

PR #23086 完整报告

sgl-project/sglang

[CI] GB200 nightly: on-demand PR/branch image build and config filter

合并时间: 2026-04-23 04:51

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/23086>

执行摘要

- 一句话: 为 GB200 夜间流水线添加按需 PR/ 分支镜像构建和配置过滤功能。
- 推荐动作: 对于负责 CI 基础设施的工程师, 此 PR 值得精读, 以了解如何扩展 GitHub Actions 工作流支持按需构建和过滤; 关注输入验证和镜像清理的设计决策, 以及如何平衡灵活性与安全性。

功能与动机

根据 PR body, 动机是解决 GB200 流水线只能使用已发布镜像的问题, 使得能够验证未合并 PR、测试 fork PR 和运行针对性基准测试, 减少集群时间浪费。具体表述为: "The GB200 nightly pipeline could previously only run against published lmsysorg/sglang:dev-cu13 images. This made it impossible to:

- Validate an unmerged PR on GB200 hardware. - Test a fork PR — fork contributors had no path to exercise their changes on the shared GB200 cluster. - Run a targeted benchmark — every manual dispatch ran the full matrix (~6 hours of cluster time), even when the tester only cared about one config."

实现拆解

1. 添加输入参数: 在 .github/workflows/nightly-72-gpu-gb200.yml 中扩展 workflow_dispatch 输入, 新增 pr_number、sglang_branch、configs 参数, 并更新 image 输入描述为可选且互斥, 以支持不同镜像来源和配置筛选。
2. 输入验证: 新增 validate-inputs job, 使用 shell 脚本检查输入冲突 (如同时设置多个镜像源) 和 pr_number 格式 (必须为正整数), 确保无效输入在占用集群资源前导致流水线失败。
3. 镜像构建与清理: 添加 build-image job 在 ARM 节点上构建 ARM64/CUDA13 镜像, 推送到临时仓库 lmsysorg/sglang-staging; cleanup-image job 保留最近 60 个标签并清理旧镜像, 防止仓库膨胀。
4. 配置过滤: 修改 scripts/ci/slurm/generate_matrix.py, 添加 --filter 命令行参数, 支持按名称筛选矩阵配置, 并提供错误提示列出可用名称, 避免拼写错误。
5. 配套调整: 包括工作空间清理步骤处理 root-owned 残留文件、提交信息打印步骤增强可追溯性, 以及更新环境变量如 CI_IMAGE_REPO 和 CI_IMAGE_KEEP_TAGS 以支持新流程。

关键文件:

- `.github/workflows/nightly-72-gpu-gb200.yml`（模块 CI 工作流；类别 infra；类型 infrastructure）：定义了整个 GB200 夜间流水线，添加了新输入参数和 jobs（如 `validate-inputs`、`build-image`、`cleanup-image`），是实现按需镜像构建和验证的核心文件。
- `scripts/ci/slurm/generate_matrix.py`（模块 CI 脚本；类别 infra；类型 infrastructure）：添加配置过滤功能，支持通过 `--filter` 参数选择性运行基准测试配置，是减少集群时间浪费的关键脚本。

关键符号：未识别

关键源码片段

`scripts/ci/slurm/generate_matrix.py`

添加配置过滤功能，支持通过 `--filter` 参数选择性运行基准测试配置，是减少集群时间浪费的关键脚本。

```
def main():
    # ... 解析命令行参数
    parser.add_argument(
        "--filter",
        default="",
        help=(
            "Optional comma-separated list of matrix entry names to include "
            "(e.g. 'dsr1-fp8-1k1k-max-tpt'). Names must match exactly."
        ),
    )
    # ... 读取配置并生成矩阵
    wanted = [n.strip() for n in args.filter.split(",") if n.strip()] # 解析过滤列表
    if wanted:
        available = [e["name"] for e in matrix] # 获取可用配置名称
        unknown = [n for n in wanted if n not in available] # 检查未知名称
        if unknown:
            print(
                f"ERROR: unknown config name(s): {' '.join(unknown)}. "
                f"Available for runner '{args.runner}': {' '.join(available)}",
                file=sys.stderr,
            )
            sys.exit(1) # 退出并提示错误
        matrix = [e for e in matrix if e["name"] in wanted] # 过滤矩阵
    print(json.dumps(matrix)) # 输出过滤后的矩阵
```

评论区精华

review 中仅有一个 bot 评论（`gemini-code-assist[bot]`），没有实质性讨论或争议，因此无重要决策或未解决疑虑。

- 暂无高价值评论线程

风险与影响

- 风险：技术风险包括：输入验证逻辑可能遗漏边缘情况（如空字符串或特殊字符处理），在 `.github/workflows/nightly-72-gpu-gb200.yml` 的 `validate-inputs` job 中依赖 `shell` 脚本，若脚本错误可能导致无效输入通过；镜像清理策略（保留 60 个标签）可能意外删除仍需使用的镜像，尤其在并发运行时；新 jobs 依赖自托管 runner（如 `arm-docker-build-node`），可能引入稳定性或性能问题；配置过滤功能在 `scripts/ci/slurm/generate_matrix.py` 中，若配置名称变更未同步更新，可能导致过滤失败。
- 影响：对用户（开发者）：现在可以更方便地测试未合并 PR 和分支在 GB200 硬件上，提升开发体验；对系统：CI 流水线更灵活，减少不必要的全矩阵运行时间，节约集群资源；对团队：提升开发效率和协作，但需维护新增加的 jobs 和脚本，可能增加运维复杂度。影响范围限于 CI 基础设施，不影响核心模型推理或业务逻辑。
- 风险标记：输入验证风险，镜像清理策略，自托管 runner 依赖

关联脉络

- PR #23492 [CI] `/rerun-stage: auto-include wheel build when PR modifies sgl-kernel/`: 都涉及 CI 流水线改进，调整了 GitHub Actions 工作流以支持更灵活的测试和构建。
- PR #23447 [CI] `Move disaggregation basic CI back to 2-gpu suite`: 关联 CI 测试套件调整，展示了仓库中 CI 基础设施的持续演进。
- PR #23465 `Add 'already' to ignore words list in .codespellrc`: 同为基础设施小改动，反映了团队对开发工具链的维护。