

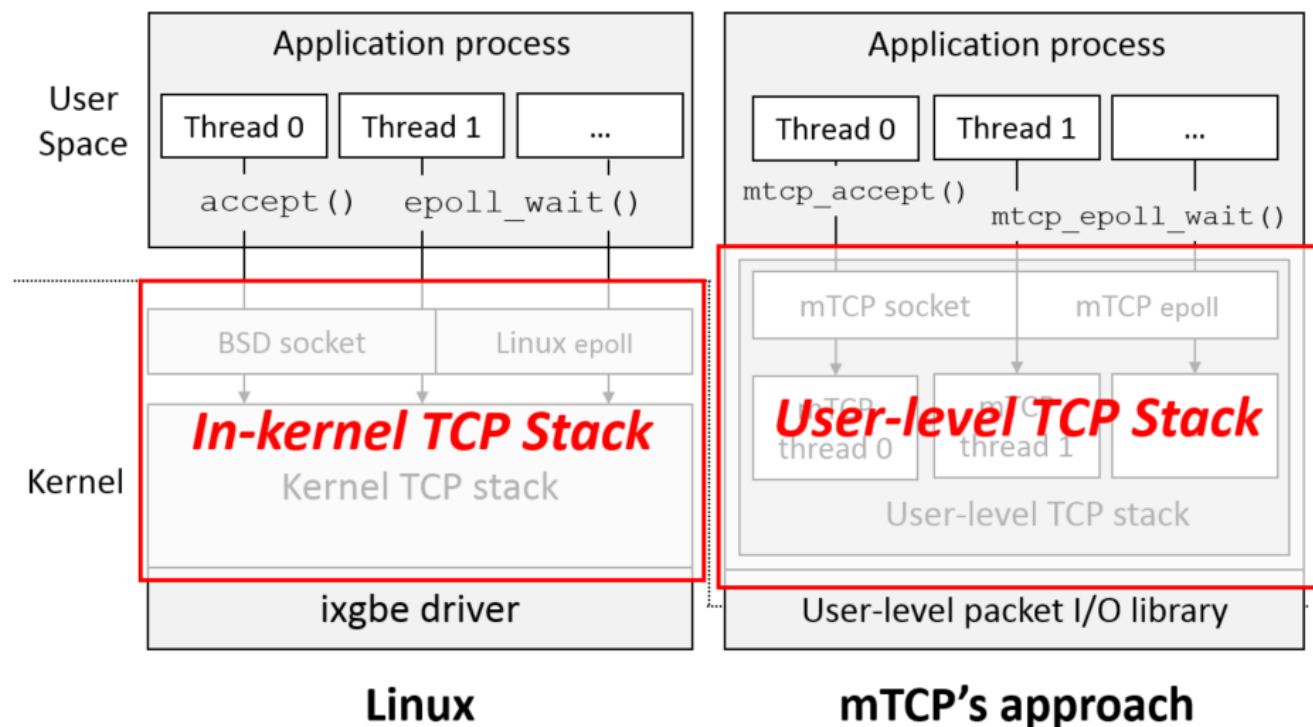
## DPDK NFV 性能提升

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，[点击查看](#)，欢迎大家关注，二维码文末可以扫。

本文是一篇翻译，翻译自<https://software.intel.com/en-us/blogs/2015/06/12/user-space-networking-fuels-nfv-pe>了，15年写的，但是文章总结了一些用户态的协议栈，很有学习参考的意义。

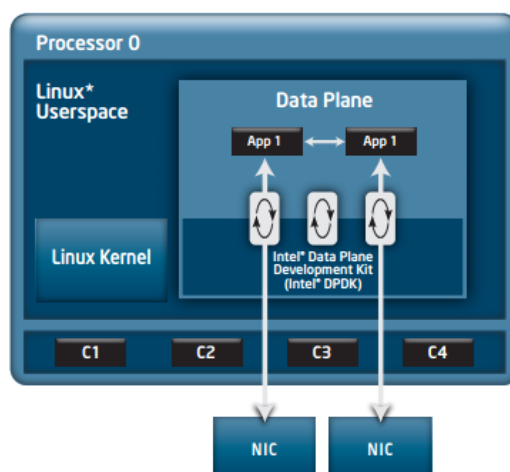
如今，作为一个网络空间的软件开发人员是非常激动人心的，因为工程师的角色随着这个世界的规则在逐渐改变。

过去这 15 年来，人们对高性能网络做了很多努力，网络模型也发生了很多改变，起初，数据包的收发都要推送到内核才能完成，现在，不用内核态的参与也可以完成。这种改变的背后是在力图解决以下的问题：1) 用户态和内核态上下文切换的开销；2) 软硬中断的开销；3) 数据拷贝的开销等等。



最近，很多人在讨论 mTCP——一个实现了用户态协议栈的开源库，这种技术在很大程度上就颠覆了传统的网络模型，使得网络收发包的效率大大提升。它充分利用了 CPU 的亲性和，共享内存，批处理等技术来实现高效的 I/O 事件。与之类似的技术也相继被提出。

实验表明，mTCP 这种用户态协议栈，相较原生的内核协议栈，在处理多种流行的应用时的性能得到较大提升，如 SSLShader 提升了 33%，lighttpd 提升了 320%。





和任务，以及多核并行处理任务。这种设计避免了上下文之间的切换开销，可以显著降低数据传输的延迟和 CPU 使用，同时提高吞吐量和带宽。另外，提供一种运行到完成（run-to-completion）的模型能够让不同的核独立并行地完成不同的任务。

随着网络的发展，我们现在看到了大量的开源项目，都在将内核协议栈移到用户空间来做。它们的做法存在一些区别。有些项目是从零开始开发的，而其他很多项目则是基于 FreeBSD 的来做，这主要是因为 FreeBSD 的协议栈具有“最健壮的网络堆栈”的声誉。很多存储解决方案也是采用的 FreeBSD 来作为其核心操作系统。当然，Linux 协议栈也是可以采用的。

这些用户态协议栈怎么做到绕过内核的，这就离不开 DPDK 的支持。利用 DPDK，用户态协议栈可以创建一个中断区直接映射到用户空间，然后利用协议栈的特性来管理 TCP/IP 数据包的处理和传输。

DPDK 还可以作为一些 vSwitch（虚拟交换机）的加速器，这些 vSwitch 包含 OpenFlow 协议的完整实现，以及与 Openstack Neutron 的整合。

下面，我们收集了一些发现的开源项目，无论你决定使用一个 vSwitch 还是一个完整的网络协议栈，网络开发人员都有很多选择，可以将应用程序移到用户空间，并在多核系统上扩展性能。

## DPDK-Enabled vSwitch：

### OVS

Open vSwitch 是一个多核虚拟交换机平台，支持标准的管理接口和开放可扩展的可编程接口，支持第三方的控制接入。

<https://github.com/openvswitch/ovs>

### Lagopus

Lagopus 是另一个多核虚拟交换的实现，功能和 OVS 差不多，支持多种网络协议，如 Ethernet，VLAN，QinQ，MAC-in-MAC，MPLS 和 PBB，以及隧道协议，如 GRE，VxLan 和 GTP。

<https://github.com/lagopus/lagopus/blob/master/QUICKSTART.md>

### Snabb

Snabb 是一个简单且快速的数据包处理工具箱。

<https://github.com/SnabbCo/snabbswitch/blob/master/README.md>

### xDPd

xDPd 是一个多平台，多 OpenFlow 版本支持的开源 datapath，主要专注在性能和可扩展性上。

<https://github.com/bisdn/xdpd/blob/stable/README>

## 从零开发的用户空间协议栈套件：

### mTCP

mTCP 是一个针对多核系统的高可扩展性的用户空间 TCP/IP 协议栈。

<https://github.com/eunyoung14/mtcp/blob/master/README>

### Mirage-Tcpip

mirage-tcpip 是一个针对 MirageOS 这种“库操作系统”而开发的一个用户态网络协议栈，开发的语言是 OCaml。

<https://github.com/mirage/mirage-tcpip>

### lwIP

lwIP 针对 RAM 平台的精简版的 TCP/IP 协议栈实现。

<http://git.savannah.gnu.org/cgit/lwip.git/tree/README>

## 移植版的用户空间协议栈套件：

### Arrakis

针对多核系统的用户空间 OS，移植于 lwIP。

[https://github.com/UWNetworksLab/arrakis/blob/master/README\\_ARRAKIS](https://github.com/UWNetworksLab/arrakis/blob/master/README_ARRAKIS)

### libuinet

用户空间的 TCP/IP 协议栈，移植于 FreeBSD。

<https://github.com/pkelsey/libuinet/blob/master/README>

### NUSE (libos)

一个基于 Linux 的库操作系统，移植于 Linux。

<https://github.com/libos-nuse/net-next-nuse/wiki/Quick-Start>

### OpenDP

一个针对 DPDK TCP/IP 协议栈的数据面，移植于 FreeBSD。

<https://github.com/pendp/dpdk-odp/wiki>

#### CONTENTS

- 0.1. DPDK-Enabled vSwitch:
- 0.2. 从零开发的用户空间协...
- 0.3. 移植版的用户空间协议...
- 1. 总结:





<http://www.openonload.org/download/openonload-201205-README.txt>

## OSv

一个针对虚拟机的开源操作系统。移植于 FreeBSD。

<https://github.com/cloudius-systems/osv/blob/master/README.md>

## Sandstorm

一个针对个人服务器安全的开源网络平台，移植于 FreeBSD。

<https://github.com/sandstorm-io/sandstorm/blob/master/README.md>

## 总结：

1、这篇文字的亮点在于总结了当前阶段业界出现的一些用户空间协议栈，对于文章标题提到的 NFV 在文中则只字未提，但其实意思也很明了了。用户空间的协议栈是随着硬件技术的发展，以及新鲜应用场景应运而生的，换句话说，对于像 NFV 这种对性能要求比较高的场景，采用用户态的协议栈是比较合适的。

2、文中是 2015 年写的，这意味着到现在为止，肯定出现了很多比上面总结还要多的方案，其中比较出名的有 SeaStar 和 腾讯开源的 F-Stack，后面找机会再进行详述，敬请期待吧。

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/8324492.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, DPDK

标签：云计算, OVS, DPDK

推荐 1

赞赏

收藏

反对 0

« 上一篇：virtio 简介

» 下一篇：vhost：一种 virtio 高性能的后端驱动实现

posted @ 2018-01-21 17:24 CloudDeveloper 阅读(1849) 评论(2) 编辑 收藏

### 评论列表

#1楼 2018-09-17 17:13 王小东大将军



Hi, 请问关于 云平台架构组件的书籍, 有推荐的吗?

支持

0



2018-09-17 17:34 CloudDeveloper

公众号：Linux云计算网络

博主：王小东大得年

OpenStack是最好的资料

博客园

首页

新随笔

联系

订阅

管理

支持(0) 反对(0)

≡ CONTENTS

×

- 0.1. DPDK-Enabled vSwitch:
- 0.2. 从零开发的用户空间协...
- 0.3. 移植版的用户空间协议...
- 1. 总结:

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。

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

