

# PR #7003 完整报告

verl-project/verl

[megatron, trainer] feat: add HY v3 Megatron GRPO example

合并时间: 2026-07-10 12:50

原文链接: <http://prhub.com.cn/verl-project/verl/pull/7003>

## 执行摘要

- 一句话: 添加 HY V3 Megatron GRPO 训练示例与 FLOPs 估算
- 推荐动作: 该 PR 提供了实用的模型支持, 值得用户参考。但应留意 use\_mbridge 的过时间题, 建议尽快为 FLOPs 估算函数补充单元测试, 并在后续迭代中迁移至原生 Megatron-Bridge 配置。

## 功能与动机

HY V3 是一种 MoE 模型, 需要专项的 GRPO 训练脚本和 FLOPs 估算支持, 以降低用户使用门槛。PR 旨在提供开箱即用方案, 并复用 Megatron-Bridge 桥接。

## 实现拆解

1. 在 verl/utils/flops\_counter.py 中新增 \_estimate\_hy\_v3\_flops 函数, 计算包含 MoE 路由、线性层及注意力机制的 FLOPs, 并通过 ESTIMATE\_FUNC 字典注册 hy\_v3 类型。
2. 新建 examples/grpo\_trainer/run\_hy\_v3\_megatron.sh 示例脚本, 提供完整的 DAPO 风格 GRPO 训练配置, 支持 vLLM / SGLang / TRTLLM rollout 后端, 可调节 TP/PP/CP/EP 等分布式参数。
3. 注意脚本中使用了 use\_mbridge=1 选项以启用 Megatron-Bridge, 该配置在后续版本中计划废弃。

关键文件:

- verl/utils/flops\_counter.py (模块 性能计算; 类别 source; 类型 core-logic; 符号 \_estimate\_hy\_v3\_flops): 核心修改: 新增 HY V3 专用 FLOPs 估算函数, 是 PR 的技术核心。
- examples/grpo\_trainer/run\_hy\_v3\_megatron.sh (模块 训练示例; 类别 other; 类型 example): 用户直接使用的训练脚本, 包含完整配置参数, 是 PR 的主要交付物。

关键符号: \_estimate\_hy\_v3\_flops

## 关键源码片段

verl/utils/flops\_counter.py

核心修改: 新增 HY V3 专用 FLOPs 估算函数, 是 PR 的技术核心。

```
def _estimate_hy_v3_flops(config, tokens_sum, batch_seqlens, delta_time):
```

```

hidden_size = config.hidden_size
vocab_size = config.vocab_size
# Hy V3 使用 expert_hidden_dim 作为 moe_intermediate_size 的遗留别名。
# 优先使用标准名称并回退。
moe_intermediate_size = getattr(config, "moe_intermediate_size", None)
if moe_intermediate_size is None:
    moe_intermediate_size = config.expert_hidden_dim
num_hidden_layers = config.num_hidden_layers
num_experts = config.num_experts
moe_topk = config.num_experts_per_tok
share_expert_num = config.num_shared_experts

num_attention_heads = config.num_attention_heads
num_key_value_heads = config.num_key_value_heads
head_dim = config.head_dim
q_size = num_attention_heads * head_dim
k_size = num_key_value_heads * head_dim
v_size = num_key_value_heads * head_dim

# 每层非注意力部分：路由门控 + MoE 专家 (topk 路由 + 共享专家)
moe_gata_N = hidden_size * num_experts
moe_expertmlp_N = hidden_size * moe_intermediate_size * (moe_topk + share_expert_num) *
3
moe_mlp_N = moe_gata_N + moe_expertmlp_N

attn_linear_N = hidden_size * (q_size + k_size + v_size + num_attention_heads * head_dim)
emd_and_lm_head_N = vocab_size * hidden_size * 2

dense_N = (moe_mlp_N + attn_linear_N) * num_hidden_layers + emd_and_lm_head_N
dense_N_flops = 6 * dense_N * tokens_sum

seqlen_square_sum = 0
for seqlen in batch_seqLens:
    seqlen_square_sum += seqlen * seqlen
attn_qkv_flops = 6 * seqlen_square_sum * head_dim * num_attention_heads * num_hidden_
layers

flops_all_token = dense_N_flops + attn_qkv_flops
flops_achieved = flops_all_token * (1.0 / delta_time) / 1e12
return flops_achieved

ESTIMATE_FUNC = {
    # ... 已有映射 ...
    "hy_v3": _estimate_hy_v3_flops, # 新增注册
}

```

## 评论区精华

Review 中 gemini-code-assist 两次指出脚本引用了已废弃的 `use_mbridge` 配置选项（第 156、209 行），建议迁移到纯 Megatron-Bridge。但 reviewer 未要求修改即直接合并，该问题未被解决。

- 使用已废弃的 `use_mbridge` 配置 (design): reviewer 未要求修改即直接合并，该问题未在本次 PR 中解决。

## 风险与影响

- 风险：
  1. 配置风险: `use_mbridge` 已废弃，后续版本可能移除，导致示例无法运行。
  2. 兼容性风险: FLOPs 估算函数依赖模型属性（如 `expert_hidden_dim`），若配置字段变化可能估算不准确。
  3. 缺少测试: 新增函数未附带单元测试，无法验证正确性。
  4. 规模风险: 示例使用 16 节点 128 GPU 的默认参数，小规模用户需大幅调整。 - 影响：  
用户: HY V3 模型用户可立即使用该示例进行 GRPO 训练，降低集成成本。系统: FlopsCounter 增加新模型支持，性能监控更全面。团队: 需后续跟进 `use_mbridge` 的迁移。影响范围限于 Megatron + MoE 训练路径，影响程度中等。 - 风险标记: 使用已废弃的 `use_mbridge` 配置，缺少测试覆盖

## 关联脉络

- PR #6951 [BREAKING][megatron] chore: deprecate vanilla mBridge and make Megatron Bridge default: 本 PR 使用的 `use_mbridge` 配置恰好是 PR #6951 计划废弃的，两者直接相关。