

公告

昵称： 于杨
园龄： 3年9个月
粉丝： 5
关注： 0
+加关注

<	2020年4月						>
日	一	二	三	四	五	六	
29	30	31	1	2	3	4	
5	6	7	8	9	10	11	
12	13	14	15	16	17	18	
19	20	21	22	23	24	25	
26	27	28	29	30	1	2	
3	4	5	6	7	8	9	

搜索

找找看

谷歌搜索

常用链接

- 我的随笔
- 我的评论
- 我的参与
- 最新评论
- 我的标签

随笔档案

- 2016年11月(4)
- 2016年10月(4)
- 2016年9月(2)
- 2016年8月(4)
- 2016年7月(7)

文章分类

- Linux 协议栈(1)

阅读排行榜

- 1. Linux Kernel Vhost 架构(5978)
- 2. [Mark] KVM 虚拟化基本原理(4709)
- 3. NFV 利器 FD.io VPP(4659)

Linux Kernel Vhost 架构

Vhost 回顾

Linux中的vhost驱动程序提供了内核virtio设备仿真。通常，QEMU用户空间进程模拟guest的I / O访问。Vhost将virtio仿真代码放到内核中，使QEMU用户空间脱离。这允许设备仿真代码直接调用到内核子系统，而不是从用户空间执行系统调用。

vhost-net驱动程序模拟主机内核中的virtio-net网卡。Vhost-net是最早的vhost设备，也是主流Linux中唯一可用的设备。也已经开发了实验vhost-blk和vhost-scsi装置。

在Linux 3.0中，vhost代码存放在drivers / vhost /中。所有设备使用的通用代码在drivers / vhost / vhost.c中。这包括virtio vring访问功能，所有virtio设备需要为了与客户进行通信。vhost-net代码存放在drivers / vhost / net.c中。

Vhost 驱动模型

vhost-net驱动程序在主机上创建一个/ dev / vhost-net字符设备。此字符设备作为配置vhost-net实例的接口。

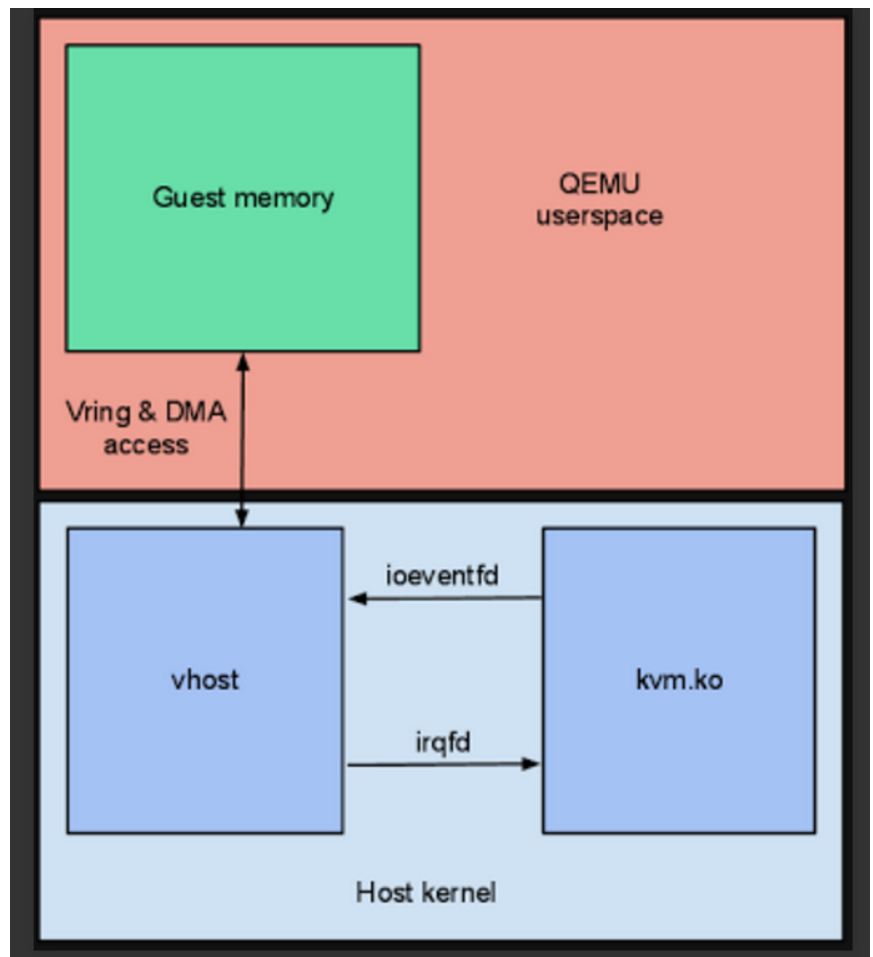
当使用-netdev tap启动QEMU时，vhost = on将打开/ dev / vhost-net并使用几个ioctl调用初始化vhost-net实例。这些必须将QEMU进程与vhost-net实例关联，准备virtio功能协商，并将guest虚拟机物理内存映射传递到vhost-net驱动程序。

在初始化期间，vhost驱动程序创建一个名为vhost- \$ pid的内核线程，其中\$ pid是QEMU进程pid。这个线程被称为“vhost工作线程”。工作线程的任务是处理I / O事件并执行设备仿真。

4. Linux 内核协议栈之TCP连接关闭 (3360)
5. FD.io VPP 技术Neutron VNF v Router 实现(2926)

推荐排行榜

1. Neutron网络性能测试与分析 (一) CVR(1)



内核Virtio仿真

Vhost不会模拟一个完整的virtio PCI适配器。相反，它只将自己限制为virtqueue操作。QEMU仍然用于执行virtio特性协商和实时迁移。这意味着vhost驱动程序不是完整的virtio设备实现，它依赖于用户空间来处理控制平面，而数据平面在内核中完成。

vhost工作线程等待virtqueue kick，然后处理放在virtqueue上的缓冲区。在vhost-net这意味着从tx virtque的数据包并且通过tap文件描述符发送它们。

文件描述符轮询也由vhost工作线程完成。在vhost-net中，当数据包进入tap文件描述符时，工作线程被唤醒，并将其放置到rx virtqueue中，以便客户端可以接收它们。

Vhost作为用户空间接口

vhost架构的一个令人惊讶的方面是它不以任何方式绑定到KVM。Vhost是一个用户空间接口，不依赖于KVM内核模块。这意味着其他用户空间代码（如libpcap）在理论上可以使用vhost设备，如果他们发现它们方便的高性能I / O接口。

当客户端kick主机，因为它已经将缓冲区放在一个virtqueue，需要一种方式来通知vhost工作线程有工作要做。由于vhost不依赖于KVM内核模块，它们不能直接通信。相反，vhost实例使用vhost工作线程监视活动的eventfd文件描述符进行设置。KVM内核模块具有称为ioeventfd的功能，用于获取

eventfd并将其挂接到特定的guest虚拟机I / O出口。 QEMU用户空间注册一个ioeventfd用于VIRTIO_PCI_QUEUE_NOTIFY硬件寄存器访问，它能kick virtqueue。这是当guest虚拟机kick virtqueue，vhost工作线程被KVM内核模块通知的方式。

在从vhost工作线程返回到中断客户端时，使用了类似的方法。 Vhost需要一个“call”文件描述符，写这个文件描述符是为了去kick guest。 KVM内核模块有一个称为irqfd的功能，它允许eventfd触发客户机中断。 QEMU用户空间注册一个irqfd为virtio PCI设备中断并将其交给vhost实例。这是vhost工作线程如何中断客户端。

因此，vhost实例只知道guest虚拟机内存映射，kick eventfd和call eventfd。

More detail

下面是开始探索代码的要点：

drivers / vhost / vhost.c - 常见的vhost驱动程序代码

drivers / vhost / net.c - vhost-net驱动

virt / kvm / eventfd.c - ioeventfd和irqfd

QEMU用户空间代码显示如何初始化vhost实例：

hw / vhost.c - 常见的vhost初始化代码

hw / vhost_net.c - vhost-net初始化

好文要顶

关注我

收藏该文



于杨

关注 - 0

粉丝 - 5

0

0

+加关注

« 上一篇: [Linux 虚拟机性能监控](#)

» 下一篇: [Linux kernel Vhost-net 和 Virtio-net代码详解](#)

posted @ 2016-11-10 14:24 于杨 阅读(5978) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

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。

【推荐】超50万行VC++源码：大型组态工控、电力仿真CAD与GIS源码库

【推荐】腾讯云产品限时秒杀，爆款1核2G云服务器99元/年！

【推荐】独家下载 | 《大数据工程师必读手册》揭秘阿里如何玩转大数据

【推荐】斩获阿里offer的必看12篇面试合辑

相关博文:

- virtIO之VHOST工作原理简析
- vhost-user 简介
- macvtap与vhost-net技术
- linux kernel 字符设备详解
- linux下TUN/TAP虚拟网卡的使用
- » 更多推荐...

2019 Flink Forward 大会最全视频来了! 5大专题不容错过

最新 IT 新闻:

- MIUI 12发布: 大幅改进界面效果 超40款机型支持升级
- 《我的世界》建筑团队还原迷宫都市欧拉丽 宏伟震撼
- .NET 5.0 Preview 3 发布
- 致癌的基因突变, 可能比预想的来得更早
- 微软表示XSX已达成多项里程碑 将为玩家带来惊喜
- » 更多新闻...

Copyright © 2020 于杨

Powered by .NET Core on Kubernetes