

PR #46356 完整报告

vllm-project/vllm

[XPU][CI]Skip v1/spec_decode/test_speculators_correctness.py in intel GPU nightly

合并时间: 2026-06-22 19:30

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/46356>

执行摘要

- 一句话: 跳过 Intel GPU nightly CI 中失败的 spec decode 测试
- 推荐动作: 该 PR 为临时性 CI 调整, 可直接合并。建议关联一个 Issue 以跟踪该测试的修复状态, 避免长期跳过。

功能与动机

该测试在 Intel GPU 上 nightly 运行时失败, 因此通过 CI 配置跳过, 以保持 CI 流程稳定。

实现拆解

1. 修改 .buildkite/scripts/hardware_ci/run-intel-ci-test.sh 脚本中 pytest -v -s v1/spec_decode 命令的 --ignore 参数列表, 增加 --ignore=v1/spec_decode/test_speculators_correctness.py。
2. 该变更使得 Intel GPU nightly CI 在运行 spec_decode 测试时跳过该文件, 不影响其他测试的执行。

关键文件:

- .buildkite/scripts/hardware_ci/run-intel-ci-test.sh (模块 部署脚本; 类别 infra; 类型 infrastructure): 该文件定义了 Intel GPU nightly CI 的测试命令, 修改通过 --ignore 参数跳过了 test_speculators_correctness.py 测试。

关键符号: 未识别

关键源码片段

[.buildkite/scripts/hardware_ci/run-intel-ci-test.sh](#)

该文件定义了 Intel GPU nightly CI 的测试命令, 修改通过 `--ignore` 参数跳过了 `test_speculators_correctness.py` 测试。

原命令:

```
# pytest -v -s v1/spec_decode --ignore=v1/spec_decode/test_max_len.py ...
```

修改后:

```
pytest -v -s v1/spec_decode \  
  --ignore=v1/spec_decode/test_max_len.py \  
  --ignore=v1/spec_decode/test_speculators_eagle3.py \  
  --ignore=v1/spec_decode/test_acceptance_length.py \
```

--ignore=v1/spec_decode/test_speculators_correctness.py # 新增跳过，避免 nightly 失败

评论区精华

无 review 讨论。

- 暂无高价值评论线程

风险与影响

- 风险：风险极低：仅修改 CI 脚本，跳过已知失败的测试，不涉及任何业务逻辑。但需要确保该测试在 Intel GPU 上确实不可用，否则可能遗漏回归问题。
- 影响：仅影响 Intel GPU nightly CI 的测试范围，不会影响其他平台或日常 CI。开发人员需要另行关注该测试的修复。
- 风险标记：临时跳过测试，需跟踪修复

关联脉络

- 暂无明显关联 PR