

PR #39531 完整报告

vllm-project/vllm

[ROCm][CI] Introducing new MI300 nodes

合并时间: 2026-04-20 16:09

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/39531>

执行摘要

- 一句话: 启用 AMD MI300 单 GPU 测试代理, 优化 CI 测试分布。
- 推荐动作: 对于 CI/ 基础设施工程师, 值得精读此 PR 以了解 AMD 硬件测试配置和资源优化策略; 对于核心开发者, 只需关注测试结果的潜在变化。建议关注 CI 配置的组织方式、测试分组逻辑和资源请求设置。

功能与动机

根据 PR body, 动机是 'Enabling new 1 GPU MI300X test agents', 即启用新的单 GPU MI300X 测试代理, 以支持 AMD 最新硬件的 CI 测试。

实现拆解

1. 修改 CI 配置文件: 在 `.buildkite/test-amd.yaml` 中新增 `'gfx942 — mi300_8 (single-GPU shadow)'` 测试组, 包含多个单 GPU 测试步骤 (如核心功能、模型推理、量化测试), 指定 `agent_pool: mi300_8` 并避免多 GPU 请求。
2. 调整测试文件: 将 `tests/quantization/test_mi3xx_moe.py` 重命名为 `test_gfx950_moe.py`, 并更新函数名从 `test_mi3xx_moe` 到 `test_mi355_moe`, 以准确反映 MI355 硬件型号, 保持测试占位符。
3. 提交历史迭代: 30 次提交显示多次合并主分支和配置调整 (如 `'Redistributed MI300 tests'`、`'Commented out a temporarily dead test group'`), 确保与最新代码兼容并优化资源使用。
4. 测试配套: 无新增源码测试, 主要是配置和重命名, 以支持未来测试扩展。

关键文件:

- `.buildkite/test-amd.yaml` (模块 CI 配置; 类别 `config`; 类型 `configuration`): 主 CI 配置文件, 新增 MI300 单 GPU 测试步骤组, 影响整个 AMD CI 管道和测试分布。
- `tests/quantization/test_gfx950_moe.py` (模块 MoE 量化测试; 类别 `test`; 类型 `rename-or-move`; 符号 `test_mi3xx_moe`, `test_mi355_moe`): 测试文件重命名和函数更新, 确保测试与 MI355 硬件匹配, 保持测试架构一致性。

关键符号: `test_mi355_moe`

关键源码片段

.buildkite/test-amd.yaml

主 CI 配置文件，新增 MI300 单 GPU 测试步骤组，影响整个 AMD CI 管道和测试分布。

```
# gfx942 — mi300_8 (single-GPU shadow) # 新增测试组头，针对 AMD MI300 单 GPU 节点
steps:
  - label: "V1 Core + KV + Metrics (1 GPU)" # 单 GPU 测试步骤标签
    timeout_in_minutes: 120 # 设置超时
    agent_pool: mi300_8 # 指定代理池为 MI300 单 GPU 队列，确保资源匹配
    command: pytest -v -s tests/v1/core/test_core.py # 运行核心功能测试
    soft_fail: false # 不允许软失败，确保测试严格性
# 其他类似步骤省略，如模型测试、量化测试等，均配置为单 GPU
```

评论区精华

review 中，gemini-code-assist[bot] 指出配置错误：有测试步骤请求 `num_gpus: 2` 但 MI300_8 是单 GPU 队列，会导致 CI 失败；tjtanaa 询问是否有测试用例没运行任何测试以节省资源。讨论结论是这些问题被识别，PR 最终被批准，但提示了 CI 配置需精细化以避免资源浪费和失败。

- CI 配置错误导致资源不匹配 (correctness): 需要移除或修正这些多 GPU 测试步骤，确保配置正确，以避免 CI 作业失败。
- 优化 CI 资源使用以减少浪费 (performance): 可能已通过提交历史中的调整（如注释掉无用测试组）来处理，以减少资源浪费。

风险与影响

- 风险：主要风险是配置错误：如果多 GPU 测试被错误地安排在单 GPU 队列（如 .buildkite/test-amd.yaml 中的某些步骤），会导致 CI 作业失败，影响测试稳定性。此外，测试文件重命名可能影响测试发现流程，需确保所有引用更新。资源浪费风险：空测试或无用测试组可能增加 CI 成本。
- 影响：对系统影响：扩展 CI 测试能力，支持 AMD MI300 硬件，提高 vLLM 在该平台上的测试覆盖和兼容性。对团队：开发者能更早发现和修复 AMD 相关问题，但可能略增 CI 运行时间和资源成本。用户无直接影响，但间接提升软件稳定性和性能。
- 风险标记：配置错误风险，测试覆盖偏差

关联脉络

- PR #40305 [ci] Make ecr authenticate non blocking: 同属 CI 基础设施改进，涉及 Buildkite 配置优化和资源管理，与本 PR 的 CI 配置变更相关。