



vhost-user 简介

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。

CONTENTS

1. 什么是 vhost-user

2. vhost-user 的实现

3. 几个例子

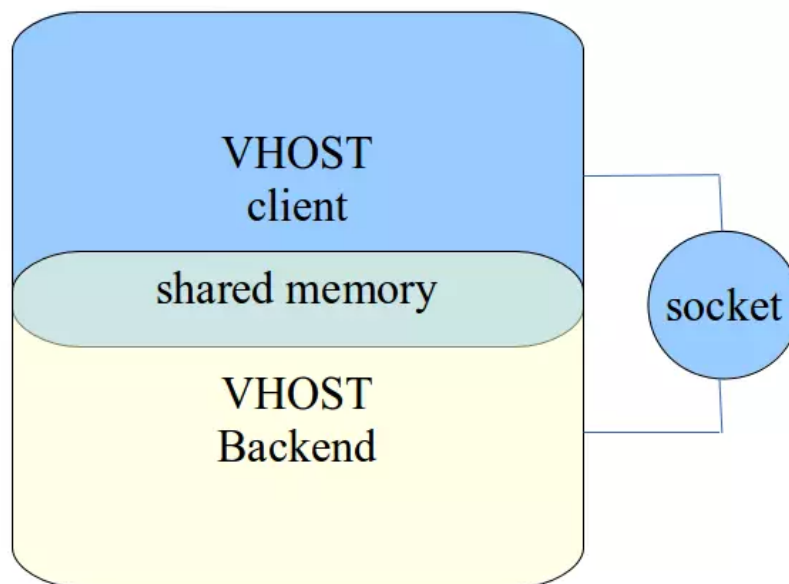
4. 总结

什么是 vhost-user

在 vhost 的方案中，由于 vhost 实现在内核中，guest 与 vhost 的通信，相较于原生的 virtio 方式性能上有了一定程度的提升，但 kvm.ko 的交互只有一次用户态的切换以及数据拷贝。这个方案对于不同 host 之间的通信，或者 guest 到 host nic 之间的通信是比较好的，但是对于某些用户态进程间的通信，比如数据面的通信方案，openvswitch 和与之类似的 SDN 的解决方案，guest 需要和 host 用户态的 vswitch 进行数据交换，如果采用 vhost 的方案，guest 和 host 之间又存在多次的上下文切换和数据拷贝，为了避免这种情况，业界就想出将 vhost 从内核态移到用户态。这就是 vhost-user 的实现。

vhost-user 的实现

vhost-user 和 vhost 的实现原理是一样，都是采用 vring 完成共享内存，eventfd 机制完成事件通知。不同在于 vhost 实现在内核中，而 vhost-user 实现在用户空间中，用于用户空间中两个进程之间的通信，其采用共享内存的通信方式。



vhost-user 基于 C/S 的模式，采用 UNIX 域套接字（UNIX domain socket）来完成进程间的事件通知和数据交互，相比 vhost 中采用 ioctl 的方式，vhost-user 采用 socket 的方式大大简化了操作。

vhost-user 基于 vring 这套通用的共享内存通信方案，只要 client 和 server 按照 vring 提供的接口实现所需功能即可，常见的实现方案是 client 实现在 guest OS 中，一般是集成在 virtio 驱动上，server 端实现在 qemu 中，也可以实现在各种数据面中，如 OVS，Snabbswitch 等虚拟交换机。

如果使用 qemu 作为 vhost-user 的 server 端实现，在启动 qemu 时，我们需要指定 -mem-path 和 -netdev 参数，如：

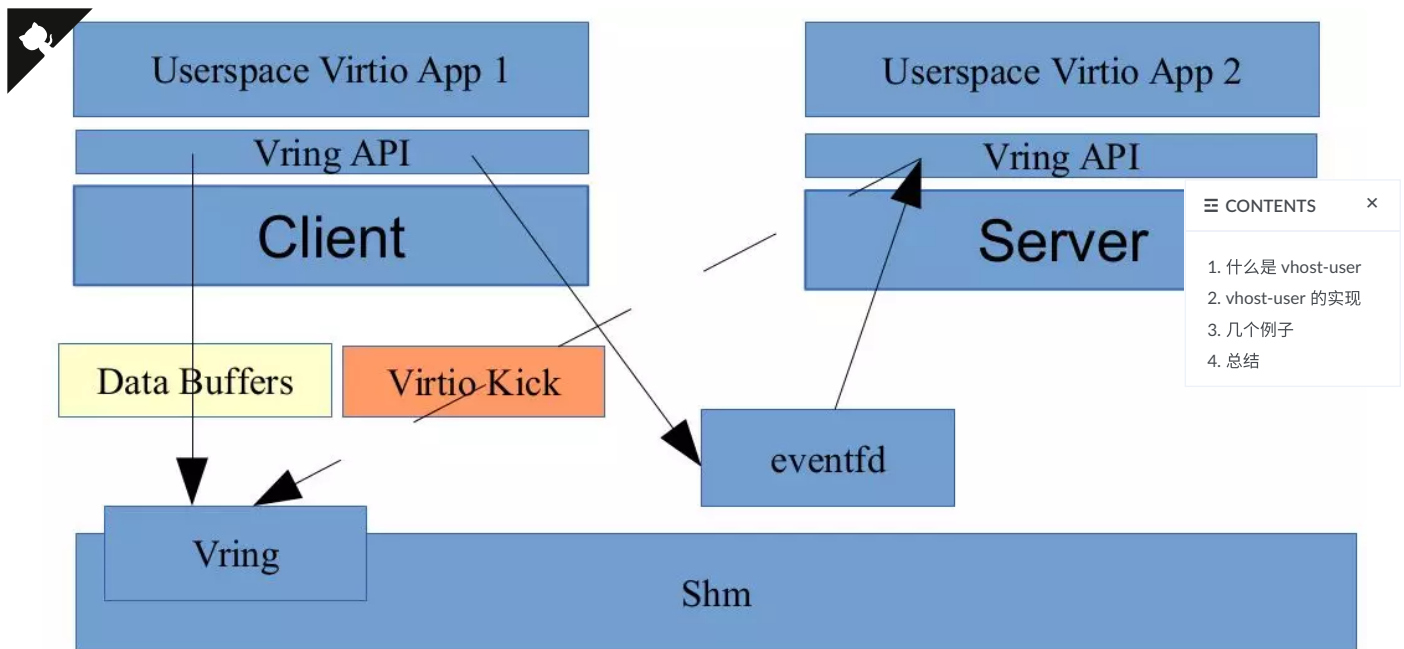
```
$ qemu -m 1024 -mem-path /hugetlbfs,prealloc=on,share=on \
-netdev type=vhost-user,id=net0,file=/path/to/socket \
-device virtio-net-pci,netdev=net0
```

Copy

指定 -mem-path 意味着 qemu 会在 guest OS 的内存中创建一个文件，share=on 选项允许其他进程访问这个文件，也就意味着能访问 guest OS 内存，达到共享内存的目的。

-netdev type=vhost-user 指定通信方案，file=/path/to/socket 指定 socket 文件。

当 qemu 启动之后，首先会进行 vring 的初始化，并通过 socket 建立 C/S 的共享内存区域和事件机制，然后 client 通过 eventfd 将 virtio 事件通知到 server 端，server 端同样通过 eventfd 进行响应，完成整个数据交互。



几个例子

开源社区中实现了一个项目 Vapp，主要是用来测试 vhost-user 的 C/S 模式的，github 地址如下：

<https://github.com/virtualopensystems/vapp.git>

使用：

```
$ git clone https://github.com/virtualopensystems/vapp.git
$ cd vapp
$ make
// 运行 server 端
$ ./vhost-s ./vhost.sock
// 运行 client 端
$ ./vhost-q ./vhost.sock
```

Copy

通过以上步骤，就可以启动 vhost-user 的 C/S 模式。

另外还有例子就是集成在虚拟交换机 Snabbswitch 上的 vhost-user，通过以下方式获得 vhost-user 分支：

```
$ git clone -b vhostuser --recursive https://github.com/SnabbCo/snabbswitch.git
$ cd snabbswitch
$ make
测试：
$ sudo src/snabbswitch -t apps.vhost.vhost_user
```

Copy

还有例子就是 qemu 上的实现，这也是最原早的实现，同样通过以下方式来获得使用：

```
$ git clone -b vhost-user-v5 https://github.com/virtualopensystems/qemu.git
$ mkdir qemu/obj
$ cd qemu/obj/
$ ../configure --target-list=x86_64-softmmu
$ make -j
```

Copy

除此之外，还有很多的实现，如 OVS 和 DPDK 上都有实现，这实际上是集成了 vhost-user 的通用 API。

总结

virtio, vhost, vhost-user 是基于场景和性能而提出的三种 guest 和 host 之间的通信方案，三种方案，各有优劣。

vhost-user 用在很多数据面之上的进程间通信，效率高。

我的博客即将入驻“云栖社区”，诚邀技术同仁一同入驻。

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”



CONTENTS

1. 什么是 vhost-user

2. vhost-user 的实现

3. 几个例子

4. 总结

stay hungry stay foolish ----jobs 希望多多烧香！

分类: 云计算, 虚拟化
标签: 云计算, 虚拟化, KVM

推荐 0

赞赏

收藏

反对 0

« 上一篇: DPDK 全面分析

» 下一篇: Docker 基础技术之 Linux namespace 详解

posted @ 2018-02-06 12:37 CloudDeveloper 阅读(5936) 评论(0) 编辑 收藏

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。

