

PR #7357 完整报告

verl-project/verl

[ci] test: migrate workflows from fully_async/one_step_off_policy to v1 separate_async

合并时间: 2026-08-12 20:44

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7357>

执行摘要

- 一句话: CI 工作流从 fully_async 迁移到 v1 separate_async
- 推荐动作: 值得阅读, 尤其关注其 CI 迁移方式和参数校验设计。建议合入后重点观察: 1) e2e_v1_separate_async_* 系列工作流首次运行是否全绿; 2) 被删除的 GenRM 覆盖是否需要补回 v1 路径; 3) 后续 fully_async recipe 化是否按此方案推进。

功能与动机

PR body 明确说明目的:

Add v1 separate_async workflow and disable fully_async/one_step_off_policy workflows, preparing to move fully_async into recipe. 即提前将 CI 迁移到 v1 架构, 为 fully_async 功能收编进 recipe 铺路, 避免旧流程持续占用维护成本。

实现拆解

实现拆解如下:

1. 新增 v1 separate_async 测试脚本: 在 tests/special_e2e/ 下新增 run_v1_separate_async.sh、run_v1_separate_async_opd.sh、run_v1_separate_async_ascend.sh, 覆盖 FSDP2 / Megatron 双后端、多教师 OPD、Ascend NPU 场景。脚本统一使用 trainer.use_v1=True + trainer.v1.trainer_mode=separate_async 新配置, 并加入 GPU 分配校验、参数同步步数等显式检查。
2. 新增对应 CI workflow: 新增 .github/workflows/e2e_v1_separate_async.yml、e2e_v1_separate_async_ascend.yml、e2e_v1_separate_async_opd.yml (由 e2e_fully_async_policy_opd.yml 重命名而来), 替换旧的并行异步测试 job, 触发路径指向新脚本。
3. 删除旧工作流和脚本: 移除 e2e_fully_async_policy.yml、e2e_one_step_off_policy.yml 及其 Ascend 变体, 同时删除 run_fully_async_policy.sh、run_one_step_off_policy.sh、run_fully_async_policy_opd.sh、run_fully_async_policy_genrm.sh 等旧 E2E 脚本, 并更新 e2e_ascend.yml 中相关引用。
4. 配套修改: 更新 docs/algo/opd.md 中的 CI 说明, 调整 run_v1_colocate_async_disrm.sh 中少量参数以保持一致性。

整个过程先新增后删除, 保证 CI 在迁移窗口内仍有覆盖。

关键文件：

- tests/special_e2e/run_v1_separate_async.sh（模块 异步测试；类别 test；类型 test-coverage）：新增的 v1 separate_async 主 E2E 脚本，支持 FSDP2/Megatron 双后端，是新 CI 的核心执行入口，包含 GPU 拆分校验和两种策略的参数组装。
- .github/workflows/e2e_v1_separate_async.yml（模块 CI 工作流；类别 infra；类型 infrastructure）：新增的 v1 separate_async 主工作流，定义 GPU runner、安装步骤和 FSDP2/Megatron 两个 job，是 CI 迁移的核心载体。
- tests/special_e2e/run_v1_separate_async_opd.sh（模块 异步蒸馏；类别 test；类型 test-coverage）：新增的多教师在线蒸馏（OPD）E2E 脚本，将原 fully_async OPD 场景迁移到 v1 separate_async，保留 GSM8K/Geo3K 双教师配置。
- .github/workflows/e2e_fully_async_policy.yml（模块 CI 工作流；类别 infra；类型 deletion）：被删除的旧 fully_async 工作流，代表旧 CI 路径移除，是迁移的另一侧。
- tests/special_e2e/run_fully_async_policy.sh（模块 异步测试；类别 test；类型 deletion）：被删除的旧 fully_async E2E 脚本，对应旧 workflow 的执行入口，移除后新脚本接管。

关键符号：未识别

关键源码片段

tests/special_e2e/run_v1_separate_async.sh

新增的 v1 separate_async 主 E2E 脚本，支持 FSDP2/Megatron 双后端，是新 CI 的核心执行入口，包含 GPU 拆分校验和两种策略的参数组装。

```
#!/usr/bin/env bash
set -xeuo pipefail

# Megatron-LM 已打底到 CI 镜像，但 Ray worker 不会继承其 PYTHONPATH，故运行时显式补上
export PYTHONPATH="/workspace/Megatron-LM${PYTHONPATH:+:${PYTHONPATH}}"
export VLLM_USE_V1=1

NUM_GPUS=${NUM_GPUS:-8}
# 默认将 GPU 对半拆给训练与 rollout
N_GPUS_TRAINING=${N_GPUS_TRAINING:-$((NUM_GPUS / 2))}
N_GPUS_ROLLOUT=${N_GPUS_ROLLOUT:-$((NUM_GPUS - N_GPUS_TRAINING))}

# 校验：训练、rollout 的 GPU 数量必须为正，且二者之和等于总数
if ((N_GPUS_TRAINING <= 0 || N_GPUS_ROLLOUT <= 0 || N_GPUS_TRAINING + N_GPUS_ROLLOUT != NUM_GPUS)); then
    echo "Invalid GPU split: total=${NUM_GPUS}, training=${N_GPUS_TRAINING}, rollout=${N_GPUS_ROLLOUT}"
    exit 1
fi

# 公共参数：v1 separate_async 模式的开关与采样器配置
common_params=(
    trainer.use_v1=True
```

```

trainer.v1.trainer_mode=separate_async
trainer.v1.separate_async.num_warmup_batches=1
trainer.v1.separate_async.parameter_sync_step=${PARAMETER_SYNC_STEP:-4}
trainer.v1.sampler.max_off_policy_threshold=8
trainer.v1.sampler.max_off_policy_strategy=drop
transfer_queue.enable=True
# ... 其余数据、算法、rollout 参数从略
)

# 按策略分支: FSDP2 或 Megatron
if [[ "${ACTOR_STRATEGY:-fsdp2}" == "fsdp2" ]]; then
    python3 -m verl.trainer.main_ppo \
        "${common_params[@]}" \
        actor_rollout_ref.actor.strategy=fsdp2 \
        actor_rollout_ref.actor.fsdp_config.fsdp_size=2 \
        "$@"
elif [[ "${ACTOR_STRATEGY}" == "megatron" ]]; then
    TRAIN_TP=${TRAIN_TP:-2}
    # 训练 GPU 数必须能被 TP 大小整除
    if ((TRAIN_TP <= 0 || N_GPUS_TRAINING % TRAIN_TP != 0)); then
        echo "N_GPUS_TRAINING must be divisible by TRAIN_TP"
        exit 1
    fi
    TRAIN_PP=$((N_GPUS_TRAINING / TRAIN_TP))
    python3 -m verl.trainer.main_ppo \
        model_engine=megatron \
        "${common_params[@]}" \
        actor_rollout_ref.actor.megatron.tensor_model_parallel_size=${TRAIN_TP} \
        actor_rollout_ref.actor.megatron.pipeline_model_parallel_size=${TRAIN_PP} \
        "$@"
else
    echo "Unknown ACTOR_STRATEGY=${ACTOR_STRATEGY}"
    exit 1
fi

```

评论区精华

该 PR 没有产生任何 review 评论，仅由维护者 wuxibin89 直接批准通过。因此没有可提炼的口头讨论或设计权衡交锋。

- 暂无高价值评论线程

风险与影响

- 风险：风险主要集中在 CI 覆盖切换：
- 回归风险：旧 `fully_async` / `one_step_off_policy` 路径的 CI 被整体移除，若 v1 `separate_async` 尚未完全等能力覆盖（如 GenRM 场景未在 v1 脚本中保留），可能出现异步训练回归未被 CI 捕获。`run_fully_async_policy_genrm.sh` 被直接删除，v1 侧没有对

等脚本。

- 新脚本健壮性：新增脚本中做了较严格的 GPU 数量和整除性校验（如 `run_v1_separate_async_opd.sh`），但在不同 CI runner 配置下可能因默认值（如 `NUM_GPUS=8`）与实际环境不符而失败。
- Ascend 兼容性：`e2e_v1_separate_async_ascend.yml` 为全新编写，依赖的容器镜像和 NPU 环境未经过长时间验证，存在首次运行失败的风险。

整体为 CI 层变更，不涉及训练核心代码路径，但需观察合入后几个 workflow 的实际运行结果。

- 影响：影响范围限于 CI 基础设施和端到端测试脚本：
- CI 覆盖变化：`fully_async_policy` 与 `one_step_off_policy` 的 nightly / PR CI 不再运行，替换为 `v1separate_async` 系列；开发者在修改异步训练相关代码时，CI 信号来源发生变化。
- 团队协作：维护者需按新工作流配置本地复现，旧脚本从仓库消失，历史 record 可能失效。
- 项目演进：这是 `fully_async` 功能向正式 v1 路径收敛的明确一步，后续 recipe 化将以此为基座。

对用户（训练框架使用者）无直接影响，对内部 CI 消费方影响较大。

- 风险标记：旧 CI 已删除，新工作流未充分验证，GenRM 覆盖丢失，Ascend 环境风险

关联脉络

- PR #7306 [fsdp] fix: add FSDP1 restore handler for Decoupled PPO: 同属 `fully_async` 功能线，修复异步训练恢复问题，本 PR 迁移其 CI 覆盖，后续应关注 v1 路径下该修复的回归测试。
- PR #7345 [ci] chore: Fix npu nightly ci: 同为 CI 工作流维护，本 PR 新增的 Ascend v1 工作流与其共享 NPU runner 环境，存在配置关联。