PR #26396 完整报告

sgl-project/sglang

[AMD] [CI] Add GLM-5.1 MXFP4 TP2 accuracy gate

合并时间：2026-05-27 16:49

原文链接：http://prhub.com.cn/sgl-project/sglang/pull/26396

执行摘要

- 一句话：为 AMD MI355X 添加 GLM-5.1-MXFP4 TP=2 GSM8K 准确率门控测试
- 推荐动作：该 PR 结构清晰、用途明确，值得在 AMD CI 相关维护者中了解。其中 _raise_nofile_limit 和 _get_model_path 是硬件测试中常见的实用工具，可复用于未来测试。建议关注其 CI 运行稳定性，并根据实际硬件资源调整 num_questions 和 parallel 等参数。

功能与动机

关联 Issue #25742 报告了 GLM-5.1-MXFP4 在 AMD MI355X 上 TP=2 时 GSM8K 准确率大幅下降（off: 0.32, EAGLE-MTP: 0.18），而 TP=8 仍保持正常。现有 8-GPU 覆盖未能捕获此回归，因此需要为 TP=2 添加特定的准确率门控测试，防止未来 aiter 或 SGLang 变更无声地重新引入该退化。

实现拆解

该 PR 仅新增一个测试文件，整体实现拆解如下：

1. 注册 CI 套件：通过 register_amd_ci(est_time=3600, suite="stage-c-test-large-8-gpu-amd-mi35x") 将测试加入 AMD 大型 GPU CI 阶段。
2. 定义模型路径和阈值：设置模型 ID amd/GLM-5.1-MXFP4，提供本地回退路径，并设定准确率阈值≥0.92、无效率≤0.02。默认并行 1200 个请求。
3. 启动服务器 (setUpClass)：在 setUpClass 中提升进程文件描述符软限制（避免 Too many open files），然后调用 popen_launch_server 启动 SGLang 服务，配置 TP=2、DSA tilelang prefill/decode、FP8 KV Cache、GLM 解析器以及其他必要参数。
4. 运行 GSM8K 测试 (test_gsm8k_accuracy)：从环境变量获取问题数量和并行度，调用 sglang.test.few_shot_gsm8k.run_eval 执行 5-shot GSM8K 评估，收集准确率和无效率指标并断言满足阈值。同时在 CI 中输出 Markdown 摘要。
5. 清理 (tearDownClass)：通过 kill_process_tree 终止服务器进程。

关键文件：

- test/registered/amd/accuracy/mi35x/test_glm51_mxfp4_tp2_gsm8k_mi35x.py（模块 CI 门控；类别 test；类型 test-coverage；符号 _raise_nofile_limit, _get_model_path, TestGLM51MXFP4TP2GSM8KMI35x, setUpClass）：该文件是 PR 的唯一变更，实现了整个准确率门控测试：包含 CI 注册、服务器启动、GSM8K 评估和阈值断言。

关键符号：_raise_nofile_limit, _get_model_path, TestGLM51MXFP4TP2GSM8KMI35x.setUpClass, TestGLM51MXFP4TP2GSM8KMI35x.tearDownClass, TestGLM51MXFP4TP2GSM8KMI35x.test_gsm8k_accuracy

评论区精华

无实质性讨论。仅作者提交后由 HaiShaw 快速批准（APPROVED），无任何 review 评论。

- 暂无高价值评论线程

风险与影响

- 风险：低风险：该 PR 仅添加测试文件，不修改任何运行时代码。风险主要来自外部依赖：
  - 测试依赖特定硬件（AMD MI355X/gfx950）和本地模型路径，若 CI 环境不匹配则可能跳过或失败。
  - est_time=3600 可能导致 CI 超时（实际约 250 秒，留有余量但仍有阻塞风险）。
  - 使用 kill_process_tree 如果进程异常可能留下孤儿进程，但已在 tearDownClass 中处理。
  - 影响：影响范围：仅限于 AMD CI 流程中的 stage-c-test-large-8-gpu-amd-mi35x 套件。每次 PR 提交（AMD 相关）将额外运行约 4 分钟的 GSM8K 评估（实际测得 249 秒），增加了 CI 总耗时但保证了关键模型 TP=2 的精度回归检测。对用户和其他平台无影响。
  - 风险标记：硬件依赖，CI 耗时增加，环境路径敏感

关联脉络

- 暂无明显关联 PR