

PR #43583 完整报告

vllm-project/vllm

[Misc] Print accuracy value for PD tests even on success

合并时间: 2026-05-25 17:10

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/43583>

执行摘要

- 一句话: 精度测试始终打印测量值
- 推荐动作: 值得合并, 对调试高方差测试场景有实际帮助, 无副作用。

功能与动机

该 PR 响应 issue 评论, 旨在解决分布式模式下多种模型组合运行测试时精度方差高于预期的问题。通过始终打印测量值, 可在断言触发前追踪其变化。

实现拆解

在 `tests/v1/kv_connector/nixl_integration/test_accuracy.py` 的 `test_accuracy` 函数中, 将原先仅在 `expected_value is None` 分支内打印 `measured_value` 的 `print(f"Measured value: {measured_value}")` 移到分支外部, 改为无条件执行 `print(f"Measured accuracy value: {measured_value}\n")`, 确保每次运行都记录精度值。

关键文件:

- `tests/v1/kv_connector/nixl_integration/test_accuracy.py` (模块 测试; 类别 test; 类型 test-coverage): 唯一变更文件, 将测量精度打印移到断言前, 确保始终输出。

关键符号: 未识别

关键源码片段

`tests/v1/kv_connector/nixl_integration/test_accuracy.py`

唯一变更文件, 将测量精度打印移到断言前, 确保始终输出。

```
# tests/v1/kv_connector/nixl_integration/test_accuracy.py

measured_value = results["results"][TASK][FILTER]
expected_value = EXPECTED_VALUES.get(MODEL_NAME)

# 始终打印测量精度值, 便于追踪跨运行变化
print(f"Measured accuracy value: {measured_value}\n")

if expected_value is None:
    print(
```

```
        f"Warning: No expected value found for {MODEL_NAME}. "  
        "Skipping accuracy check."  
    )  
    return  
  
    assert (  
        measured_value - RTOL < expected_value  
        and measured_value + RTOL > expected_value  
    ), f"Expected: {expected_value} | Measured: {measured_value}"
```

评论区精华

无 reviewer 讨论。

- 暂无高价值评论线程

风险与影响

- 风险：极低风险，仅修改了测试中的日志打印逻辑，不影响测试逻辑或系统行为。
- 影响：仅影响精度集成测试的输出，使其始终打印测量值，便于开发者在 CI 或本地调试时诊断精度波动。
- 风险标记：暂无

关联脉络

- PR #43186 [KVConnector] fix: Remove spurious copy in test: PR body 引用该 PR 的 issue 评论，说明变更动机。