

PR #40453 完整报告

vllm-project/vllm

Update Dockerfile.rocm for AINIC & Thor NIC

合并时间: 2026-05-14 15:24

原文链接: <http://prhub.com.cn/vllm-project/vllm/pull/40453>

执行摘要

- 一句话: 为 ROCm 镜像添加 AINIC 和 Thor-2 NIC 驱动支持
- 推荐动作: 该 PR 是基础设施增强而非功能性变更, 对 ROCm 多节点用户至关重要。值得关注的设计决策包括: 默认安装所有 NIC 以减少认知负担 (类似 SGLang 做法)、提供 none 选项兼容不使用 MoRI 的场景、以及明确的版本参数化。对于需要构建定制 ROCm 镜像的团队, 此 PR 提供了可复用的模式。建议阅读 Dockerfile 中的函数封装和 case 分支, 理解如何在 Docker 构建中灵活管理多方驱动。

功能与动机

关联 Issue #38687 指出 ROCm 镜像未内置 AINIC 或 Thor-2 NIC 驱动, 导致 MoRI 多节点功能无法在 AMD 集群中直接使用。PR 目的是使 ROCm vLLM 镜像达到与 CUDA 类似的使用体验, 并参照 SGLang 的 ROCm Dockerfile 实现。社区成员强调需要让 Pollara AINIC 和 Thor-2 NIC 开箱即用, 避免用户手动安装驱动和重建 MoRI。

实现拆解

1. 修改构建参数默认值: 将 NIC_BACKEND 从 "none" 改为 "all", 并添加注释说明可用选项。将 AINIC_VERSION 从 1.117.5 改为 1.117.3-hydra (经测试兼容常见主机驱动)。
2. 定义安装函数: 在 RUN 层中新增 install_ainic() 和 install_bnxt() Shell 函数, 分别处理两个 NIC 的 apt 源配置、GPG 密钥验证、包安装和后续清理。
3. 扩展 case 语句: 支持 all (依次调用两个函数)、ainic (仅安装 AINIC)、bnxt (仅安装 Thor-2)、none (无操作) 四种模式, 并为无效输入报错。
4. 安装 MoRI Python 依赖: 保留 pip 安装 quart、msgpack 等, 确保 MoRI proxy 可正常运行。
5. 更新注释和文档: 在文件头部和 case 分支处添加详细说明, 帮助构建者理解选项含义和版本依赖。

关键文件:

- docker/Dockerfile.rocm (模块 构建脚本; 类别 infra; 类型 infrastructure; 符号 install_ainic, install_bnxt, NIC_BACKEND): 唯一变更文件, 实现所有 NIC 驱动安装逻辑, 包括新定义的 install_ainic 和 install_bnxt 函数以及基于 NIC_BACKEND 参数的分支选择。

关键符号: install_ainic, install_bnxt, NIC_BACKEND case 选择逻辑

关键源码片段

docker/Dockerfile.rocm

唯一变更文件, 实现所有 NIC 驱动安装逻辑, 包括新定义的 install_ainic 和 install_bnxt 函数以及基于 NIC_BACKEND 参数的分支选择。

```
# 定义 AINIC 用户空间库安装函数
```

```
install_ainic() {  
    # 配置 AMD AINIC apt 源, 使用 GPG 验证  
    apt-get update && apt-get install -y --no-install-recommends ca-certificates curl gnupg apt-  
transport-https  
    rm -rf /var/lib/apt/lists/*  
    mkdir -p /etc/apt/keyrings  
    curl -fsSL https://repo.radeon.com/rocm/rocm.gpg.key | gpg --dearmor > /etc/apt/keyrings/  
amdainic.gpg  
    echo "deb [arch=amd64 signed-by=/etc/apt/keyrings/amdainic.gpg] https://repo.radeon.com/  
amdainic/pensando/ubuntu/${AINIC_VERSION} ${UBUNTU_CODENAME} main" \  
    > /etc/apt/sources.list.d/amdainic.list  
    apt-get update && apt-get install -y --no-install-recommends \  
    libionic-dev \  
    ionic-common  
    rm -rf /var/lib/apt/lists/*  
}
```

```
# 定义 Broadcom Thor-2 用户空间库安装函数
```

```
install_bnxt() {  
    # 添加 Broadcom NIC apt 源, 使用公钥验证  
    install -m 0755 -d /etc/apt/keyrings  
    curl -fsSL https://packages.broadcom.com/artifactory/api/security/keypair/PackagesKey/  
public \  
    -o /etc/apt/keyrings/broadcom-nic.asc  
    chmod a+r /etc/apt/keyrings/broadcom-nic.asc  
    echo "deb [arch=amd64 signed-by=/etc/apt/keyrings/broadcom-nic.asc] https://packages.  
broadcom.com/artifactory/ethernet-nic-debian-public jammy main" \  
    > /etc/apt/sources.list.d/broadcom-nic.list  
    apt-get update && apt-get install -y --no-install-recommends \  
    bnxt-rocelib=235.2.86.0  
    # 手动复制库文件并更新链接  
    cp -a /usr/local/lib/x86_64-linux-gnu/libbnxt_re* /usr/local/lib/  
ldconfig  
    rm -rf /var/lib/apt/lists/*  
}
```

```
# 根据 NIC_BACKEND 参数执行对应安装
```

```
case "${NIC_BACKEND}" in  
    none) ;; # 不安装任何驱动  
    all) install_ainic; install_bnxt ;; # 默认: 安装所有支持的 NIC
```

```
ainic) install_ainic ;; # 仅安装 AMD AINIC
bnxt) install_bnxt ;; # 仅安装 Broadcom Thor-2
*) echo "ERROR: unknown NIC_BACKEND=${NIC_BACKEND}. Use one of: none, ainic, bnxt,
all"; exit 2 ;;
esac
```

评论区精华

Review 讨论聚焦于默认值选择和镜像质量: tjtanaa 建议将默认 NIC_BACKEND 设为 all, 并提供 none 作为不安装的选项, 避免对无需 MoRI 的用户增加负担; gemini-code-assist[bot] 指出 bnxt 安装时未清除 cmake、ninja 等构建依赖, 建议 purge 以减小镜像体积和安全风险, 但 PR 最终未采纳该建议; simondanielsson 在集群上验证了 AINIC 版本兼容性, 并将默认 AINIC_VERSION 从 1.117.5 下调至 1.117.3-hydra, 以匹配更多主机驱动。

- 默认 NIC_BACKEND 值选择 (design): 采纳, 最终默认值为 all, 并保留 none 分支。
- none 分支注释不准确 (documentation): 最终注释已修正, none 分支不再指示为 default。
- bnxt 安装中构建依赖残留 (performance): 未采纳; 最终代码中仍保留了这些构建依赖的安装, 未添加清理步骤。

风险与影响

- 风险:
 1. 镜像体积风险: 默认安装两个 NIC 的用户空间库, 增加约 10-20 MB, 但相对于模型权重可忽略;
 2. 版本兼容风险: 用户空间库版本需与主机内核模块匹配, 若版本不匹配 MoRI 将报错。PR 允许通过 AINIC_VERSION 参数自定义, 但默认版本可能不兼容所有集群;
 3. 外部源安全风险: 添加了 repo.radeon.com 和 packages.broadcom.com 两个 apt 源, 但均使用了 GPG 密钥验证, 风险可控;
 4. 构建依赖残留: bnxt 安装中 cmake、ninja-build 等工具未清理, 可能增加镜像体积和攻击面 (reviewer 已指出但未解决)。- 影响: 用户影响: 使用 AMD MI300/MI355 等 ROCm 平台并搭配 AINIC 或 Thor-2 NIC 进行多节点推理的用户现在可以直接使用官方镜像, 无需手动安装驱动; 不使用 MoRI 的用户可通过 --build-arg NIC_BACKEND=none 保持原有轻量化镜像。系统影响: 镜像大小增加, 但 MoRI 运行时自动检测 NIC 类型, 无性能开销。团队影响: 需要维护 NIC 版本与主机驱动的同步, 未来版本升级时需关注兼容性。- 风险标记: 镜像大小增加, 构建依赖残留, 版本兼容风险, 外部源安全验证

关联脉络

- 暂无明显关联 PR