

PR #2876 完整报告

radixark/miles

ci(e2e): Harbor trials on real e2b sandboxes through the full rollout path (manual for now)

合并时间: 2026-09-10 11:23

原文链接: <http://prhub.com.cn/radixark/miles/pull/2876>

执行摘要

- 一句话: 新增 Harbor E2B 沙箱端到端测试, 验证 agentic 路径完整链路, 为 CI 自动化铺路。
- 推荐动作: 本 PR 值得精读, 尤其是测试脚本的设计和演进。它展示了如何为复杂的 agentic 路径构建端到端测试, 包括环境预检、资源准备、执行和结果验证的清晰分离。其中关于解决 Qwen3 模型特定问题、处理陈旧状态和增强预检逻辑的迭代过程 (通过提交历史可见) 具有学习价值。最终的 `--debug-rollout-only` 方法巧妙地降低了测试的资源需求 (仅需 2 GPU)。对于从事 agentic 系统、沙箱集成或端到端测试开发的工程师, 此 PR 提供了实用的模式参考。

功能与动机

为了验证 Harbor 在真实 E2B 沙箱上的完整端到端流程, 因为现有的沙箱 smoke 测试 (`scripts/sandbox_smoke`) 无法覆盖 launcher 接线到 rollout workers、session server + TITO 使用真实模型、terminus-2 驱动沙箱以及奖励回流的完整训练路径。PR body 中明确说明, 此前的训练路径形状 (一个 GRPO 步骤) 因 Qwen3-0.6B 模型的退化输出需要严格门控的豁免, 而当前变更通过仅运行 rollout 并使用更强的模型 (GLM-4.7-Flash) 解决了此问题。

实现拆解

1. 新增端到端测试脚本: 在 `tests/e2e/agentic/test_harbor_e2b_rollout.py` 创建了完整的测试脚本。它定义了 `preflight()` (快速检查环境变量、密钥和网络连接)、`prepare()` (准备模型、任务数据集和提示数据)、`harbor_worker_env()` (构建 rollout workers 的 Harbor 环境变量)、`execute()` (组装并执行训练参数, 使用 `--debug-rollout-only` 仅运行 rollout) 和 `check_trials()` (验证至少有一个 Harbor 试验成功完成并获得奖励)。该脚本被注册为禁用状态, 需要手动运行。
2. 注册 CI 标签: 在 `tests/ci/labels.py` 的 `KNOWN_LABELS` 字典中添加了新的 agentic 标签, 描述为“需要沙箱服务路由和密钥的 Agentic 沙箱端到端测试”, 用于在 CI 系统中识别此类测试。
3. 更新 CI 评论命令策略: 在 `.github/workflows/policies/comment-command-access.json` 的 `allowed_labels` 列表中添加了 `run-ci-agentic`, 允许通过评论命令触发带有该标签的 CI 测试。

关键文件:

- tests/e2e/agentica/test_harbor_e2b_rollout.py (模块 端到端测试; 类别 test; 类型 test-coverage; 符号 preflight, prepare, harbor_worker_env, execute) : 核心测试文件, 实现了 Harbor E2B 沙箱端到端测试的完整流程, 是验证 agentica 路径的关键。
- tests/ci/labels.py (模块 CI 标签; 类别 test; 类型 configuration) : CI 标签注册表, 新增 'agentica' 标签以在 CI 系统中识别此类测试。
- .github/workflows/policies/comment-command-access.json (模块 CI 策略; 类别 infra; 类型 infrastructure) : CI 评论命令策略文件, 授权通过评论命令触发带有 'agentica' 标签的 CI 测试。

关键符号: preflight, prepare, harbor_worker_env, execute, check_trials

关键源码片段

tests/e2e/agentica/test_harbor_e2b_rollout.py

核心测试文件, 实现了 Harbor E2B 沙箱端到端测试的完整流程, 是验证 agentica 路径的关键。

端到端测试脚本的核心执行流程

```
if __name__ == "__main__":
    preflight() # 1. 快速失败检查: 环境变量、密钥文件、网络连通性、harbor 库
    prepare() # 2. 准备资源: 清理旧试验目录、下载模型、克隆任务仓库、生成提示数据
    # 清除代理变量, 确保 rollout workers 直接访问沙箱服务
    for proxy_var in ("http_proxy", "https_proxy", "HTTP_PROXY", "HTTPS_PROXY"):
        os.environ.pop(proxy_var, None)
    execute() # 3. 执行训练 (仅 rollout) : 组装参数并调用 execute_train
    check_trials() # 4. 验证结果: 检查试验目录存在且至少有一个试验无异常
```

preflight() 函数: 关键预检逻辑

```
def preflight():
    """Fail fast with instructions instead of failing every trial later."""
    api_url = os.environ.get("E2B_API_URL", "").strip()
    if not api_url:
        sys.exit("set E2B_API_URL (and E2B_SANDBOX_URL) to your E2B-compatible service; see the module docstring")
    key_file = Path(os.environ.get("E2B_API_KEY_FILE", "~/.config/e2b/api_key")).expanduser()
    # 镜像真实检查 (credentials.sandbox_key_supply): 空占位符文件必须在此失败
    file_has_key = key_file.is_file() and bool(key_file.read_text().strip())
    if not os.environ.get("E2B_API_KEY", "").strip() and not file_has_key:
        sys.exit(f"no e2b credential: set E2B_API_KEY or put a non-empty key at {key_file}")
    try:
        request = urllib.request.Request(f"{api_url}/nodes", headers={"X-API-Key": "preflight"})
        urllib.request.urlopen(request, timeout=10).read()
    except urllib.error.HTTPError:
        pass # 401 仍然证明控制平面有响应
    except OSError as e:
        sys.exit(f"sandbox service unreachable at {api_url} ({e}); does this machine have a route to it?")
    try:
        import harbor # noqa: F401
```

```
except ImportError:
    sys.exit("harbor is not importable; see the module docstring for the install line")
```

评论区精华

Review 讨论主要由 claude[bot] 发起，集中于测试脚本的健壮性和正确性。核心争议点包括：

- 陈旧试验目录：claude[bot] 指出 `prepare()` 未清理 `TRIALS_DIR`，可能导致旧的成功试验掩盖当前运行的失败。作者 nblintao 在后续提交中通过在 `prepare()` 开头添加 `shutil.rmtree(TRIALS_DIR, ignore_errors=True)` 修复了此问题。
- 预检密钥文件检查：claude[bot] 发现 `preflight()` 仅检查密钥文件是否存在 (`is_file()`)，但实际运行时会检查内容是否非空，这可能导致空占位文件通过预检但在训练深处失败。作者在后续提交中更新了检查逻辑为 `key_file.is_file()` and `bool(key_file.read_text().strip())`。
- TITO 严格门控豁免范围：claude[bot] 提出 `--ci-disable-tito-strict-checker` 标志会禁用所有三种不匹配类型的断言，而不仅仅是 Qwen3 的 `special_token_count` 误报，可能掩盖其他 bug。此评论指向的是被重写为仅 rollout 的旧版本测试 (`test_harbor_e2b_training.py`)，在最终合并的版本中，由于采用了 GLM-4.7-Flash 模型和正确的解析器配置，该标志已被移除，此问题不再存在。
- 试运行超时设置：claude[bot] 发现 `harbor_worker_env()` 中的 `AGENT_TRIAL_TIMEOUT` 设置可能因环境变量条件而不生效。作者在后续提交中将其改为无条件设置（使用 `os.environ.setdefault()`）。

所有讨论均以作者修复代码或问题因架构重写而不再存在告终，无未解决疑虑。

- 陈旧试验目录可能掩盖测试失败 (correctness)：作者在后续提交中修复了此问题，在 `prepare()` 开头添加了 `shutil.rmtree(TRIALS_DIR, ignore_errors=True)`。
- 预检密钥文件检查不充分 (correctness)：作者在后续提交中更新了检查逻辑为 `key_file.is_file()` and `bool(key_file.read_text().strip())`。
- TITO 严格门控豁免范围过宽 (correctness)：此评论针对的是旧版本测试（包含训练步骤）。最终合并的版本采用仅 rollout 模式和 GLM-4.7-Flash 模型，移除了该标志，问题不再存在。
- 试运行超时设置可能不生效 (correctness)：作者在后续提交中将 `AGENT_TRIAL_TIMEOUT` 的设置改为无条件（使用 `os.environ.setdefault()`）。

风险与影响

- 风险：
 1. 测试有效性风险：测试脚本注册为 disabled，依赖手动在特定开发机上运行，缺乏 CI 自动化，可能无法及时发现回归。虽然 PR 旨在为 CI 自动化铺路，但启用取决于 runner 配置就绪。
 2. 环境依赖风险：测试强依赖外部 E2B 兼容服务的网络路由和 API 密钥。`preflight()` 函数已尽力提前验证，但网络环境变化仍可能导致测试失败。
 3. 模型特定风险：测试固定使用 GLM-4.7-Flash 模型和特定的 TITO 家族 (glm47)，验证范围限于该组合。其他模型或配置的兼容性未覆盖。

4. 代码演进风险：从提交历史看，该 PR 经历了从包含训练步骤到仅 rollout 的重大重写。最终版本移除了对 Qwen3-0.6B 的支持以及相关的严格门控豁免逻辑，这简化了代码，但也意味着针对特定模型问题的修复可能被后续引入。- 影响：对用户的直接影响很小，因为这是一个测试变更。它增强了对 Harbor agentic 路径（特别是 E2B 沙箱集成）的端到端验证能力，为团队在集成和发布新功能时提供了更高的信心。引入的 agentic CI 标签为后续自动化测试奠定了基础。对系统的影响主要在于测试基础设施的扩展，增加了对 agentic 场景的覆盖。对团队的影响是提供了一个可复用的、结构清晰的手动测试模板，用于验证类似的 agentic 滚动更新路径。- 风险标记：测试依赖外部服务，CI 自动化待启用，模型特定验证，手动运行易出错

关联脉络

- PR #2868 smoke: manual golden episodes on real sandbox APIs: 同为沙箱相关测试，本 PR 的测试覆盖了更完整的 rollout 路径，是其上层集成验证。
- PR #2837 examples: launcher for the in-process Harbor path: 本 PR 测试的 launcher 接线和环境变量构建逻辑来源于该 PR 引入的 launch_common 模块。
- PR #3113 fix: chunk true-on-policy logprob computation: PR body 中提及 Qwen3-0.6B 模型的退化 turn 问题与 #3113 的分析相关，本 PR 最终通过更换模型解决了该问题。
- PR #3167 docs(tito): explain abnormal Qwen3 endoftext output: 评论中 guapisolo 提及了 Qwen3 模型的异常 EOS 输出行为，本 PR 的演进过程（从 Qwen3 到 GLM）与此背景相关。