

技术贴：利用DPDK加速容器网络

原创 DPDK开源社区 2016-11-02

作者 谈鉴锋



↑↑↑ 点击蓝字，轻松关注

问题的提出

随着应用场景的不断增加，用户对容器网络的要求也越来越高。这些场景包括：

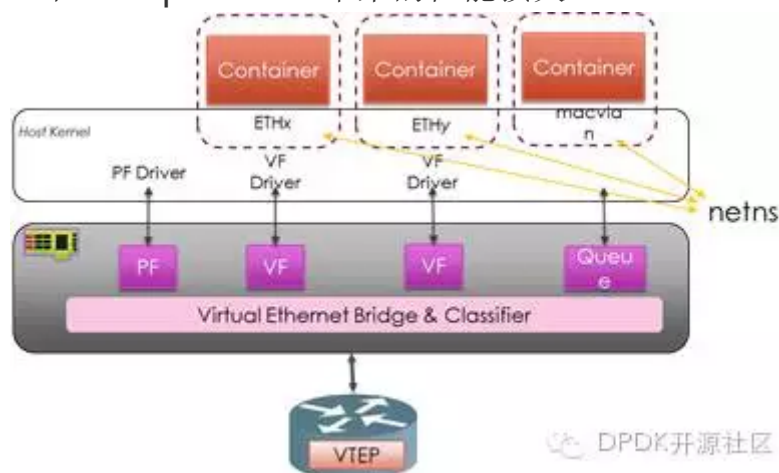
- a. 基于容器的VNF，如vRouter、vFirewall等；
- b. 数据中心里面的网络节点上的各种服务，如LB、VPN等；
- c. 需要高性能网络的关键应用服务，如Key-Value数据库。

在这些场景下对容器网络的要求包括：

- a. 从性能角度保证高带宽、低时延，并且尽可能少的网络抖动；
- b. 从安全角度保证隔离性（特别是多租户环境下）；
- c. 从服务质量角度保证对不同容器提供QoS差分服务。这里看到，不同于传统容器网络里常常被提及的挑战（如服务发现、IP地址分配），我们这里主要关心的还是数据面。

问题的分析

一直以来，容器网络都是构建在整个内核网络协议栈之上，并且利用net namespace来进行隔离的。受内核网络协议栈的限制，数据层面的性能优化空间很小。比如，a.避免使用vethpair和Linuxbridge或OVS，直接将VF或者NIC的queue包装为interface，分配给某个容器网络空间（如下图所示）；b.避免使用overlay网络来规避encapsulation和decapsulation带来的性能损失。



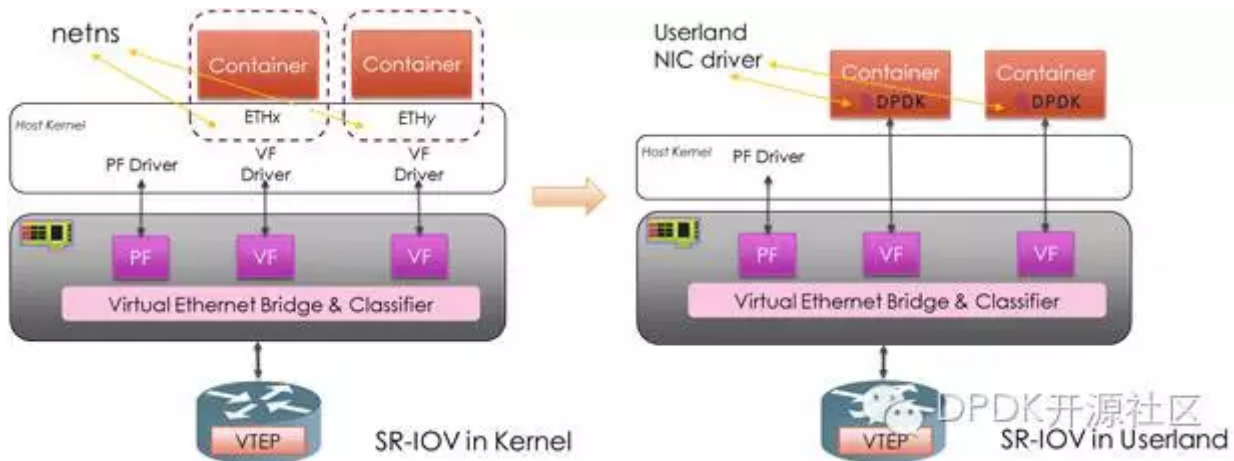
DPDK开源社区

问题的解决

这些优化技术带来的性能提升有限，于是我们考虑通过DPDK这样的用户态驱动来加速容器网络。DPDK利用hugepage、poll-mode、batching等技术大大提升了收发包能力。使用DPDK加速容器网络主要有两种使用模型：a.设备pass-through模型；b. vSwitch模型。

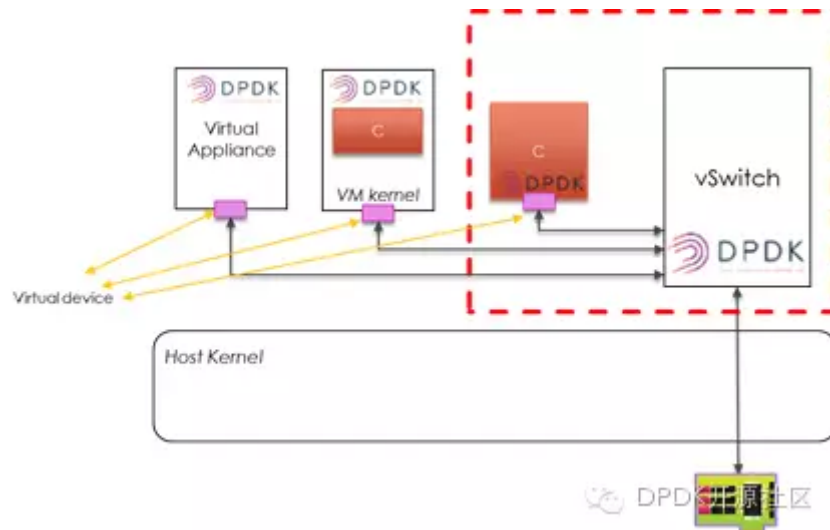
设备pass-through模型

如下图所示，所生成的Virtual Function（VF）将被用户态的DPDK来直接驱动。应用程序将通过DPDK直接从物理网络中收发包。整个数据通道仅仅需要物理网卡到内存之间的数据拷贝，并且最大程度发挥了DPDK的收发包优势。



vSwitch模型

在vswitch模型中，和虚拟机类似，容器中的网络流都汇集到经过DPDK加速的vSwitch（如OVS-DPDK、VPP）中。容器和vSwitch之间的通信通过DPDK里面虚拟设备virtio_user来进行（该设备主要提供了一种高速的进程间通信机制）。



下面是一些开放性问题，欢迎留言探讨。

- 1、怎样在非特权容器实例里使用DPDK？
- 2、Orchestrator在部署DPDK容器和一般容器何不同？
- 3、哪些应用适合跑在利用DPDK加速的容器里？

作者简介：

谈鉴锋，英特尔工程师，清华大学计算机硕士，主要从事于DPDK中半虚拟化网卡virtio的驱动开发，目前专注于使用DPDK加速容器网络以及推动容器的NFV技术。
关于本文详情，请点击[阅读全文](#)



微信ID：
DPDK开源社区



长按指纹识别二维
码关注

[阅读原文](#)

[投诉](#)