

PR #2046 完整报告

THUDM/slime

Revert "[docker] always re-register mooncake addr during offloading"

合并时间: 2026-06-09 22:46

原文链接: <http://prhub.com.cn/THUDM/slime/pull/2046>

实现拆解

虽然本 PR 是 revert 操作, 但为了直观展示前后变化, 下面给出两个关键文件的对比。

`docker/patch/latest/sglang.patch` (移除内容示例)

该补丁在 #2041 中添加了 mooncake 地址重新注册逻辑, revert 后这些代码被移除。以下是 revert 后保留的补丁中一段取自其他 PR 的代码 (仅供参考):

```
def release_memory_occupation(self):
    self.queue.clear()
    self.retracted_queue.clear()
    if hasattr(self.kv_manager, "deregister_buffer_to_engine"):
        self.kv_manager.deregister_buffer_to_engine()

def resume_memory_occupation(self):
    if hasattr(self.kv_manager, "register_buffer_to_engine"):
        self.kv_manager.register_buffer_to_engine()
```

`docker/version.txt`

版本号回退至:

nightly-dev-20260608b