

PR #27730 完整报告

sgl-project/sglang

[AMD]: docker(rocm) bump Mooncake to latest main + enable multi-protocol

合并时间: 2026-07-02 08:41

原文链接: <http://prhub.com.cn/sgl-project/sglang/pull/27730>

执行摘要

- 一句话: 升级 AMD ROCm Docker 中 Mooncake 并启用多协议
- 推荐动作: 该 PR 值得关注, 特别是对 AMD 平台部署 PD Disaggregation 的团队, 其多协议运行时选择的设计思路清晰, 可参考。

功能与动机

旧版 Mooncake 缺少多协议传输支持, 在 `USE_HIP=ON` 构建下仅支持 XGMI, 导致跨节点 RDMA 返回 `NotSupportedTransport` 错误。升级版本并启用多协议后, 同一进程可自动根据对端选择传输协议, 无需重建或手动切换。

实现拆解

1. 升级 Mooncake 版本与构建参数: 在 `docker/rocm.Dockerfile` 中将 `MOONCAKE_COMMIT` 更新至 `01d1eb2a7ec37fd5e20a88573e9b4956e7846e9a`, 并在 `cmake` 命令中添加 `-DENABLE_MULTI_PROTOCOL=ON -DWITH_STORE=ON -DBUILD_UNIT_TESTS=OFF`。其中 `BUILD_UNIT_TESTS=OFF` 用于规避 Mooncake 测试目标 `mp_transport_test` 因未启用 CXL 导致的链接错误。
2. 优化 CI 调试与设备配置: 在两个 CI 工作流文件 (`pr-test-amd-rocm720.yml` 和 `pr-test-amd.yml`) 的 `disaggregation` 测试步骤中添加 `rocm-smi --showtopotype` 输出、Mooncake 引擎的协议广告字符串检查、`MC_LOG_LEVEL=TRACE` 和 `GLOG_logtostderr=1` 环境变量。同时将 RDMA 设备从自动检测所有 HCA (`rdma0..rdma7`) 硬编码为单个 `rdma0`, 以解决单节点 1P1D 测试中多设备导致 P2P 握手失败的问题。

关键文件:

- `docker/rocm.Dockerfile` (模块 部署脚本; 类别 `infra`; 类型 `infrastructure`): 核心变更: 升级 Mooncake commit 并添加多协议构建标志。
- `.github/workflows/pr-test-amd-rocm720.yml` (模块 CI 配置; 类别 `infra`; 类型 `infrastructure`): 为 ROCm 7.2 CI 添加调试日志、环境变量和 RDMA 设备固定。
- `.github/workflows/pr-test-amd.yml` (模块 CI 配置; 类别 `infra`; 类型 `infrastructure`): 为主 AMD CI (ROCm 7.0) 同步添加相同的调试日志、环境变量和 RDMA 设备固定。

关键符号: 未识别

评论区精华

讨论主要集中在两部分：一是 Mooncake 自身测试编译失败（`undefined reference to typeinfo for mooncake::CxlTransport`），通过添加 `-DBUILD_UNIT_TESTS=OFF` 解决；二是单节点 1P1D disagg 测试中 RDMA 多设备导致 P2P 握手失败（`received packet mismatch`），最终固定为单个 `rdma0` 解决。

- Mooncake 多协议构建链接错误 (correctness): 通过添加 `-DBUILD_UNIT_TESTS=OFF` 禁用测试编译。
- 单节点 RDMA 多设备 P2P 失败 (correctness): 将 RDMA 设备固定为单个 `rdma0`。

风险与影响

- 风险：主要风险来自 Mooncake 版本升级可能引入新的 bug 或不兼容行为，但由于仅改动 Docker 构建参数和 CI 脚本，不涉及模型前向或内核代码，影响范围可控。CI 中添加的硬编码 `rdma0` 可能在未来多 NIC 环境下需要重新评估，但当前仅用于单节点调试场景。
- 影响：对 AMD ROCm 用户影响较大：PD Disaggregation 现在可以在多节点间通过 RDMA 传输 KV，同一镜像同时支持 XGMI 和 RDMA，简化部署。CI 中添加的调试日志有助于未来排查相关问题，但设备固定可能掩盖部分环境问题。
- 风险标记：依赖外部仓库版本迭代，CI 调试变化可能影响测试稳定性

关联脉络

- 暂无明显关联 PR