

# PR #46494 完整报告

vllm-project/vllm

[CI/Build] Remove BaiChuanForCausalLM from the LoRA test

合并时间: 2026-06-23 22:09

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/46494>

## 执行摘要

- 一句话: 移除 LoRA 测试对 BaiChuan 模型的直接依赖
- 推荐动作: 可快速合并。该 PR 是良好的测试优化实践, 适合作为代码整洁和维护性改进的参考。

## 功能与动机

根据 PR body 和讨论链接 [https://github.com/vllm-project/vllm/pull/46362#discussion\\_r3452479264](https://github.com/vllm-project/vllm/pull/46362#discussion_r3452479264), 目的是移除 LoRA 测试对 BaiChuan 模型的直接依赖, 减少因模型实现变更导致的测试维护负担。

## 实现拆解

1. 移除 BaiChuan 模型导入: 删除 tests/lora/test\_lora\_checkpoints.py 中的 `from vllm.model_executor.models.baichuan import BaiChuanBaseForCausalLM`。
2. 引入 MOCK\_PACKED\_MAPPING: 新增一个硬编码的字典 `MOCK_PACKED_MAPPING`, 模拟 Baichuan 模型的 `packed_modules_mapping`, 包含 `W_pack` 和 `gate_up_proj` 两个条目。
3. 更新测试函数: 将 `test_load_checkpoints` 和 `test_lora_weights_mapping` 中原本引用 `BaiChuanBaseForCausalLM.packed_modules_mapping` 的逻辑替换为对 `MOCK_PACKED_MAPPING` 的引用, 其余测试逻辑不变。
4. 保持 Gemma4 测试不变: `test_gemma4_lora_weights_mapping` 不受影响。

关键文件:

- tests/lora/test\_lora\_checkpoints.py (模块 LoRA 测试; 类别 test; 类型 test-coverage) : 唯一修改的文件, 移除对 BaiChuan 模型的直接依赖, 引入 `MOCK_PACKED_MAPPING`。

关键符号: 未识别

## 关键源码片段

`tests/lora/test_lora_checkpoints.py`

唯一修改的文件, 移除对 BaiChuan 模型的直接依赖, 引入 `MOCK_PACKED_MAPPING`。

```
# SPDX-License-Identifier: Apache-2.0
```

```
# tests/lora/test_lora_checkpoints.py (变更后)
```

```
import pytest

from vllm.lora.lora_model import LoRAModel
from vllm.lora.peft_helper import PEFTHelper
from vllm.lora.utils import parse_fine_tuned_lora_name
# 移除了 : from vllm.model_executor.models.baichuan import BaiChuanBaseForCausalLM
from vllm.model_executor.models.gemma4 import Gemma4ForCausalLM
from vllm.model_executor.models.utils import WeightsMapper

lora_lst = ["baichuan7B", "baichuan7B-zero", "baichuan7B-zero-regex", "chatglm3-6b"]
BAICHUAN_LORA_MODULES = [
    "W_pack",
    "o_proj",
    "gate_up_proj",
    "down_proj",
]

# 新增: 内联的模拟 packed_modules_mapping, 替代直接依赖 BaiChuan 模型类
MOCK_PACKED_MAPPING = {
    "W_pack": ["W_pack"],
    "gate_up_proj": [
        "gate_proj",
        "up_proj",
    ],
}

# 测试函数其余部分保持不变, 仅将 BaiChuanBaseForCausalLM.packed_modules_mapping 替换为
MOCK_PACKED_MAPPING
```

## 评论区精华

PR 仅有一条来自 DarkLight1337 的 approved 评论, 无其他讨论。从 PR body 引用的 discussion 推测, 原始讨论建议移除对 BaiChuan 的依赖以简化测试。

- 暂无高价值评论线程

## 风险与影响

- 风险: 风险极低。变更仅限于测试文件, 未修改生产代码。测试逻辑本身不变, 只是将依赖的外部模型常量替换为内联常量。如果未来 Baichuan 模型打包映射发生变化, 测试需要手动同步 MOCK\_PACKED\_MAPPING, 但这是可接受的权衡。
- 影响: 影响范围仅限于 LoRA 测试文件。其他模块无影响。测试可维护性提升: 移除了对特定模型实现的直接依赖, 避免因模型重构导致的测试失败。
- 风险标记: 暂无

## 关联脉络

- PR #46362 (inferred) 可能的相关 PR: PR body 引用该 PR 的 discussion 作为变更动机。