

PR #48738 完整报告

vllm-project/vllm

[Rust Frontend] Fix mock engine test shutdown race

合并时间: 2026-07-16 15:27

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/48738>

执行摘要

该 PR 通过交换 `mock-engine/src/tests.rs` 中 `shutdown_mock` 函数的两行代码，修复了 mock engine 测试因关闭顺序竞态导致的 `Connection reset by peer` 错误。改动仅 1 行删除、1 行添加，但根治了 CI 中偶现的测试失败。

功能与动机

在 Buildkite CI 中，`abort_cancels_active_request_and_emits_terminal_output` 测试偶尔因 `Connection reset by peer` 失败。根因是 `shutdown_mock` 先关闭客户端传输，再取消 mock engine；若 mock engine 在处理客户端关闭时尚未收到取消信号，会将连接断开视为意外 I/O 错误。修复后先通过 `CancellationToken` 通知 mock engine 进入关闭流程，再关闭客户端 socket。

实现拆解

1. 变更入口：文件 `rust/src/mock-engine/src/tests.rs`，仅修改 `shutdown_mock` 辅助函数。
2. 核心逻辑调整：将 `shutdown.cancel()` 从第 94 行移至第 93 行，`client.shutdown().await` 相应后移。
3. 影响：所有使用 `shutdown_mock` 的测试（`mock_engine_connects_over_tcp`、`mock_engine_connects_over_ipc`、`mock_engine_registers_multiple_identities` 等）均受益，关闭路径更可靠。
4. 测试配套：无额外测试，现有测试即覆盖该修复场景。

`rust/src/mock-engine/src/tests.rs`

核心变更文件，调整 `shutdown_mock` 中取消信号与客户端关闭的顺序。

```
// 测试关闭辅助函数：先取消 mock engine，再关闭客户端传输
async fn shutdown_mock(
    client: EngineCoreClient,
    shutdown: CancellationToken,
    task: tokio::task::JoinHandle<Result<>>>,
) {
    // 先发送取消信号，让 mock engine 走正常关闭路径
    shutdown.cancel();
    // 再关闭客户端传输，避免 mock engine 将传输关闭误判为意外 I/O 错误
    client.shutdown().await.expect("client shutdown");
    task.await.expect("mock join").expect("mock run");
}
```

```
}
```

评论区精华

无实质讨论，[BugenZhao](#) 直接审批并表示感谢。

风险与影响

- 风险：极低。仅改变两个异步操作的顺序；mock engine 的取消信号是独立的 CancellationToken，不依赖客户端状态。
- 影响：仅影响 Rust 测试模块的关闭路径，消除 CI 竞态失败，不影响生产逻辑。

关联脉络

无直接关联 PR。本修复确保了 Rust mock engine 测试的稳定性，为后续 Rust 前端开发提供可靠的基础设施。