PR #6582 完整报告

verl-project/verl

[model] feat: add Qwen3.5-122B Ascend support

合并时间：2026-06-03 16:17

原文链接：http://prhub.com.cn/verl-project/verl/pull/6582

执行摘要

- 一句话：支持 Qwen3.5-122B-A10B 在 Ascend NPU 上运行
- 推荐动作：该 PR 展示了多硬件场景下使用 shell 脚本进行配置分离的模式，结构清晰，但仍存在未解决的 review 问题。建议阅读以了解硬件适配方法，但在实际使用 NPU 任务前，应检查并补全缺失的 Ref 配置项和全局化公共步骤。

功能与动机

PR body 说明："This PR makes the Qwen3.5-122B-A10B Megatron GRPO script compatible with both GPU and Ascend NPU. It keeps the original GPU defaults and flow, adds DEVICE=gpu|npu auto-detection, and configures the NPU path for 4 nodes with Megatron settings aligned with the existing Qwen3.5-35B compatible script."

实现拆解

1. 设备自动检测：在脚本开头通过 python3 -c 'import torch_npu' 探测是否存在 torch_npu，自动设置 DEVICE=npu 或 gpu；用户也可通过环境变量覆盖。
2. 环境变量分离：针对 GPU 和 NPU 分别配置环境变量（如 CUDA_VISIBLE_DEVICES 和 NPU 特有的 OMP_NUM_THREADS、HCCL_* 等），封装在 *case* 语句内。
3. 并行与超参数差异化：TP/PP/EP/GEN_TP 等根据设备赋予不同的默认值（GPU：PP=2 EP=8；NPU：PP=4 EP=16）；训练微批次大小、log_prob 批次大小、参考模型批次大小等也分别设置。
4. Actor 模型配置：GPU 时设置 MoE 负载均衡和 rope_fusion；NPU 时设置 vanilla_mbridge=False 禁用标准 Megatron-Bridge 以适配 MindSpeed。Ref 模型对应配置（use_remove_padding=False、vanilla_mbridge=False）在合并版本中缺失，构成主要 review 争议。

关键文件：

- *examples/grpo_trainer/run_qwen3_5_122b_a10b_megatron.sh*（模块 启动脚本；类别 other；类型 core-logic）：唯一变更文件，增加 NPU 支持的核心脚本，包含设备检测、环境配置、并行参数分离和模型配置调整。

关键符号：未识别

关键源码片段

examples/grpo_trainer/run_qwen3_5_122b_a10b_megatron.sh

唯一变更文件，增加 NPU 支持的核心脚本，包含设备检测、环境配置、并行参数分离和模型配置调整。

```
# DEVICE auto-detect via torch_npu; override via environment variable DEVICE=${DEVICE:-$(python3-c'import torch_npu'2>/dev/null&&echonpu||echogpu)} case"${DEVICE}"in gpu) # GPU: set CUDA_VISIBLE_DEVICES, unset proxy (should be global per review) exportCUDA_VISIBLE_DEVICES=${CUDA_VISIBLE_DEVICES:-0,1,2,3,4,5,6,7} unsethttp_proxy unsethttps_proxy # Dataset download only in gpu branch; review suggested global scope hfdownloadtyzhu/geo3k--repo-typedataset--local-dir$HOME/data/geo3k ;; npu) # NPU: configure environment for Ascend/MindSpeed exportCPU_AFFINITY_CONF=${CPU_AFFINITY_CONF:-1} exportOMP_NUM_THREADS=${OMP_NUM_THREADS:-1} exportPYTORCH_NPU_ALLOC_CONF=${PYTORCH_NPU_ALLOC_CONF:-garbage_collection_threshold:0.8} exportHCCL_CONNECT_TIMEOUT=${HCCL_CONNECT_TIMEOUT:-5400} exportHCCL_BUFFSIZE=${HCCL_BUFFSIZE:-300} exportTASK_QUEUE_ENABLE=${TASK_QUEUE_ENABLE:-1} exportCOMBINED_ENABLE=${COMBINED_ENABLE:-1} exportTOKENIZERS_PARALLELISM=${TOKENIZERS_PARALLELISM:-false} exportRAY_DEDUP_LOGS=${RAY_DEDUP_LOGS:-0} exportVLLM_ASCEND_ENABLE_NZ=${VLLM_ASCEND_ENABLE_NZ:-0} ;; *) echo"Unsupported DEVICE=${DEVICE}. Expected 'gpu' or 'npu'.">&2 exit1 ;; esac
```

（代码展示了设备自动检测和基础环境配置，但注意 review 意见中建议 proxy unset 和数据集下载应全局化，当前版本仅在 gpu 分支执行。）

评论区精华

- Critical 配置缺失：gemini-code-assist[bot] 指出 actor_rollout_ref.ref.megatron.use_remove_padding=False 和 actor_rollout_ref.ref.megatron.vanilla_mbridge=False 未设置，可能导致 GDN 线性注意力运行时失败或 NPU 初始化崩溃。最终代码未修正。
- 全局化建议：gemini-code-assist[bot] 建议 http_proxy / https_proxy 的 unset 以及 geo3k 数据集下载应移至全局范围，而非仅在 GPU 分支执行；否则 NPU 运行可能因代理或缺少数据出错。
- 公共参数位置：wucong25 指出公共参数（如 TP、PP）不应在 case 内重复定义，应统一在脚本外部设置。最终代码中部分参数仍在 case 内。
- 容器镜像链接：beirong8kmiles 要求添加 NPU 容器镜像链接以方便环境搭建，合并版本是否已添加未确认。
  - Ref 模型 use_remove_padding 缺失 (correctness): 未在最终代码中修复；PR 仍被合并。
  - Ref 模型 vanilla_mbridge 缺失 (correctness): 未在最终代码中修复；PR 仍被合并。
  - 公共参数位置 (design): 最终代码中部分参数仍在 case 内定义，未完全采纳。

风险与影响

- 风险：
  - 脚本中 Ref 模型缺少 use_remove_padding=False，在 Qwen3.5 GDN 架构下可能导致序列打包模式下的运行时错误。
  - 缺少 vanilla_mbridge=False 使 Ref 模型在 NPU 上可能加载错误的 Megatron-Bridge，导致初始化崩溃。
  - 环境变量 http_proxy 和 https_proxy 仅在 GPU 分支 unset，当 NPU 环境存在代理时，多节点通信可能被代理拦截；数据集下载仅限 GPU 分支，NPU 执行前需手动准备数据。
  - 但这些风险仅影响 NPU 运行路径，GPU 路径保持不变。
- 影响：
  - 用户：需要运行 Qwen3.5-122B 在 NPU 上的 RL 训练的用户可以基于此脚本快速配置；但应注意 review 中暴露的未解决问题，可能需额外调试。
  - 团队：新增一条硬件适配路径，但维护需跟进；review 中的问题可作为后续 PR 改善方向。
  - 系统：无核心库改动，仅示例脚本变更，风险可控。
  - 风险标记：Ref 模型配置遗漏，环境变量未全局化，数据集下载仅限 GPU

关联脉络

- PR #6318 [model] feat: add Qwen3.5-35B Ascend support: PR body 明确提及搜索到类似 PR；是 Qwen3.5 系列在 Ascend 上的早期支持，此 PR 在此基础上扩展至 122B 模型。
- PR #6520 [ci] chore: npu ci use cann9.0.0: 同为 NPU 相关基础设施变更，为 NPU 上运行提供 CANN 9.0.0 支持。