PR #7014 完整报告

verl-project/verl

[fsdp] fix: sync merged LoRA weights before context exit

合并时间：2026-07-15 20:32

原文链接：http://prhub.com.cn/verl-project/verl/pull/7014

执行摘要

- 一句话：修复 LoRA merge 时权重同步的陈旧权重 bug
- 推荐动作：推荐精读。此 PR 展示了在 PyTorch FSDP 环境中处理上下文管理器和张量生命周期的典型陷阱，_merged_lora_per_tensor_param 的设计模式（生成器 + try/finally 确保 offload）值得在类似同步场景中复用。

功能与动机

修复 Issue #6782 中描述的训练异常：training/rollout_probs_diff_mean 持续增大，rollout_actor_probs_pearson_corr 逐渐下降，奖励不提升。根本原因是 LoRA 合并权重同步时，因 context 生命周期与张量消费错位，vLLM 获得了未合并的基础权重。

实现拆解

1. 新增 _merged_lora_per_tensor_param 生成器方法：在 verl/workers/engine/fsdp/transformer_impl.py 中新增该方法，在 merged_lora_context 内部调用了 normalize_peft_param_name 和 convert_weight_keys，并对每个参数立即 yield，包括对 DTensor 调用 full_tensor()，对普通张量调用 detach().clone()，防止 context 退出后引用失效。
2. 修改 get_per_tensor_param 的分支逻辑：当检测到 LoRA 合并路径时，直接 return self._merged_lora_per_tensor_param(), None，跳过后续的通用参数处理和 QAT 逻辑，因为团队确定不支持 LoRA + QAT 组合。
3. 恢复非 LoRA 路径的 FSDP2 staging 跳过逻辑：在 reviewer 指出之前，作者曾误移除了 PR #7005 中引入的 _skip_staging 逻辑，后修复为保留它，并且移除了一个无关的 meta_info.pop("model_output", None) 删除。

关键文件：

- verl/workers/engine/fsdp/transformer_impl.py（模块 引擎层；类别 source；类型 core-logic；符号 _merged_lora_per_tensor_param, get_per_tensor_param）：核心修改文件，新增 _merged_lora_per_tensor_param 方法并调整 get_per_tensor_param 分支，修复了 LoRA 合并权重同步的陈旧权重 bug。

关键符号：_merged_lora_per_tensor_param, get_per_tensor_param

关键源码片段

verl/workers/engine/fsdp/transformer_impl.py

核心修改文件，新增 _merged_lora_per_tensor_param 方法并调整 get_per_tensor_param 分支，修复了 LoRA 合并权重同步的陈旧权重 bug。

```
def _merged_lora_per_tensor_param(self):
    """Stream merged (base + LoRA) weights for rollout weight sync.

    ``state_dict()`` returns tensors that alias the live FSDP parameter
    storage, and ``merged_lora_context`` restores the un-merged base
    weights when it exits. The context therefore must stay open until the
    consumer has materialized every tensor. ``DTensor.full_tensor()``
    produces a copy, so yielded tensors remain valid after the restore.
    """
    device = get_device_id()
    try:
        # Keep context open while we yield; after exit weights revert to base.
        with merged_lora_context(self.module, backup_adapters=True):
            params = normalize_peft_param_name(self.module.state_dict())
            params = convert_weight_keys(params,
                                         getattr(self.module, "_fsdp_wrapped_module", self.module))
            for name, param in params.items():
                yield (
                    name,
                    param.to(device, non_blocking=True).full_tensor()
                        .to(torch.bfloat16, non_blocking=True)
                    if isinstance(param, DTensor)
                    # Clone plain tensors to avoid aliasing restored storage.
                    else param.detach().clone(),
                )
    finally:
        # Offload only after consumer has fully consumed the generator.
        log_gpu_memory_usage("Before offload_fsdp_model_to_cpu", logger=logger)
        if self._is_offload_param:
            offload_fsdp_model_to_cpu(self.module)
        log_gpu_memory_usage("After offload_fsdp_model_to_cpu", logger=logger)
```

评论区精华

Reviewer wuxibin89 指出了两个问题：

- 误移除了 FSDP2 的 staging 跳过逻辑（来自 PR #7005），作者在后续 commit 中恢复。
- 删除 meta_info.pop("model_output", None) 可能影响训练，作者确认后恢复。
- 关于早期返回 self._merged_lora_per_tensor_param() 会跳过 QAT 的问题，reviewer 明确表示不支持 LoRA+QAT，作者据此简化，最终版本直接早期返回。
- FSDP2 staging 跳过逻辑误删 (correctness): 作者 rongkunxue 在 ee5d95b7 中恢复被误删的逻辑。
- meta_info.pop("model_output", None) 误删 (correctness): 作者恢复该行。

- 早期返回与 QAT 兼容性 (design): wuxibin89 表示不支持 LoRA+QAT，因此作者简化设计，直接早期返回。

风险与影响

- 风险：风险较低。变更只影响 actor_rollout_ref.model.lora.merge=True 路径，该路径之前有 bug，修复后更可靠。未合并 LoRA（merge=False）路径不受影响。因为明确不支持 LoRA + QAT，早期返回排除了 QAT 处理，但不会导致错误。
- 影响：影响范围：仅影响使用 LoRA 训练且 merge=True 的用户。这部分用户的训练将恢复正常，之前可能因权重滞后导致奖励不提升。无 API 变更或配置变更。
- 风险标记：核心路径变更，缺少 LoRA+QAT 兼容确认

关联脉络

- PR #7005 [fsdp] fix: skip the whole-shard staging round trip in FSDP2 weight export: 修改了同一文件 transformer_impl.py，PR 中 reviewer 指出误删了来自 #7005 的 staging 跳过逻辑，后恢复。