

PR #6604 完整报告

verl-project/verl

[fsdp] fix: do not manually move model to GPU

合并时间: 2026-06-05 13:30

原文链接: <http://prhub.com.cn/verl-project/verl/pull/6604>

执行摘要

- 一句话: 禁止 FSDP2 保存检查点时手动移动模型到 GPU
- 推荐动作: 值得精读。该 PR 展示了 FSDP1 与 FSDP2 在参数卸载 / 加载上的根本差异, 以及如何通过增加一个属性判断来安全处理分支。对于使用 FSDP2 进行全权重训练的团队, 此修复关键。

功能与动机

根据 Issue #5995, FSDP2 启用 CPUOffloadPolicy 后, `save_checkpoint` 中的 `load_fsdp_model_to_gpu` 无法真正移动参数 (FSDP2 内部管理), 后续 `state_dict()` 会因设备不匹配而崩溃。需要在条件中跳过手动移动。

实现拆解

1. 在 `verl/workers/engine/fsdp/transformer_impl.py` 的 `save_checkpoint` 方法中, 修改条件判断。
2. 原条件 `if self._is_offload_param or origin_module_device == "cpu"` 会无条件触发 `load_fsdp_model_to_gpu`。
3. 新条件添加 `and not getattr(self, '_uses_fsdp2_cpu_offload_policy', False)`, 当 FSDP2 CPU offload 策略激活时跳过手动移动。
4. 使用 `getattr` 安全访问属性, 避免子类未初始化该属性时抛出 `AttributeError`。

关键文件:

- `verl/workers/engine/fsdp/transformer_impl.py` (模块 调度器; 类别 source; 类型 core-logic; 符号 `save_checkpoint`): 核心变更文件, 修改了 `save_checkpoint` 方法中的条件判断, 修复 FSDP2 CPU offload 下的保存崩溃。

关键符号: `save_checkpoint`

关键源码片段

`verl/workers/engine/fsdp/transformer_impl.py`

核心变更文件, 修改了 `save_checkpoint` 方法中的条件判断, 修复 FSDP2 CPU offload 下的保存崩溃。

```
def save_checkpoint(self, local_path, hdfs_path=None, global_step=0, max_ckpt_to_keep=None,
```

```

**kwargs):
    """
    Save FSDP checkpoint, handling parameter offload as needed.
    """

    origin_module_device = next(self.module.parameters()).device.type
    # 在 FSDP2 使用 CPUOffloadPolicy 时，不要手动将模型移动到 GPU，
    # 因为 FSDP2 内部管理参数位置，手动移动会导致后续 state_dict() 报错
    if (self._is_offload_param or origin_module_device == "cpu") and not getattr(
        self, "_uses_fsdp2_cpu_offload_policy", False
    ):
        load_fsdp_model_to_gpu(self.module)

    self.checkpoint_manager.save_checkpoint(
        local_path=local_path, hdfs_path=hdfs_path,
        global_step=global_step, max_ckpt_to_keep=max_ckpt_to_keep
    )

    torch.distributed.barrier()
    if self._is_offload_param:
        offload_fsdp_model_to_cpu(self.module)

```

评论区精华

gemini-code-assist[bot] 建议使用 `getattr(self, '_uses_fsdp2_cpu_offload_policy', False)` 替代直接访问 `self._uses_fsdp2_cpu_offload_policy`，防止子类（如 VeOmniEngine）未初始化该属性导致 `AttributeError`。该建议被采纳并体现在最终提交中。

- 使用 `getattr` 安全访问属性以避免子类 `AttributeError (correctness)`: 作者采纳建议，最终提交中使用 `getattr` 方式。

风险与影响

- 风险：风险较低。修改仅增加一个额外条件判断，仅在 FSDP2 CPU offload policy 启用时改变行为。已有 `_uses_fsdp2_cpu_offload_policy` 属性在 `FSDPEngine.__init__` 中根据配置设置，默认 `False`。需确保子类或自定义实现正确设置该属性，否则可能导致错误跳过。
- 影响：直接影响 FSDP2 + CPUOffloadPolicy 训练场景的检查点保存，修复了更新权重时的设备不匹配崩溃。不影响不使用 CPUOffloadPolicy 或使用 FSDP1 的情况。变更极小（+3/-1），无 API 或配置变化。
- 风险标记：子类属性未定义

关联脉络

- PR #6463 Similar fix for offload issue: PR body 中提及，与 #6463 相似修复，处理参数卸载相关崩溃。