

# PR #46945 完整报告

vllm-project/vllm

[Bugfix][Responses] Set completed status for Harmony function calls

合并时间: 2026-06-30 15:55

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/46945>

## 执行摘要

- 一句话: 修复 Harmony 模型 `function_call` 状态为 `null` 的问题
- 推荐动作: 值得合并。该 PR 修复了一个明确的合规性问题, 变更小、测试完备、无副作用。阅读者可以重点关注 `_parse_function_call()` 函数的构造模式, 理解 Harmony 适配层的实现方式。

## 功能与动机

Issue #46940 指出, 使用 Harmony 模型时, `/v1/responses` 返回的 `function_call` 输出项的 `status` 字段为 `null`, 而非 OpenAI API 规范要求的 `"completed"`、`"in_progress"` 或 `"incomplete"` 之一。该问题仅影响 Harmony 模型, 非 Harmony 模型行为正常。PR 旨在对齐 Harmony 模型的响应行为与 OpenAI 规范以及其他模型的实际表现。

## 实现拆解

1. 修改核心逻辑: 在 `vllm/entrypoints/openai/responses/harmony.py` 的 `_parse_function_call()` 函数中, 构造 `ResponseFunctionToolCall` 对象时, 新增 `status="completed"` 参数 (第 321 行)。该函数负责将 Harmony 消息中的函数调用解析为 RESP API 输出项, 此前遗漏了 `status` 字段, 导致返回 `null`。
2. 增加回归测试: 在 `tests/entrypoints/openai/responses/test_harmony_utils.py` 的现有测试方法 `test_commentary_with_function_recipient_creates_function_call` 中, 添加断言 `assert output_items[0].status == "completed"` (第 145 行), 确保函数调用输出项的 `status` 正确设置为 `completed`。
3. 变更规模: 仅两处单行改动 (+2/-0), 无其他文件或模块影响。

关键文件:

- `vllm/entrypoints/openai/responses/harmony.py` (模块 响应入口; 类别 `source`; 类型 `core-logic`; 符号 `_parse_function_call`): 核心修改文件: 在 `_parse_function_call()` 函数中为 `ResponseFunctionToolCall` 对象添加 `status="completed"` 参数, 修复 `function_call.status` 为 `null` 的 bug。
- `tests/entrypoints/openai/responses/test_harmony_utils.py` (模块 测试; 类别 `test`; 类型 `test-coverage`; 符号 `test_commentary_with_function_recipient_creates_function_call`): 测试文件: 在已有测试函数中添加断言, 验证 `function_call` 输出项具有正确的 `status` 值, 防止回归。

关键符号: `_parse_function_call`

## 关键源码片段

### `vllm/entrypoints/openai/responses/harmony.py`

核心修改文件: 在 `_parse_function_call()` 函数中为 `ResponseFunctionToolCall` 对象添加 `status="completed"` 参数, 修复 `function_call.status` 为 `null` 的 bug。

# `vllm/entrypoints/openai/responses/harmony.py` 第 309-324 行

```
def _parse_function_call(message: Message, recipient: str) -> list[ResponseOutputItem]:
    """将 Harmony 消息中的函数调用解析为 RESP API 输出项。"""
    function_name = extract_function_from_recipient(recipient)
    output_items = []
    for content in message.content:
        random_id = random_uuid()
        # 创建函数调用输出项; 必须显式设置 status='completed'
        # 符合 OpenAI Responses API 规范 (status 不允许为 null)
        response_item = ResponseFunctionToolCall(
            arguments=content.text,
            call_id=f"call_{random_id}",
            type="function_call",
            name=function_name,
            id=f"fc_{random_id}",
            status="completed", # 新增: 修复 bug#46940, 否则 status 为 null
        )
        output_items.append(response_item)
    return output_items
```

### `tests/entrypoints/openai/responses/test_harmony_utils.py`

测试文件: 在已有测试函数中添加断言, 验证 `function_call` 输出项具有正确的 `status` 值, 防止回归。

# `tests/entrypoints/openai/responses/test_harmony_utils.py` 第 125-145 行

```
def test_commentary_with_function_recipient_creates_function_call(self):
    """Test commentary with recipient='functions.X' creates function calls."""
    message = Message.from_role_and_content(
        Role.ASSISTANT, '{"location": "San Francisco", "units": "celsius"}'
    )
    message = message.with_channel("commentary")
    message = message.with_recipient("functions.get_weather")

    output_items = harmony_to_response_output(message)

    assert len(output_items) == 1
    assert isinstance(output_items[0], ResponseFunctionToolCall)
    assert output_items[0].type == "function_call"
    assert output_items[0].name == "get_weather"
```

```
assert (
    output_items[0].arguments
    == '{"location": "San Francisco", "units": "celsius"}'
)
assert output_items[0].call_id.startswith("call_")
assert output_items[0].id.startswith("fc_")
# 新增回归断言：确保 function_call 的 status 不为 null
assert output_items[0].status == "completed"
```

## 评论区精华

该 PR 的 Review 较为简单，未产生实质性讨论。审核者 [chaunceyjiang](#) 批准了 PR，仅标注了 [Thnaks~ LGTM](#)。没有其他 Reviewer 提出争议或设计问题。

- 暂无高价值评论线程

## 风险与影响

- 风险：风险极低。变更仅涉及两行代码：一行在源码中添加参数，一行在测试中添加断言。`_parse_function_call()` 是 Harmony 模型专有函数，不会影响非 Harmony 模型或其他 RESP API 路径。参数 `status="completed"` 的添加不会破坏现有逻辑，且已有完整测试覆盖。
- 影响：影响范围：仅影响 Harmony 模型通过 `/v1/responses` API 返回的 `function_call` 输出项。修复后，`status` 字段从 `null` 变为 `"completed"`，与 OpenAI API 规范一致。影响程度：较小，但提升了 API 的合规性和客户端兼容性。
- 风险标记：无显著风险

## 关联脉络

- 暂无明显关联 PR