

PR #2199 完整报告

THUDM/slime

fix(npu): bracket IPv6 hosts in distributed init methods

合并时间: 2026-08-12 14:12

原文链接: <http://prhub.com.cn/THUDM/slime/pull/2199>

执行摘要

- 一句话: 扩展 NPU 补丁以括号包裹 IPv6 地址
- 推荐动作: 建议精读该 PR 以理解 NPU 补丁的维护方式, 关注 `_wrap_ipv6` 的实现和补丁的版本兼容性。

功能与动机

PR #2151 修复了当前主分支的权重更新 rendezvous URL, 但文档中的 NPU 设置使用 v0.2.2 和 `slime.patch`, 该版本中仍有三个 actor/critic 和权重更新的 `tcp init` 方法未处理 IPv6 地址。为保持补丁与固定版本兼容, 需要扩展补丁以使用已有的 `_wrap_ipv6` 辅助函数。

实现拆解

1. 修改 `docker/npu_patch/slime.patch`, 为三个文件 (`update_weight_utils.py`、`actor.py`、`update_weight_from_distributed.py`) 添加 `_wrap_ipv6` 导入和调用。
2. 在 `update_weight_utils.py` 和 `megatron_utils/actor.py` 中, 将 `init_method` 的 `tcp://` 地址用 `_wrap_ipv6` 包裹。
3. 在 `megatron_utils/update_weight/update_weight_from_distributed.py` 中同样处理。
4. 补丁使得三个模块编译通过, 且目标 URL 被正确处理。

关键文件:

- `docker/npu_patch/slime.patch` (模块 NPU 补丁; 类别 `infra`; 类型 `infrastructure`): 扩展补丁以修复 IPv6 地址处理

关键符号: 未识别

评论区精华

无 review 讨论或评论。

- 暂无高价值评论线程

风险与影响

- 风险: 补丁主要影响 NPU 部署场景, 若 `_wrap_ipv6` 对 IPv4 地址处理不当, 可能影响原有 IPv4 环境, 但该函数已存在, 风险较低。补丁针对固定版本 v0.2.2, 未来版本升级可能需

要同步更新。

- 影响：影响 NPU 部署的分布式初始化，确保 IPv6 环境下的稳定性。对团队而言，为 NPU 用户提供了修复，但需确保补丁与主分支同步。
- 风险标记：补丁版本兼容性，IPv6 处理依赖已有函数

关联脉络

- PR #2151 Fix weight-update rendezvous URL: 本 PR 修复了当前主分支的 URL，但补丁仍需针对 v0.2.2 扩展