

PR #41072 完整报告

vllm-project/vllm

[CI][AMD][BugFix] Patch has_flashinfer decorator for test_select_rocm_aiter_backend

合并时间: 2026-04-29 10:12

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/41072>

执行摘要

- 一句话: 修复 has_flashinfer mock 路径, 使 ROCm 测试通过
- 推荐动作: 无特殊关注必要。此 PR 是典型的后续维护修复, 可快速合并。

功能与动机

has_flashinfer 函数已从 `vllm.model_executor.layers.fused_moe.oracle.unquantized` 移至 `vllm.utils.flashinfer.has_flashinfer`, 但测试中的 `@patch` 装饰器未更新, 导致 `test_select_rocm_aiter_backend` 测试失败。PR 目的就是修复此路径不一致问题, 使测试重新通过。

实现拆解

仅修改 `tests/kernels/moe/test_unquantized_backend_selection.py` 第 88 行, 将 `@patch` 的路径参数从 `"vllm.model_executor.layers.fused_moe.oracle.unquantized.has_flashinfer"` 改为 `"vllm.utils.flashinfer.has_flashinfer"`, 以匹配 `has_flashinfer` 函数的最新位置。

关键文件:

- `tests/kernels/moe/test_unquantized_backend_selection.py` (模块测试; 类别 test; 类型 test-coverage): 唯一变更文件, 修复 mock 路径以匹配 `has_flashinfer` 函数的新位置

关键符号: `test_select_rocm_aiter_backend`

关键源码片段

`tests/kernels/moe/test_unquantized_backend_selection.py`

唯一变更文件, 修复 mock 路径以匹配 `has_flashinfer` 函数的新位置

```
# tests/kernels/moe/test_unquantized_backend_selection.py 中修复后的 mock 装饰器
# 将路径从旧地址更新为 has_flashinfer 函数的新位置
@patch(
    "vllm.utils.flashinfer.has_flashinfer", # 修复前: "vllm.model_executor.layers.fused_moe.oracle.
    unquantized.has_flashinfer"
    return_value=False,
)
@patch(
    "vllm.model_executor.layers.fused_moe.oracle.unquantized.rocm_aiter_ops.is_fused_moe_
```

```
    enabled",
    return_value=True,
)
@pytest.mark.skipif(
    not current_platform.is_rocm(), reason="ROCm-specific backend selection test"
)
def test_select_rocm_aiter_backend(mock_aiter_enabled, mock_has_flashinfer):
    """Test ROCm backend selection when AITER is available."""
    # 测试主体未变化
```

评论区精华

无实际 review 讨论; [gemini-code-assist\[bot\]](#) 和 [claude\[bot\]](#) 的评论均无实质性反馈, [yewentao256](#) 和 [tjtanaa](#) 均直接批准。

- 暂无高价值评论线程

风险与影响

- 风险: 风险极低。只修改测试文件中的 mock 路径, 与被 mock 的生产代码逻辑无关。唯一风险是该路径在未来再次移动时未同步更新, 但属测试维护范畴。
- 影响: 仅影响 ROCm 环境下的 test_select_rocm_aiter_backend 测试用例。修复后该测试可通过, 使得 CI 中 ROCm 相关测试更可靠。对生产代码无影响。
- 风险标记: 低风险 - 仅测试变更

关联脉络

- 暂无明显关联 PR