

PR #44146 完整报告

vllm-project/vllm

[XPU][CI] Fix test_audio_in_video flake by using module-scoped server fixture

合并时间: 2026-06-01 19:21

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/44146>

执行摘要

- 一句话: 修复 XPU 上测试音频视频的间歇失败
- 推荐动作: 值得合并。该 PR 修复了 XPU 特定的 CI flake, 变更简单且经过审核。对于维护者, 建议关注测试 fixture 作用域与隔离性的权衡, 但当前方案是合理的。无需深入精读。

功能与动机

修复 XPU CI 中 `test_online_audio_in_video_multi_videos` 测试的间歇性失败:

`RuntimeError: Server failed to start in time`。在 XPU 上, 每个测试启动新服务器会导致第二个服务器静默挂起并超时。

实现拆解

1. 修改 `tests/entrypoints/openai/chat_completion/test_audio_in_video.py` 中 `server fixture` 的 `scope` 参数: 从默认的 `function` 改为 `"module"`, 使该文件内的所有测试共享同一个服务器实例。
2. 在 `fixture` 定义前添加注释说明改动机: XPU 上每个测试启动新服务器会导致静默挂起和超时, 模块作用域可避免此问题。
3. 这是唯一一处变更, 仅涉及测试代码, 无源码或配置修改。

关键文件:

- `tests/entrypoints/openai/chat_completion/test_audio_in_video.py` (模块测试; 类别 `test`; 类型 `test-coverage`): 唯一变更文件, 将 `server fixture` 的 `scope` 从 `function` 改为 `module`, 并添加注释解释动机。

关键符号: 未识别

关键源码片段

`tests/entrypoints/openai/chat_completion/test_audio_in_video.py`

唯一变更文件, 将 `server fixture` 的 `scope` 从 `function` 改为 `module`, 并添加注释解释动机。

```
@pytest.fixture(scope="module") # 原为默认 function 作用域
# 使用 module 作用域, 使文件内所有测试共享一个服务器实例。
# 在 XPU 上, 每个测试启动新服务器会导致第二个服务器静默挂起
# 并超过等待超时, 引发 RuntimeError。
```

```
def server():
    args = [
        "--max-model-len",
        "16384",
        "--enforce-eager",
        "--limit-mm-per-prompt",
        json.dumps({"audio": 3, "video": 3}),
        *ROCM_EXTRA_ARGS,
    ]
    with RemoteOpenAIServer(MODEL_NAME, args) as remote_server:
        yield remote_server
```

评论区精华

审核人 AndreasKaratzas 评论: “Looks good but reduces a bit test isolation.” (减少了测试隔离性)。合并者 DarkLight1337 批准了 PR, 未进一步讨论。这是一个已知的 trade-off: 模块级 fixture 减少了隔离性, 但避免了 XPU 特有的服务器启动挂起问题。

- 测试隔离性降低 (testing): 合并者已知晓该 trade-off, 并批准了 PR。未进一步讨论。

风险与影响

- 风险: 风险较低。变更仅影响测试 fixture 的作用域, 将函数级改为模块级。主要风险是测试隔离性降低: 一个测试中的副作用 (如模型状态污染) 可能影响同一模块中的其他测试。但该测试文件原本只有一个测试函数, 且测试设计是独立的, 因此实际影响很小。此外, 如果 XPU 环境的变更解除了该挂起问题, 此 workaround 可能导致后续测试难以发现回归, 但当前这是必要措施。
- 影响: 影响范围限于 XPU CI 上的 test_audio_in_video.py 测试文件。该变更消除了该测试的间歇性失败, 使 XPU CI 更可靠。对非 XPU 平台无影响。对其他测试或生产代码无影响。
- 风险标记: 测试隔离性降低, XPU 平台特定

关联脉络

- 暂无明显关联 PR