

PR #2237 完整报告

radixark/miles

fix(openenv): build E2B task templates as root

合并时间: 2026-08-08 08:54

原文链接: <http://prhub.com.cn/radixark/miles/pull/2237>

执行摘要

- 一句话: E2B 模板构建固定 root 用户并纳入别名摘要
- 推荐动作: 建议快速精读。变更范围小但设计完整: 把影响产物内容的因素 (构建用户) 纳入 alias digest, 与既有“资源变更重烘焙”策略一致, 是 E2B 模板缓存正确性的通用模式。同时值得学习作者在 PR 描述中区分“修复”与“契约声明”的表达方式。团队在使用 openenv/e2b 后端时无需额外操作, 仅注意首次构建会因 alias 变化触发一次性重建。

功能与动机

PR body 明确指出根因与隐蔽性: "E2B Cloud runs template-build commands as a non-root user, so every layer of the per-task TB2 recipe fails there — apt-get exits 100 and /opt is not writable. A self-hosted AgentENV builds as root, which is why the e2b backend passed its bring-up without this surfacing." 同一套代码在自托管环境自测通过, 却在云环境处处失败, 属于典型 provider 行为差异导致的隐性故障。作者还特意区分了两处改动的性质: 模板构建的 set_user("root") 是真正修复; server exec 固定用户 "That one is not a fix — both endpoints were measured to already default to root at runtime", 作用是声明契约, 防止 provider 默认值改变后, agent 在期望 root 的任务镜像里静默失权 (TB2 的 solution 和测试会自由 apt-install)。

实现拆解

1. 新增构建用户常量: 在 examples/experimental/openenv/tb2_sandbox_e2b.py 模块顶层新增 _BUILD_USER = "root", 构建与运行时共用。注释明确: TB2 任务镜像按 root agent 设计, 任何非 root 用户都会改变任务环境本身, 而不仅是构建过程。
2. 修复模板构建: ensure_task_template 中把 Template().from_image(base) 改为链式 Template().from_image(base).set_user(_BUILD_USER), 保证用户切换发生在任何 run_cmd 之前——set_user 只影响其后的命令, 顺序是正确性的关键。
3. 摘要纳入构建用户: template_alias 的 digest 输入由 [base, *commands, resources] 扩展为 [base, _BUILD_USER, *commands, resources]。构建用户是烘焙产物的组成部分, 变化时 alias 必须跟着变, 才能强制重新构建而不是继续使用第一个构建者留下的旧模板。这与既有“任务资源变化触发重烤”的策略一脉相承。
4. 固定 server 运行用户: create_task_sandbox 中 sandbox.commands.run(..., background=True, user=_BUILD_USER)。如动机所述这是契约而非修复, 防止 provider 默认值漂移。

5. 测试配套: tests/test_tb2_sandbox_e2b.py 扩展 _FakeTemplateCls (新增有序 steps 记录) 与 _FakeSandbox (新增 ran 记录), 新增 3 个用例:
test_ensure_template_builds_as_root (断言 user == "root" 且 steps[0] == ("set_user", "root"))、test_template_alias_tracks_the_build_user (monkeypatch _BUILD_USER 后 alias 变化)、test_create_runs_the_env_server_as_root (断言 server exec 的 user、background、命令内容)。离线 60 通过 2 跳过, ruff 与 black 通过。

关键文件:

- examples/experimental/openenv/tb2_sandbox_e2b.py (模块 E2B 沙箱; 类别 source; 类型 core-logic; 符号 _BUILD_USER, template_alias, ensure_task_template, create_task_sandbox): 修复主路径: 新增 _BUILD_USER 常量、ensure_task_template 链式 set_user、template_alias 摘要纳入构建用户、server exec 固定用户。
- examples/experimental/openenv/tests/test_tb2_sandbox_e2b.py (模块 E2B 沙箱; 类别 test; 类型 test-coverage; 符号 set_user, test_ensure_template_builds_as_root, test_template_alias_tracks_the_build_user, test_create_runs_the_env_server_as_root): 新增 3 个测试锁定 set_user 顺序、digest 对构建用户敏感、server exec 用户参数; fake 类扩展用于断言调用顺序。

关键符号: template_alias, ensure_task_template, create_task_sandbox,
test_ensure_template_builds_as_root, test_template_alias_tracks_the_build_user,
test_create_runs_the_env_server_as_root

关键源码片段

examples/experimental/openenv/tb2_sandbox_e2b.py

修复主路径: 新增 _BUILD_USER 常量、ensure_task_template 链式 set_user、
template_alias 摘要纳入构建用户、server exec 固定用户。

```
# 构建与运行时统一使用的用户。TB2 任务镜像按 root agent 设计
# (其测试会直接 apt-install), 任何非 root 用户都会改变任务
# 环境本身。构建阶段是真正修复: E2B Cloud 默认以非 root 执行
# 构建命令, 导致 apt-get 退出码 100、/opt 不可写; 自托管
# AgentENV 以 root 构建, 所以此前未暴露。运行时两端实测本就
# 默认 root, 显式传入只为固定契约, 防止 provider 改变默认值
# 时静默收窄 agent 权限。
_BUILD_USER = "root"
```

```
def ensure_task_template(
    task_dir: Path,
    *,
    force: bool = False,
    build_timeout_s: float = 1800.0,
    on_logs=None,
) -> str:
    """按任务构建模板, alias 已存在则直接复用; 返回 alias。"""
    # 惰性导入: e2b SDK 是可选依赖, 离线单测用 sys.modules 里的
```

```

# fake 替换，模块导入本身不能依赖真实 SDK。
from e2b import Template

task_dir = Path(task_dir)
alias = template_alias(task_dir)
with _build_lock(alias):
    # alias 已存在且未强制时短路，避免重复支付数分钟构建。
    if not force and Template.alias_exists(alias, **_connection_opts()):
        return alias
    base = resolve_docker_image(task_dir, None)
    # set_user 必须先于第一条 build command，它只影响其后的
    # 命令；alias 摘要已纳入 _BUILD_USER，改变构建用户会得到
    # 新 alias，强制重烤而不是复用旧产物。
    template = Template().from_image(base).set_user(_BUILD_USER)
    for command in server_layer_commands(task_dir):
        template = template.run_cmd(command)

def _build() -> None:
    Template.build(
        template,
        alias,
        **task_build_resources(task_dir),
        skip_cache=force,
        on_build_logs=on_logs,
        **_connection_opts(),
    )

# 不用 with：其 __exit__ 会 join worker，等于等完整个
# 本应被 timeout 放弃的构建。
pool = concurrent.futures.ThreadPoolExecutor(max_workers=1)
try:
    pool.submit(_build).result(timeout=build_timeout_s)
finally:
    pool.shutdown(wait=False)
return alias

cmd = server_cmd(command_timeout_s, default_task_id=task_dir.name)
# 环境服务器执行 agent 的命令，它运行的用户就是任务环境
# 的用户。E2B 默认非 root，而 TB2 任务镜像预期 root，
# 因此显式固定，避免 agent 静默丢失应有的权限（例如
# apt-install）。
sandbox.commands.run(
    f"bash -c {shlex.quote(cmd)} > /tmp/openenv-server.log 2>&1",
    background=True,
    user=_BUILD_USER,
)

```

评论区精华

本 PR 没有任何 review 评论或要求变更，Shi-Dong 直接 APPROVED (LGTM!)。讨论价值集中在 PR 作者自己的设计说明：正文明确写出 That one is not a fix，主动区分“修复”与“契约声明”两处改动。这种自述避免了维护者把 server exec 的 user 参数误读为另一项功能修复，从而放松对 provider 默认值漂移的警惕——是 bugfix PR 中值得借鉴的沟通方式。

- 暂无高价值评论线程

风险与影响

- 风险：
 - 一次性重建成本：template_alias 摘要新增 _BUILD_USER 后，所有既有 E2B 模板的 alias 都会变化，已烘焙模板会被视为缺失并重新构建一次。作者合并后实测 4 个模板 2m57s 重建干净，成本可接受，但大规模使用前需注意首次触发的构建并发。
 - 模块级硬编码 root：_BUILD_USER 同时约束构建与运行时两处语义。若未来引入非 root 设计的任务镜像，该常量需要演进为按任务配置，否则会改变任务环境而非仅构建过程。
 - SDK 参数兼容性：sandbox.commands.run(..., user=...) 依赖 E2B SDK 支持该参数。作者已对 E2B Cloud 与自托管 AgentENV 实测通过，但离线测试使用 fake 模块，无法捕获真实 SDK 版本间的参数差异。
 - 影响范围：仅限 examples/experimental/openenv 的 e2b backend；daytona 后端与训练主路径不受影响。
 - 影响：对用户而言，修复使 e2b backend 在 E2B Cloud 上真正可用——此前模板构建必然失败，训练无法走通；自托管 AgentENV 用户无行为变化。对系统而言，alias digest 变化会触发一次性的模板重建，之后恢复常态。对团队而言，这是 openenv 实验线从“单 provider 自测可用”走向“多 provider 契约固定”的关键一步，为后续 provider 扩展提供了可复制的用户语义处理模式。
- 风险标记：跨 provider 行为差异，模板一次性重建，SDK 参数兼容，模块级硬编码 root

关联脉络

- PR #2274 refactor(openenv): move duplicated sandbox helpers into the TB2 recipe: 同一 openenv/tb2_sandbox 系列，本 PR 的 ensure_task_template 依赖共享 recipe 的 server_layer_commands、task_build_resources 等；两 PR 共同收敛 e2b/daytona 后端逻辑。
- PR #2220 Add the GLM-5.2 744B x terminal-bench-2 Daytona example: tbench2 openenv 实验线的 Daytona 示例；本次修复使 e2b backend 在 E2B Cloud 可运行，扩展同一实验线的 provider 覆盖面。
- PR #2276 : PR body 明确提到 Details in #2276（合并后四个模板 2m57s 重建 + GRPO 训练走通），是本次修复的配套验证；该 PR 未包含在提供的历史列表中，标题未知。