

PR #47065 完整报告

vllm-project/vllm

[ROCm][CI] Move PyTorch Compilation Unit Tests to MI300(gfx942)

合并时间: 2026-07-01 00:32

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/47065>

执行摘要

- 一句话: 将 PyTorch 编译单元测试从 MI250 迁移到 MI300
- 推荐动作: 属于常规 CI 硬件迁移, 实现简单, 无精读必要。但可作为 CI 配置变更的参考案例。

功能与动机

为了让 PyTorch 编译单元测试在 MI300 硬件上运行, 利用其更强的计算能力。PR 标题明确为 '[ROCm][CI] Move PyTorch Compilation Unit Tests to MI300(gfx942)'。

实现拆解

1. 移除 MI250 下的旧测试: 从 MI250 (gfx90a) 测试区块中删除 PyTorch Compilation Unit Tests 的完整定义 (共 18 行)。
2. 在 MI300 区块新增测试: 在 MI300 (gfx942) 测试区块中插入相同配置的测试, 关键差异在于:
 - mirror_hardware 从 [amdexperimental, amdproduction, amdgfx90anightly, amdmi250] 改为 [amdexperimental, amdproduction, amdgfx942nightly, amdmi300]
 - agent_pool 从 mi250_1 改为 mi300_1
3. 修复 review 中发现的遗漏: 第二次提交修正了首次提交中未正确更新的 mirror_hardware 字段, 将 amdgfx90anightly 改为 amdgfx942nightly。
4. 清理与对齐: 确保其他相关硬件标签一致性, 如将 amdmi250 改为 amdmi300。

关键文件:

- .buildkite/test-amd.yaml (模块 CI 配置; 类别 config; 类型 configuration): 唯一变更文件, 核心配置调整: 将 PyTorch 编译单元测试从 MI250 迁移到 MI300 硬件。

关键符号: 未识别

评论区精华

Review 中 AndreasKaratzas 指出了两次配置错误:

1. 第一次提交中 mirror_hardware 仍包含 amdgfx90anightly 和 amdmi250, 未更新为 MI300 对应的值。

2. 第二次提交后再次提醒 `amdgfx90anightly` 应改为 `amdgfx942nightly`。提交者随后通过第三次提交修正了这些问题，最终获得批准。
- 硬件标签未完全更新 (`correctness`): 提交者后续提交修正了该问题。

风险与影响

- 风险：低风险。变更仅涉及 CI 配置文件，无业务逻辑或运行时影响。可能的风险是 MI300 硬件资源不足导致测试排队或超时，但测试原为 `optional: true`，不会阻塞流水线。
- 影响：直接影响 AMD CI 流水线：PyTorch 编译单元测试将从 MI250 迁移到 MI300 执行。MI300 性能更强，可能缩短测试时间，但测试的 `timeout_in_minutes` 仍为 180 分钟。对其他平台（NVIDIA、Intel 等）无影响。
- 风险标记：仅 CI 配置变更

关联脉络

- 暂无明显关联 PR