并行功能。
3. review 发现配置问题：gemini-code-assist[bot] 在 review 中指出，直接设置 enable_expert_parallel 是无效的，因为该标志在 vLLMHttpServer.launch_server 中会被显式覆盖。正确的做法是设置 actor_rollout_ref.rollout.expert_parallel_size > 1 来启用专家并行。

关键文件：

- examples/gspo_trainer/run_qwen3_32b_gspo_npu.sh（模块 示例脚本；类别 other；类型 configuration）：GSPO 训练脚本，移除了可能过时的环境变量配置
- verl/experimental/fully_async_policy/shell/dapo_30b_a3b_math_fsdp_npu.sh（模块 实验模块；类别 other；类型 configuration）：DAPO 训练脚本，尝试启用专家并行但配置方式被 review 指出无效

关键符号：未识别

关键源码片段

examples/gspo_trainer/run_qwen3_32b_gspo_npu.sh

GSPO 训练脚本，移除了可能过时的环境变量配置

```
# 环境变量配置部分
# 移除了 VLLM_ATTENTION_BACKEND=XFORMERS，该变量在 NPU 环境中可能不再需要
# 其他 NPU 相关配置保持不变
export HCCL_CONNECT_TIMEOUT=3600
export HCCL_ASYNC_ERROR_HANDLING=0
export CPU_AFFINITY_CONF=1
export VLLM_USE_V1=1
# export VLLM_ATTENTION_BACKEND=XFORMERS # 已移除
export VLLM_ASCEND_ENABLE_FLASHCOMM=1
export VLLM_ASCEND_ENABLE_PREFETCH_MLP=1
export VLLM_ASCEND_ENABLE_DENSE_OPTIMIZE=1
```

verl/experimental/fully_async_policy/shell/dapo_30b_a3b_math_fsdp_npu.sh

DAPO 训练脚本，尝试启用专家并行但配置方式被 review 指出无效

```
# 训练启动命令参数部分
# 添加了 enable_expert_parallel 参数，但 review 指出该参数会被内部逻辑覆盖
# 正确的启用方式应该是设置 expert_parallel_size > 1
python3 -m verl.experimental.fully_async_policy.fully_async_main \
  actor_rollout_ref.rollout.val_kwargs.top_k=${top_k} \
  actor_rollout_ref.rollout.val_kwargs.do_sample=True \
  actor_rollout_ref.rollout.val_kwargs.n=1 \
  actor_rollout_ref.rollout.engine_kwargs.vllm.enable_expert_parallel=True \  # 新增但无效
  actor_rollout_ref.ref.fsdp_config.param_offload=${ref_offload} \
  actor_rollout_ref.actor.fsdp_config.fsdp_size=${fsdp_size} \
  # 应该改为：actor_rollout_ref.rollout.expert_parallel_size=8 # 例如在 8-GPU 节点上
```

评论区精华

gemini-code-assist[bot] 在 review 中指出了关键问题：在 DAPO 脚本中添加的 enable_expert_parallel=True 参数是无效的，因为该标志在 verl/workers/rollout/vllm_rollout/vllm_async_server.py 第 311 行会被显式覆盖。正确的启用方式应该是设置 actor_rollout_ref.rollout.expert_parallel_size 大于 1（例如在 8-GPU 节点上为 Qwen3-32B 设置 8）。这个讨论揭示了配置参数的实际生效逻辑，但 PR 作者没有根据这个反馈进行进一步修改。

- enable_expert_parallel 参数无效性问题 (design): PR 中采用的配置方式不会生效，但作者没有根据反馈修改代码。

风险与影响

- 风险：

1. 配置无效风险：DAPO 脚本中添加的 *enable_expert_parallel=True* 参数实际上不会生效，可能导致用户误以为专家并行已启用，但实际运行时仍使用默认配置。
2. 兼容性风险：移除 VLLM_ATTENTION_BACKEND 环境变量可能在某些特定硬件或软件版本下影响注意力后端的选择，但考虑到这是 NPU 专用脚本，风险较低。
3. 回归风险：两个脚本的修改都很小，且不涉及核心训练逻辑，回归风险较低。

- 影响：
  1. 对用户的影响：使用这两个脚本进行 NPU 训练的用户会获得更简洁的环境配置（GSPO 脚本），但 DAPO 脚本的专家并行配置实际上并未正确生效，用户需要手动调整 *expert_parallel_size* 参数。
  2. 对系统的影响：仅影响脚本级别的配置，不改变训练框架的核心逻辑或数据流。
  3. 对团队的影响：这是一个小的配置修复，维护成本低，但揭示了配置参数传递机制的知识点。 - 风险标记：配置参数无效，缺少验证

关联脉络

- PR #5967 [fully_async] fix: Add Mindspeed Patch for Async Training on Ascend NPUs: 同样涉及 NPU 环境下的训练脚本修复，属于 NPU 兼容性相关 PR
- PR #6062 [fully_async, rollout, trainer, tool, cfg] fix: ROCm async training compatibility for AMD MI300X: 类似的环境配置修复 PR，针对不同硬件平台