



DPDK virtio-user

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」即可领取，欢迎大家关注，二维码文末可以扫。

CONTENTS

1. virtio_user 用于容器网络
2. virtio_user 作为 exception p...

virtio-user 是 DPDK 针对特定场景提出的一种解决方案，它主要有两种场景的用途，一种是用于 DPDK 应用容器对 DPDK v16.07 开始支持的；另一种是用于和内核通信，这是 DPDK v17.02 推出的。

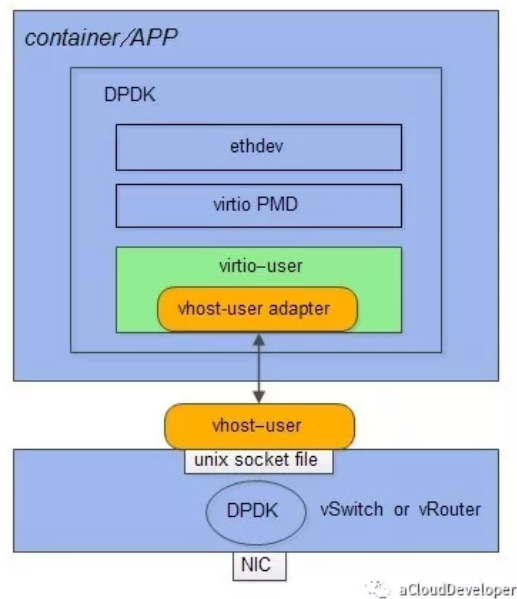
virtio_user 用于容器网络

我们知道，对于虚拟机，有 virtio 这套半虚拟化的标准协议来指导虚拟机和宿主机之间的通信，但对于容器的环境，直接沿用 virtio 是不行的，原因是虚拟机是通过 Qemu 来模拟的，Qemu 会将它虚拟出的整个 KVM 虚拟机的信息共享给宿主机，但对于 DPDK 加速的容器化环境来说显然是不合理的。因为 DPDK 容器与宿主机的通信只用得到虚拟内存中的大页内存部分，其他都是用不到的，全部共享也没有任何意义，DPDK 主要基于大页内存来收发数据包的。

所以，virtio_user 其实就是在 virtio PMD 的基础上进行了少量修改形成的，简单来说，就是添加大页共享的部分逻辑，并精简了整块共享内存部分的逻辑。

有兴趣可以对照 /driver/net/virtio 中的代码和 DPDK virtio_user 代码，其实大部分是相同的。

从 DPDK 的角度看，virtio_user 是作为一个虚拟设备（vdev）来加载的，它充当的是一个 virtio 前端驱动，与之对应的后端通信驱动，是用户态的 vhost_user，在使用的时候，我们只需要定义好相应的适配接口即可，如下：



vhost 和 vhost_user 本质上是采用共享内存的 IPC 方式，通过在 host 端创建 vhost_user 共享内存文件，然后 virtio_user 启动的时候指定该文件即可，如：

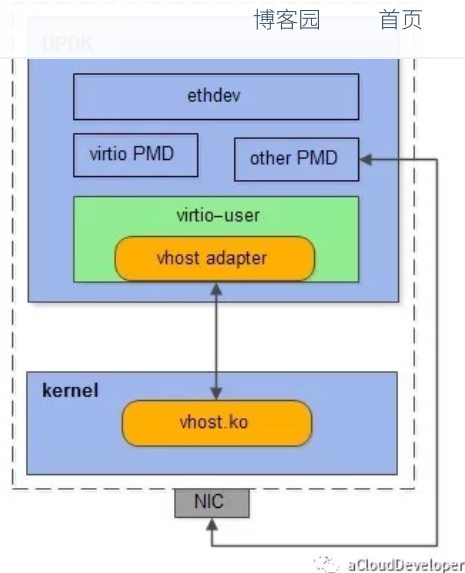
- 1) 首先创建 vhost_user 共享内存文件
--vdev 'eth_vhost_user0,iface=/tmp/vhost_user0'
- 2) 启动 virtio_user 指定文件路径
--vdev=virtio_user0,path=/tmp/vhost_user0

Copy

virtio_user 作为 exception path 用于与内核通信

virtio_user 的一个用途就是作为 exception path 用于与内核通信。我们知道，DPDK 是旁路内核的转包方案，这也是它高性能的原因，但有些时候从 DPDK 收到的包（如控制报文）需要丢到内核网络协议栈去做进一步的处理，这个路径在 DPDK 中就被称为 exception path。

在这之前，已经存在几种 exception path 的方案，如传统的 Tun/Tap，KNI（Kernel NIC Interface），AF_Packet 以及基于 SR-IOV 的 FI-Bifurcation。这些方案就不做过多介绍了，感兴趣的可看 DPDK 官网，上面都有介绍。



CONTENTS

1. virtio_user 用于容器网络
2. virtio_user 作为 exception p...

和容器网络的方案使用 vhost_user 作为后端驱动一样，要使得 virtio_user 和内核通信，只需加载内核模块 vhost.ko，让它充当的是 virtio_user 的后端通信驱动即可。

所以，我们可以看到，其实这两种方案本质上是一样，只是换了个后端驱动而已，这也是 virtio 的优势所在，定义一套通用的接口标准，需要什么类型的通信方式只需加载相应驱动即可，改动非常少，扩展性非常高。

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/8971710.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：虚拟化，云计算，DPDK

标签：虚拟化，云计算，DPDK

推荐 0

赞赏

收藏

反对 0

« 上一篇：Docker 基础技术之 Linux cgroups 详解

» 下一篇：Docker 跨主机网络方案分析



≡ CONTENTS



1. virtio_user 用于容器网络
2. virtio_user 作为 exception p...

