

PR #35764 完整报告

sgl-project/sglang

[AMD][CI] Fix ROCm 7.0's dead apt index fail the MORI dependency install

合并时间: 2026-08-21 13:13

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/35764>

执行摘要

- 一句话: 修复 ROCm 7.0 CI 因失效 apt 索引导致依赖安装失败
- 推荐动作: 值得精读。本 PR 虽然只改了一个 shell 脚本和一个测试文件, 但 PR body 是一份高质量的根因分析案例: 通过对比三个镜像 flavor 的 base image lineage 与 job 日志, 把“基础设施故障”精确判定为“代码回归”, 并区分了修复根因与收窄爆炸半径两个层次。值得关注的设计决策包括: 用镜像 flavor 门控恢复旧行为、用 `ll echo` 而非静默吞错保持可见性、以及把测试从 subprocess 重放精简为静态不变式。对于维护 AMD CI 或类似多外部源依赖的团队, 这是很好的参考模板。

功能与动机

PR body 指出集群 R387 的 27 个 job 与 nightly-7.0 的 kimi-k26 job 全部在“Install dependencies”阶段以相同方式死亡, 日志显示 `Err:13 ... compute-rocm-rel-7.0/38 amd64 Packages 404 Not Found` 与 `E: Some index files failed to download. They have been ignored, or old ones used instead.`。作者明确判断这是代码回归: `apt-get update` 是 2026-08-20 由 c7478228dd (#30984, ROCm 7.2.4 / Python 3.12 / torch 2.11 升级) 新增的未保护命令; 只有 ROCm 7.0 的 AMD 内部基础镜像继承了 `rocm-osdb-22.04-deb` 这个失效源, 而该源并非本仓库安装任何包所需; 此外该步骤还触及 6 个外部 apt 主机, 任何一个索引失败都会在 `set -e` 下拖垮整个 AMD CI。

实现拆解

1. 根因定位: 在 `scripts/ci/amd/amd_ci_install_dependency.sh` 的 MORI 重装流程中, `docker exec ci_sglang bash -c "..."` 内以 `set -euo pipefail` 运行 `apt-get update` 与 `apt-get install -y --no-install-recommends libgrpc++-dev`。 `apt-get update` 对单个索引返回退出码 100 时会直接中止整个步骤; 该段是 #30984 中唯一未按镜像 flavor 门控的改动。
2. 脚本修复: 把两条 apt 命令放入 `if [ "${IMAGE_STAGE_SUFFIX}" = '-rocm724' ]` 分支。只有 rocm724 (noble 系) 基础镜像确实缺少 `libgrpc++-dev`; ROCm 7.0 与 7.2.0 在该步骤引入前一直不装此包也能构建 MORI, 因此跳过 apt 往返并恢复升级前行为。在 rocm724 分支内, 两条命令分别追加 `ll echo '[MORI] ...'`, 以显式日志代替原 `2>/dev/null ll true`, 既保留降级可见性又避免 `set -e` 中止。作者没有添加重试: 镜像已通过 `/etc/apt/apt.conf.d/80-net-hardening` 配置 `Acquire::Retries "5"`, 且 404 本身不可重试。

3. 新增回归测试：test/registered/unit/tools/test_amd_ci_install_dependency.py (45 行) 静态扫描安装脚本中所有匹配 (sudo)?apt-get 的行，断言每条调用都带 II 保护。测试注册到 base-a-test-cpu suite，无需 docker、ROCm 或 GPU，因此任何 PR 的 CPU 测试都能覆盖。测试最初采用 subprocess stub apt-get 重放 404，后精简为纯静态断言，只保留“不得存在未保护 apt-get”这一不变式。
4. 配套说明：无其他配置文件或部署变更；由于 pr-test-amd.yml (ROCm 7.0) 在 #34204 中被降级为 daily shadow，已无 pull_request 触发，常规 PR 流程无法捕获此类回归，所以该静态测试是唯一且必要的防线。

关键文件：

- scripts/ci/amd/amd_ci_install_dependency.sh (模块 安装脚本；类别 infra；类型 infrastructure)：本 PR 的核心修复文件：将 MORI 重装流程中的 apt 步骤以 IMAGE_STAGE_SUFFIX 门控限定到 rocm724，并将两条命令非致命化，直接消除 R387 集群 27 个 job 失败的根因。
- test/registered/unit/tools/test_amd_ci_install_dependency.py (模块 回归测试；类别 test；类型 test-coverage；符号 TestAmdCiInstallDependencyApt, test_apt_calls_cannot_abort_the_dependency_install)：新增回归测试：由于 pr-test-amd.yml (ROCm 7.0) 已被降级为 daily shadow，没有 PR 触发的工作流能捕获此类回归，这个运行在 CPU suite 上的静态断言成为唯一防线。

关键符号：test_apt_calls_cannot_abort_the_dependency_install

关键源码片段

scripts/ci/amd/amd_ci_install_dependency.sh

本 PR 的核心修复文件：将 MORI 重装流程中的 apt 步骤以 IMAGE_STAGE_SUFFIX 门控限定到 rocm724，并将两条命令非致命化，直接消除 R387 集群 27 个 job 失败的根因。

```
# MORI 重装流程 (docker exec ci_sglang bash -c "...") 之内)
# 仅 rocm724 (noble 系) 基础镜像缺少 libgrpc++-dev; ROCm 7.0 / 7.2.0
# 在该步骤引入前几个月一直不装此包也能构建 MORI，因此只在 -rocm724
# 时执行 apt 往返，对其余 flavor 恢复升级前的行为。
if [ "${IMAGE_STAGE_SUFFIX}" = '-rocm724' ]; then
    # apt-get update 只要有一个索引 404 就会以退出码 100 结束，但已拉到的
    # 索引仍会被保留；该步骤涉及 6 个外部 apt 源，任何一个不可达都会在
    # set -e 下令整个 "Install dependencies" 失败。因此两条命令都必须
    # 非致命化，并用显式 echo 日志替代原来的 2>/dev/null 吞错，让降级状态可见。
    apt-get update \
        || echo '[MORI] apt-get update reported errors; continuing with the indexes it did fetch'
    apt-get install -y --no-install-recommends libgrpc++-dev \
        || echo '[MORI] libgrpc++-dev unavailable; building MORI without it'
fi
```

test/registered/unit/tools/test_amd_ci_install_dependency.py

新增回归测试：由于 pr-test-amd.yml (ROCm 7.0) 已被降级为 daily shadow，没有 PR 触发的工作流能捕获此类回归，这个运行在 CPU suite 上的静态断言成为唯一防线。

```
"""Guard on the apt calls in scripts/ci/amd/amd_ci_install_dependency.sh.
```

Those calls run under ``set -euo pipefail``, and ``apt-get update`` exits 100 when any single index is unreachable -- even though it keeps every index it did fetch. An unguarded call therefore fails the whole 'Install dependencies' step on every AMD runner at once, which is what took out ~25 of 27 jobs in pr-test-amd run 32399046576 when AMD's internal rocm-osdb artifactory started 404ing on an index this repo never installs from.

The packages involved are optional -- rocm.Dockerfile builds MORI without them -- so no apt call here may be able to abort the run.

```
"""
```

```
import re
import unittest
from pathlib import Path
```

```
from sglang.test.ci.ci_register import register_cpu_ci
from sglang.test.test_utils import CustomTestCase
```

```
register_cpu_ci(est_time=1, suite='base-a-test-cpu')
```

```
# 通过相对路径解析仓库根目录下的安装脚本，任意工作目录下都可运行
```

```
INSTALL_SCRIPT = (
    Path(__file__).resolve().parents[4] / 'scripts/ci/amd/amd_ci_install_dependency.sh'
)
```

```
class TestAmdCiInstallDependencyApt(CustomTestCase):
```

```
    def test_apt_calls_cannot_abort_the_dependency_install(self):
```

```
        # 逐行扫描脚本，找出所有可能“裸奔”的 apt-get 调用：在 set -e 下，
        # 没有 || fallback 的 apt-get 一旦遇到某个外部索引不可达，就会中止
        # 整个依赖安装步骤，因此这里要求未加保护的 apt-get 列表必须为空。
```

```
        unguarded = [
            line.strip()
            for line in INSTALL_SCRIPT.read_text().splitlines()
            if re.match(r"\s*(sudo\s+)?apt-get\b", line) and "||" not in line
        ]
```

```
        self.assertEqual(
            unguarded,
            [],
            'an unguarded apt-get under `set -e` fails the dependency install on '
            'every AMD runner whenever one apt source is unreachable; give it an '
            '`|| echo ...` fallback',
        )
```

```
if __name__ == '__main__':
    unittest.main()
```

评论区精华

本 PR 没有 review 评论，HaiShaw 直接批准合并；核心讨论体现在 commit 历史中。第二次提交首次用 `stub apt-get` 在 `set -e` 下重放 404 来验证修复；第三次提交主动将该方案精简为静态行扫描，理由是 subprocess harness 主要验证的是 shell 的 `||` 语义而非业务不变式，静态断言更轻量且更直接。PR body 还用 R391（同一天 apt.llvm.org GPG/curl 失败）佐证“6 个外部 apt 源都是单点”的判断，说明守卫并非只针对 rocm-osdb。

- 暂无高价值评论线程

风险与影响

- 风险：

1. 门控条件脆弱：修复依赖 `'${IMAGE_STAGE_SUFFIX}' = '-rocm724'` 这一精确字符串匹配；未来新增镜像 flavor 若需要 `libgrpc++-dev` 而未匹配该后缀，会静默跳过安装。不过该包对 MORI 是可选的，历史上 7.0/7.2.0 未装也能构建，风险可控。
2. 静态扫描可能漏检：测试用正则逐行匹配 `apt-get` 且要求行内含 `||`，若未来有人把命令改成多行写法（反斜杠续行、变量拼接或函数封装），静态断言可能误报或漏报；这是作者有意为之的取舍。
3. 降级构建语义：在 7.2.4 上若 `libgrpc++-dev` 安装失败，日志会明确提示“building MORI without it”，与 #30984 之前行为一致；但若未来 MORI 真的依赖该包，降级可能引入运行时隐患，需配合最近日志检查。
4. 7.2.0 未实测：PR body 承认 7.2.0 的 apt 源 lineage 是推断而非观察，当前无 job 运行该版本，若重新启用需验证。
 - 影响：对用户：无直接运行时影响，属于 CI 基础设施修复。对系统：直接恢复 R387 上 ROCm 7.0 的所有 job 及 nightly kimi-k26 job，并消除 6 个外部 apt 源对 AMD CI 的单点故障；显式 [MORI] 日志让降级状态在 CI 日志中可见，可观测性提升。对团队：新增的静态回归测试运行在通用 CPU suite 上，无需 AMD 硬件即可守护安装脚本，降低未来维护成本；对 PR #30984 引入的回归形成了明确的防再犯机制。影响范围限定在 `rucrpts/ci/amd/amd_ci_install_dependency.sh` 与测试文件，不触及 SRT 运行时逻辑。
 - 风险标记：外部 apt 源单点故障，rocm724 门控依赖镜像 flavor 匹配，静态脚本扫描可能漏检复杂 shell，可选依赖降级构建 MORI

关联脉络

- PR #30984 ROCm 7.2.4 / Python 3.12 / torch 2.11 upgrade: 本 PR 修复的回归正是该 PR 在 `scripts/ci/amd/amd_ci_install_dependency.sh` 中引入的未保护 `apt-get update` 步骤；PR body 明确指出这是根因。
- PR #34204 Demote ROCm 7.0 pr-test-amd workflow to daily shadow: 正文指出该 PR 将 `pr-test-amd.yml` (ROCm 7.0) 降级为 daily shadow，使 `pull_request` 触发无法捕获本回归，从而催生了静态回归测试这一防线。
- PR #35654 [AMD] Retry transient network failures in ROCm Dockerfile curl fetches: 同属 AMD CI / Docker 依赖安装稳定性修复，聚焦外部网络源不可用场景，方向一致。