



vhost：一种 virtio 高性能的后端驱动实现

本文首发于我的公众号 **Linux云计算网络 (id: cloud_dev)**，专注于干货分享，号内有 **10T** 书籍和视频资源，后台回复 **1024** 获取，欢迎大家关注，二维码文末可以扫。

什么是 vhost

vhost 是 virtio 的一种后端实现方案，在 virtio 简介中，我们已经提到 virtio 是一种半虚拟化的实现方案，需要虚拟机端和主机端都能完成通信，通常，virtio 主机端的驱动是实现在用户空间的 qemu 中，而 vhost 是实现在内核中，是内核的一个模块 vhost-net.ko。为什么要实现在内核中，有什么好处呢，请接着往下看。

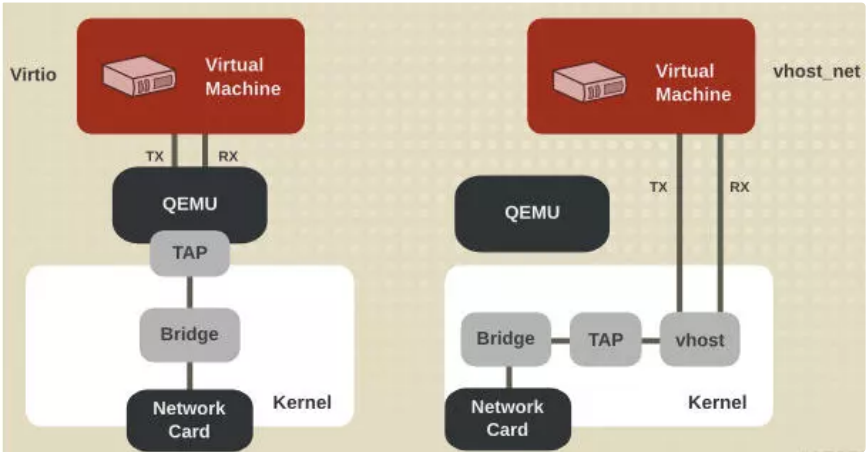
CONTENTS

1. 什么是 vhost

2. 为什么要用 vhost

3. vhost 的数据流程

4. 总结



为什么要用 vhost

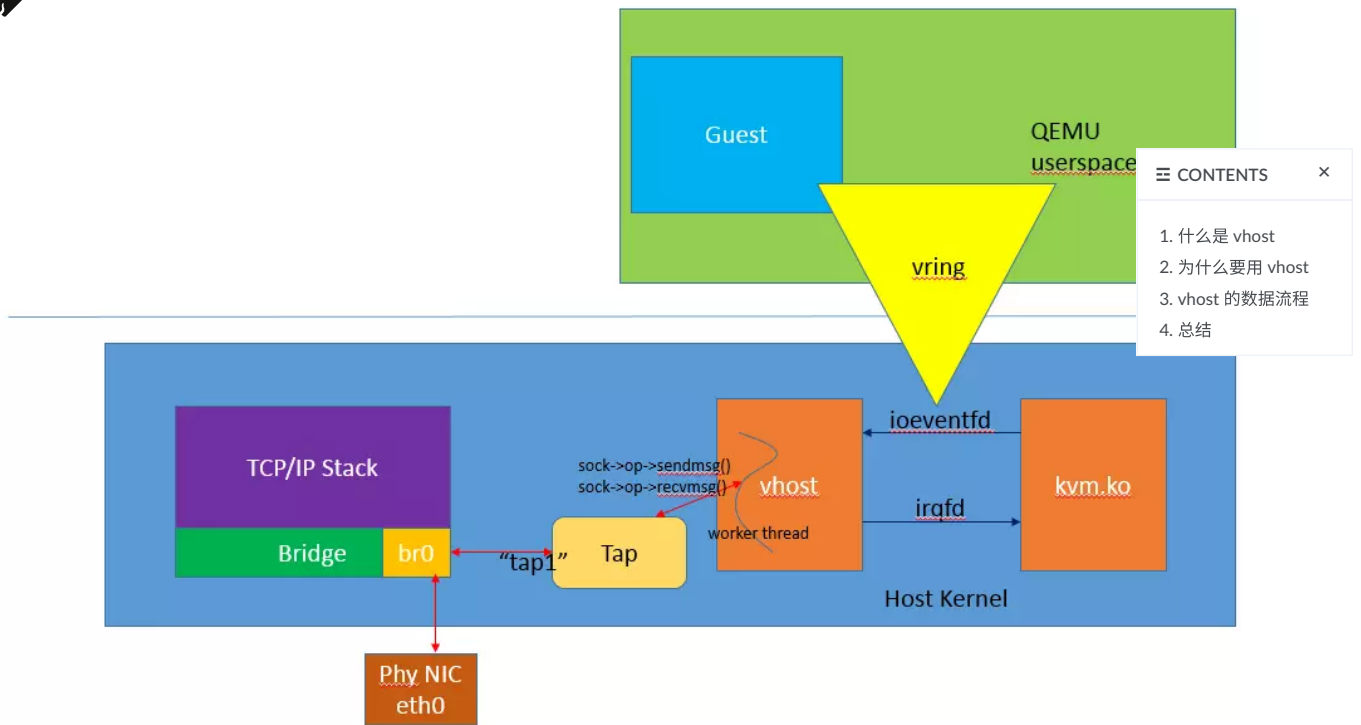
在 virtio 的机制中，guest 与 用户空间的 Hypervisor 通信，会造成多次的数据拷贝和 CPU 特权级的上下文切换。例如 guest 发包给外部网络，首先，guest 需要切换到 host kernel，然后 host kernel 会切换到 qemu 来处理 guest 的请求，Hypervisor 通过系统调用将数据包发送到外部网络后，会切换回 host kernel，最后再切换回 guest。这样漫长的路径无疑会带来性能上的损失。

vhost 正是在这样的背景下提出的一种改善方案，它是位于 host kernel 的一个模块，用于和 guest 直接通信，数据交换直接在 guest 和 host kernel 之间通过 virtqueue 来进行，qemu 不参与通信，但也没有完全退出舞台，它还要负责一些控制层面的事情，比如和 KVM 之间的控制指令的下发等。

vhost 的数据流程

下图左半部分是 vhost 负责将数据发往外部网络的过程，右半部分是 vhost 大概的数据交互流程图。其中，qemu 还是需要负责 virtio 设备的适配模拟，负责用户空间某些管理控制事件的处理，而 vhost 实现较为纯净，以一个独立的模块完成 guest 和 host kernel 的数据交换过程。





vhost 与 virtio 前端的通信主要采用一种事件驱动 eventfd 的机制来实现，guest 通知 vhost 的事件要借助 kvm.ko 模块来完成，vhost 初始化期间，会启动一个工作线程 work 来监听 eventfd，一旦 guest 发出对 vhost 的 kick event，kvm.ko 触发 ioeventfd 通知到 vhost，vhost 通过 virtqueue 的 avail ring 获取数据，并设置 used ring。同样，从 vhost 工作线程向 guest 通信时，也采用同样的机制，只不过这种情况发的是一个回调的 call event，kvm.ko 触发 irqfd 通知 guest。

总结

vhost 与 kvm 的事件通信通过 eventfd 机制来实现，主要包括两个方向的 event，一个是 guest 到 vhost 方向的 kick event，通过 ioeventfd 实现；另一个是 vhost 到 guest 方向的 call event，通过 irqfd 实现。

代码分析整个通信的流程：

<http://royluo.org/2014/08/22/vhost/>

我的公众号「Linux云计算网络」(id: cloud_dev)，号内有 10T 书籍和视频资源，后台回复「1024」即可领取，分享的内容包括但不限于 Linux、网络、云计算虚拟化、容器 Docker、OpenStack、Kubernetes、工具、SDN、OVS、DPDK、Go、Python、C/C++ 编程技术等内 容，欢迎大家关注。

Linux云计算网络

云计算 | 网络 | Linux | 干货

获取学习大礼包后台
回复“1024”

加群交流后台回复“加群”



作者：公众号「Linux云计算网络」，专注于Linux、云计算、网络领域技术干货分享

出处：<https://www.cnblogs.com/bakari/p/8341133.html>

本站使用「署名 4.0 国际」创作共享协议，转载请在文章明显位置注明作者及出处。

分类：云计算，虚拟化

标签：云计算，虚拟化，KVM



推荐 1

赞赏

收藏

反对 0

« 上一篇: [DPDK NFV 性能提升](#)

» 下一篇: [DPDK 全面分析](#)

posted @ 2018-01-24 13:04 CloudDeveloper 阅读(4

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问](#) 网站首页。

≡ CONTENTS ×

1. 什么是 vhost
2. 为什么要用 vhost
3. vhost 的数据流程
4. 总结

Copyright © 2020 CloudDeveloper
Powered by .NET Core on Kubernetes
Powered By Cnblogs | Theme Silence v2.0.0

