

PR #25356 完整报告

sgl-project/sglang

[AMD] test(sgl-kernel): seed RNG on ROCm in test_moe_topk_sigmoid to fix tie-break flake

合并时间: 2026-05-20 14:01

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/25356>

执行摘要

- 一句话: 为 AMD 测试添加确定性种子修复偶发失败
- 推荐动作: 建议精读。这是一个小而精的测试修复案例, 展示了如何通过最小侵入性修改 (单文件 +8 行) 解决平台相关的随机性问题, 值得测试工程师参考。

功能与动机

`test_moe_topk_sigmoid` 在 AMD CI (ROCm MI30x/MI35x) 上因 `atol=0` 的索引比较而间歇性失败, 原因是 `torch.randn` 的未种子化输入导致接近平局的分数, 而 hipCUB 的 `torch.topk tie-break` 与 `sgl_kernel.topk_sigmoid` 不同。该 bug 仅影响 AMD 平台。

实现拆解

1. 添加 HIP 检测标志: 在文件顶部定义 `_IS_HIP = torch.version.hip is not None`, 作为平台条件。
2. 创建确定性种子 fixture: 新增一个 `@pytest.fixture(autouse=True)` 的 `_deterministic_seed` 函数, 当 `_IS_HIP` 为真时调用 `torch.manual_seed(0)`, 确保每个测试用例的输入可重现。
3. 保持 CUDA 不变: 该 fixture 在非 ROCm 平台上无操作, 因此 CUDA CI 的测试用例仍然使用未固定的随机输入, 不降低 CUDA 覆盖质量。

关键文件:

- `sgl-kernel/tests/test_moe_topk_sigmoid.py` (模块测试; 类别 test; 类型 test-coverage; 符号 `_deterministic_seed`): 唯一被修改的文件, 添加了仅在 ROCm 上生效的确定性种子 fixture。

关键符号: `_deterministic_seed`

关键源码片段

`sgl-kernel/tests/test_moe_topk_sigmoid.py`

唯一被修改的文件, 添加了仅在 ROCm 上生效的确定性种子 fixture。

```
import itertools
import sys

import pytest
```

```
import torch
from sgl_kernel import topk_sigmoid

# 仅在 ROCm 平台上启用固定种子, 避免 hipCUB 与 sgl_kernel 在 tie-break 上的差异导致偶发失败
_IS_HIP = torch.version.hip is not None

@pytest.fixture(autouse=True)
def _deterministic_seed():
    # AMD/ROCM only: pin RNG so torch.randn produces identical gating scores
    # across runs. atol=0 indices comparison is otherwise tripped by near-tied
    # sigmoid scores where hipCUB's tie-break inside torch.topk disagrees with
    # sgl_kernel.topk_sigmoid. Not observed on CUDA, so leave CUDA behavior
    # unchanged.
    if _IS_HIP:
        torch.manual_seed(0)
```

评论区精华

无 review 评论。PR 由 HaiShaw 批准。

- 暂无高价值评论线程

风险与影响

- 风险：无显著风险。仅修改了测试文件，生产代码未受影响。在 AMD 上固定种子可能会掩盖输入依赖的 bug，但该测试的目的本就是验证索引正确性，而非随机性。
- 影响：影响范围极小，仅限于 sgl-kernel/tests/test_moe_topk_sigmoid.py 中的四个参数化测试。AMD CI 的稳定性得到改善，CUDA CI 不受影响。
- 风险标记：测试平台特定

关联脉络

- 暂无明显关联 PR