

PR #51271 完整报告

vllm-project/vllm

[CI] Run basic fullgraph correctness on one GPU

合并时间: 2026-08-07 00:09

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/51271>

执行摘要

- 一句话: fullgraph 正确性测试改为 1-GPU 专用, 删除多 GPU 冗余
- 推荐动作: 作为 CI 流程修正值得快速浏览, 不必精读。值得借鉴的决策是遇到测试在多套件中的偶发不稳定时, 先定位冗余覆盖与共享断言的不兼容, 再选择移除冗余而非放宽断言; 同时需注意删除 device_count 检查后测试与 1-GPU 套件的隐式绑定。

功能与动机

PR body 指出: PR #51074 将测试中唯一的 2-GPU 模型替换为 TP=1 Granite 模型; 测试宽松的 GPU 数检查导致 TP=1 用例在 2-GPU 和 4-GPU 套件中冗余运行; 同样的 Granite dummy-weight 对比在这些 worker 上失败, 因为相等的 top-5 logprobs 在 cutoff 处可能选中不同 token。作者选择删除多 GPU 冗余执行而非放宽共享正确性比较器, 保留 1-GPU H200 fullgraph 套件中的 Granite 与 embedding 覆盖。

实现拆解

1. 定位问题: PR #51074 将 test_basic_correctness.py 中唯一的 2-GPU 模型替换为 TP=1 的 Granite 模型后, 宽松的 GPU 数检查 (`device_count() < pp_size * tp_size` 才跳过) 使 TP=1 用例在 2-GPU、4-GPU 分布式套件中冗余执行, 且 Granite 的 dummy-weight 对比在这些 worker 上不稳定。
2. 清理测试文件: 从 tests/compile/fullgraph/test_basic_correctness.py 删除 meta-llama/Llama-3.2-1B-Instruct 的 PP=2、TP=2 参数化用例, 并删除 test_compile_correctness 中的 `current_platform.device_count()` 检查与 `pytest.skip` 逻辑, 让测试仅由 1-GPU H200 fullgraph 套件承载。
3. 调整 Buildkite 配置: 从 .buildkite/test_areas/distributed.yaml 的 distributed-compile-rpc-tests-2-gpus 与 distributed-compile-comm-4-gpus 两个步骤中移除 tests/compile/fullgraph/test_basic_correctness.py 的 source_file_dependencies 依赖与该测试的 pytest 命令, 避免重复执行。
4. 验证与配套: pre-commit (对两个文件)、git diff --check、以及 pytest --collect-only 均通过; 由于只涉及 CI 测试选择, PR 明确说明模型评估不适用。

关键文件:

- tests/compile/fullgraph/test_basic_correctness.py (模块 编译测试; 类别 test; 类型 test-coverage; 符号 TestSetting, test_compile_correctness) : fullgraph 编译正确性测

试的核心文件，删除 PP=2、TP=2 用例与 GPU 数检查，使测试转为 1-GPU 套件专用

- .buildkite/test_areas/distributed.yaml (模块 CI 配置；类别 config；类型 configuration)
：CI 编排入口，从 2-GPU、4-GPU 两个分布式步骤中移除该测试的依赖与命令

关键符号：test_compile_correctness, TestSetting

评论区精华

无技术性 review 评论：tlrmchlsmth 直接批准，claude[bot] 仅提示手动 review 配置。PR 作者 mgoin 在描述中说明核心决策——删除多 GPU 冗余执行而非放宽共享正确性比较器；并指出 Granite dummy-weight 对比失败是因为 top-5 logprobs 相等时 cutoff 可能选中不同 token。

- 移除 2-GPU / 4-GPU 冗余执行而非放宽比较器 (testing): 删除多 GPU 套件中的该测试，保留 1-GPU H200 fullgraph suite 覆盖；不修改共享的正确性比较器。

风险与影响

- 风险：
 - 环境假设收紧：删除 device_count() 检查后，test_basic_correctness.py 只能在 1-GPU 套件中运行；若未来被人加入多 GPU 套件，会直接以 TP=1 执行并可能重复或误报，需要依赖 Buildkite 配置兜底。
 - 覆盖范围收窄：2-GPU、4-GPU 套件不再执行该文件，但其中用例本就是 TP=1，并非真正的分布式编译路径，因此多 GPU 编译覆盖损失有限。
 - 未从根因修复比较器脆弱性：Granite 在 logprobs 相等时的 cutoff 抖动只是被移到了不会触发的环境，后续若 1-GPU H200 套件变化仍可能重现。
 - 无性能与安全风险。
- 影响：
 - CI 层面：两个分布式步骤各减少一次测试执行，节省 H100 / MI300 资源并消除偶发失败，CI 更稳定。
 - 团队流程：fullgraph 正确性验证职责清晰归属 1-GPU 套件；后续新增用例需保持 TP=1 并遵守单卡套件约定。
 - 用户侧无行为影响。
 - 风险标记：测试覆盖调整，CI 稳定性修复，环境假设收紧

关联脉络

- PR #51074 Replace 2-GPU model with TP=1 Granite in fullgraph correctness test: PR body 指明 #51074 将 test_basic_correctness.py 中唯一的 2-GPU 模型替换为 TP=1 Granite 模型，直接引发本 PR 的 CI 冗余与失败问题。