

PR #29543 完整报告

sgl-project/sglang

Fix DP-attention SHM feature finalization race

合并时间: 2026-06-28 12:36

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/29543>

执行摘要

- 一句话: 修复 DP-attention 下 SHM 内存的竞态条件
- 推荐动作: 该 PR 修复了一个明确的并发竞态条件, 逻辑清晰, 改动量小, 可以审阅后合并。
值得关注的是其对 DP-attention 路径的同步策略: 使用 `attn_tp` 和 `attn_cp` 分组 barrier 而非全 tp 组 barrier, 避免了不必要的同步开销, 是一个好的设计选择。

功能与动机

在 DP-attention 模式下, 广播源 rank 返回原始对象时, 其他 rank 可能仍在反序列化 `ShmPointerMMData` (调用 `shm_open`)。之前仅对非 DP-attention 路径做了同步, 导致 DP-attention 下 source rank 提前调用 `materialize()/shm_unlink`, 而其他 rank 还没完成 `shm_open`, 产生竞态条件。

实现拆解

1. 在 `_finalize_shm_features` 方法中, 移除外层条件中对 `enable_dp_attention` 的排除判断, 改为统一的 `self.model_config.is_multimodal and has_shm_features(recv_reqs)` 入口。
2. 在内部分支中根据 `enable_dp_attention` 分别处理: 若启用 DP-attention, 则根据 `attn_tp_size` 和 `attn_cp_size` 在对应的 CPU 组上执行 barrier; 否则回退到原 `tp_cpu_group barrier`。
3. 清理了过时的注释和冗余的逻辑判断, 使同步屏障覆盖所有 SHM 承载的通信路径。

关键文件:

- `python/sglang/srt/managers/scheduler_components/request_receiver.py` (模块 调度器; 类别 source; 类型 core-logic; 符号 `_finalize_shm_features`): 唯一变更文件, 修改了 `_finalize_shm_features` 方法, 调整了同步屏障的覆盖范围, 将 DP-attention 路径纳入同步保护。

关键符号: `_finalize_shm_features`

关键源码片段

`python/sglang/srt/managers/scheduler_components/request_receiver.py`

唯一变更文件, 修改了 `_finalize_shm_features` 方法, 调整了同步屏障的覆盖范围, 将 DP-attention 路径纳入同步保护。

```

def _finalize_shm_features(self, recv_reqs: Optional[List]) -> None:
    # Unwrap shared memory features AFTER all broadcasts complete,
    # so that ShmPointerMMData metadata (not full tensor data) is what
    # gets serialized during broadcast_pyobj.
    if recv_reqs:
        if self.model_config.is_multimodal and has_shm_features(recv_reqs):
            # 广播源 rank 返回原始对象时，其他 rank 可能仍在反序列化
            # ShmPointerMMData（调用 shm_open）。在 materialize() 解除
            # 映射之前，需要在承载 SHM 工作请求的同一 CPU 组上同步。
            if self.server_args.enable_dp_attention:
                # DP-attention 路径：在 attn_tp 和 attn_cp 组上分别同步，
                # 这些组是实际传递 SHM 指针的通信组。
                if self.ps.attn_tp_size > 1:
                    barrier(group=self.attn_tp_cpu_group)
                if self.ps.attn_cp_size > 1:
                    barrier(group=self.attn_cp_cpu_group)
            elif self.ps.tp_size > 1:
                # 非 DP-attention 路径：原有 tp_cpu_group 同步保持不变。
                barrier(group=self.tp_cpu_group)
        for req in recv_reqs:
            unwrap_shm_features(req)

```

评论区精华

PR 没有 review 评论或审核记录，仅由作者 merrymercy 直接合并。

- 暂无高价值评论线程

风险与影响

- 风险：变更仅涉及同一个方法内的控制流调整，范围很小。风险较低，但需要确认 DP-attention 路径下 attn_tp 和 attn_cp CPU 组的正确初始化，以及 barrier 调用不会引入死锁。由于没有新增测试文件，回归风险依赖现有 CI。
- 影响：仅影响启用 DP-attention 且使用 SHM 特征的 multimodal 模型路径。修复了潜在的竞态崩溃，提升了 DP-attention 在共享内存场景下的稳定性。非 DP-attention 路径无行为变化。
- 风险标记：缺少测试覆盖

关联脉络

- 暂无明显关联 PR