

# PR #37783 完整报告

vllm-project/vllm

[release] Move agent queue to Release cluster queues

合并时间: 2026-03-24 11:36

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/37783>

## 执行摘要

- 一句话: 将 release pipeline 的构建队列从 postmerge 迁移到 release 队列, 提升隔离与安全性。
- 推荐动作: 此 PR 变更简单直接, 无需深入代码阅读; 关注点在于 CI 配置的隔离改进, 建议验证构建流程在合并后的稳定性, 并考虑添加相关测试覆盖。

## 功能与动机

PR body 中说明动机为 'for better isolation & secret protection', 即通过使用专用的 release 队列来提升构建环境的隔离性和秘密保护能力。

## 实现拆解

仅修改一个文件: .buildkite/release-pipeline.yaml。在该文件中, 将所有构建步骤的 'agents: queue:' 字段从 'arm64\_cpu\_queue\_postmerge' 和 'cpu\_queue\_postmerge' 分别改为 'arm64\_cpu\_queue\_release' 和 'cpu\_queue\_release'。这覆盖了 ARM64 架构的 CUDA 12.9、CUDA 13.0、CPU 构建, 以及 x86 架构的相应构建步骤和 release 镜像构建, 总计 8 处变更。

关键文件:

- .buildkite/release-pipeline.yaml (模块 CI/Infrastructure): 这是 CI release pipeline 的唯一配置文件, 所有队列变更均在此文件中进行, 直接影响构建环境隔离和安全性。

关键符号: 未识别

## 评论区精华

review 评论来自 gemini-code-assist[bot], 提到了一个新脚本 tools/ci/generate\_test\_deps.py 中的 bug, 但本 PR 未修改该文件, 因此讨论不直接相关。无其他实质性讨论或争议点。

- review 评论可能不相关 (other): 评论未针对本 PR 变更, 无实质性结论, 可能为误贴或无关讨论。

## 风险与影响

- 风险：主要风险包括：1. 配置变更可能导致构建失败，如果 release 队列未正确配置或资源不足；2. 缺少自动化测试验证队列变更后的构建流程，回归风险较高；3. 秘密保护依赖队列隔离，但具体实现细节未在 PR 中说明，可能存在配置错误或安全漏洞风险。
- 影响：对最终用户无直接影响；系统层面影响 CI release pipeline 的构建环境，可能改变构建性能或稳定性，需监控构建结果；团队需要确保新队列的配置与预期一致，并可能调整 CI 维护流程。
- 风险标记：配置变更未测试，队列资源未知，秘密保护依赖外部配置

## 关联脉络

- PR #37906 [ROCm][CI] Split Entrypoints Integration (API Server 1) into 3 jobs: 同样修改 Buildkite CI 配置文件 (.buildkite/test-amd.yaml)，优化测试执行，与本 PR 同属 CI 基础设施调整。
- PR #37882 [CI] split Entrypoints Integration (API Server 1) into 3 jobs: 涉及 CI 作业拆分，修改 .buildkite/test\_areas/entrypoints.yaml，与本 PR 的队列迁移都聚焦 CI 流程优化。