

OPEN SOURCE NETWORKING DAYS

Rebuild Cloud Network Protocol Stacks via DMM

DMM: 重构云网络协议栈生态

Waterman Cao (曹水), Huawei

云网络挑战：面向应用的差异化网络传输



- VR/AR 类的应用带宽需求是普通视频的5~10倍。例如单眼8K的360度全景VR需要5Gbps的实时带宽，交互时延在20ms以内。
- 自动驾驶中为了避免相撞网络延时越短越理想，4G 时延在100 毫秒，5G 时延将在5毫秒以内。



云网络挑战:Linux内核的协议栈更新缓慢



- 在过去的25年, Linux内核仅集成了不到5个传输协议栈。
- IEFT在2013年发布了MPTCP标准, 但直到目前还没有进入Linux Kernel。





行业现状：多厂家积极布局云网络协议栈 内核态 → 用户态

用户态

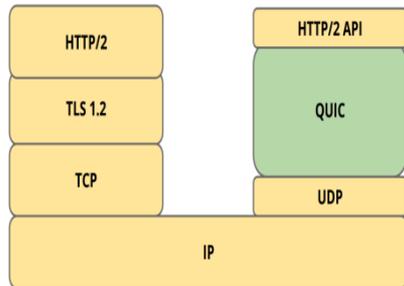
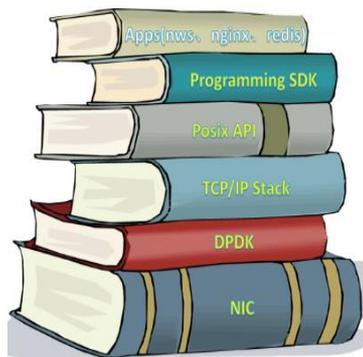
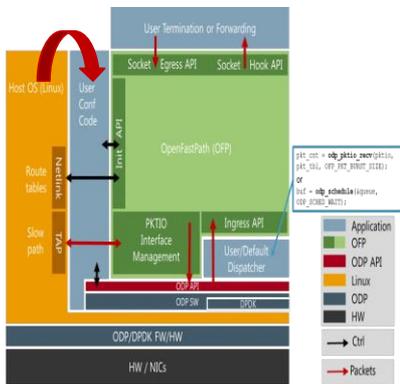
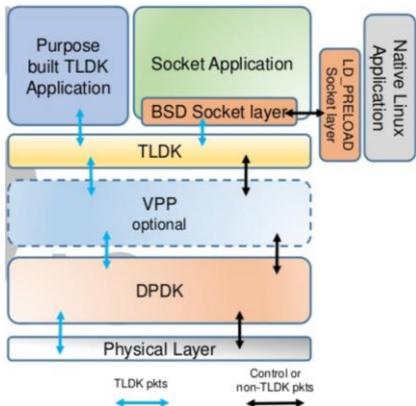
内核态

行业1: 2016.10.12 Cisco主推的fd.io的TLDK

行业2: 2015.12.2 诺基亚推出OpenFastPath

行业3: 2017.4.22 腾讯推出的F-Stack

行业: 2013 谷歌QUIC尝试替代TCP/IP



特点:

1. 基于矢量并行处理的底层IO;
2. 支持应用无感适配;但不支持多协议,多实例

特点:

1. 重点面向转发场景, 移植libuinet TCPIP, 对IP、路由、Socket API进行了优化。

特点:

1. 高性能的用户态TCPIP协议栈, 移植FreeBSD 11.01, 已集成nginx、redis等常用应用。

特点:

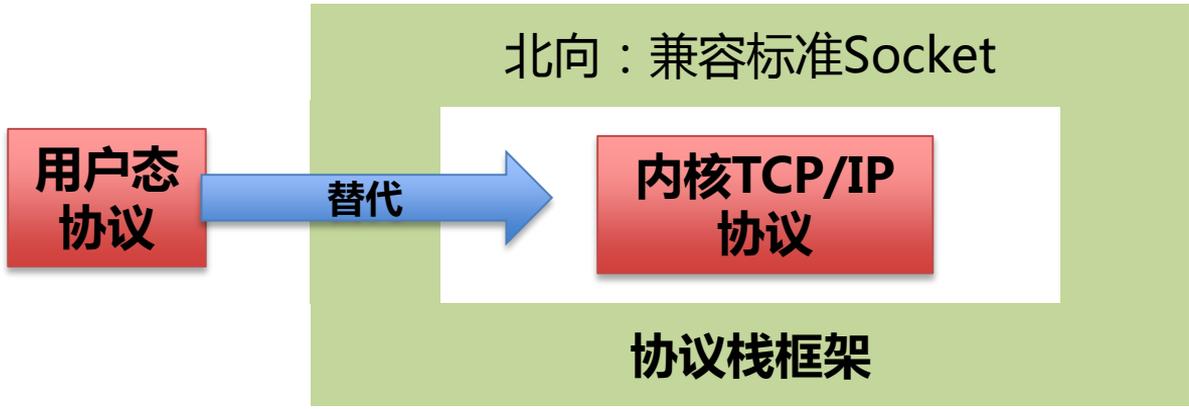
1. 平台无关:纯协议/算法变革, 与底层平台无关, 依托内核(非用户态);
2. 新协议:颠覆TCP, 需要双端部署(生态/标准)。

共性:对TCP协议本身未作优化;均以Posix Socket作为北向数据接口;支持DPDK为代表的用户态IO, 绕过内核



Possible Solution: 可按需替换协议栈

- 根据应用需求进行编排的插件化协议栈框架
- 支持不同类型的协议、协议栈实例等同时处理, 服务不同的应用
- 支持按需集中编排和对端协商



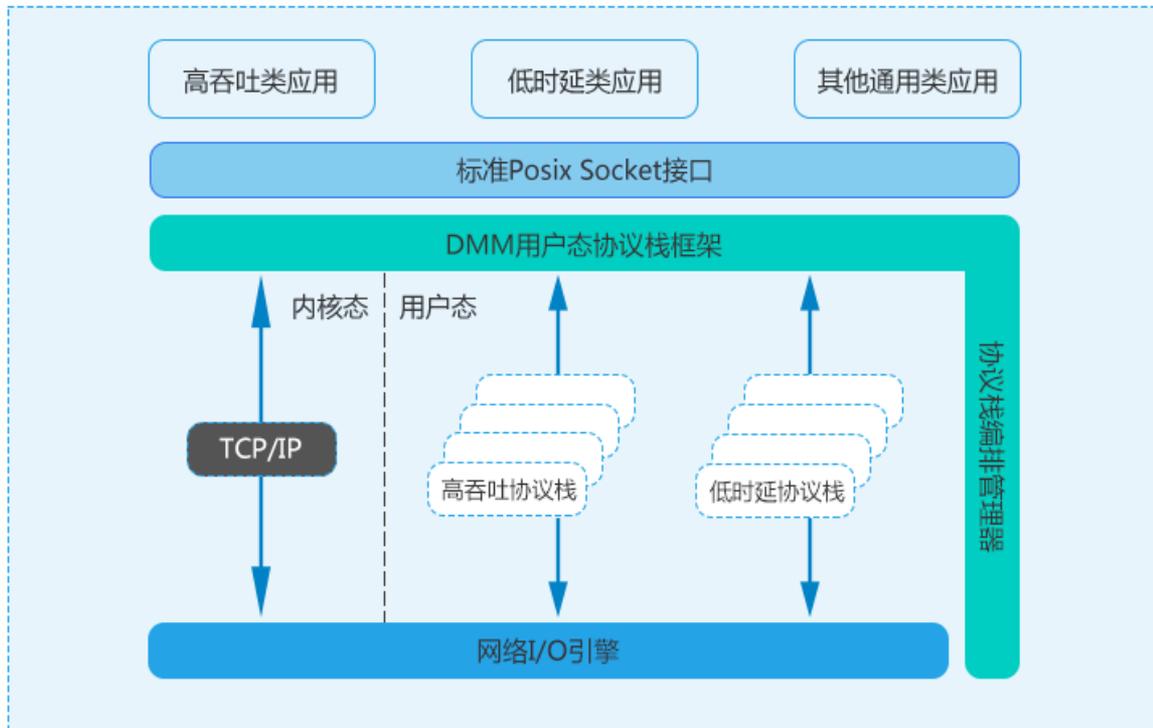
DMM:全新的开源协议栈框架



DMM (Dual Mode, Multi-Protocols, Multi-Instances)

DMM是一个全新的开源协议栈框架，支持标准POSIX socket API，避免逐个应用的适配。

- 支持双态协议栈：同时支持内核态协议和用户态协议栈。
- 支持多种不同的用户态协议栈。
- 配合框架可以实习协议栈的多实例化。
- 底层可以适配不同的IO驱动，DPDK/NetMap/..
- 根据应用对QoS需要对于网络协议进行编排，选择合适的协议栈提服务。



DMM框架是“柜子” 协议栈是“抽屉” 可按需抽屉式替换



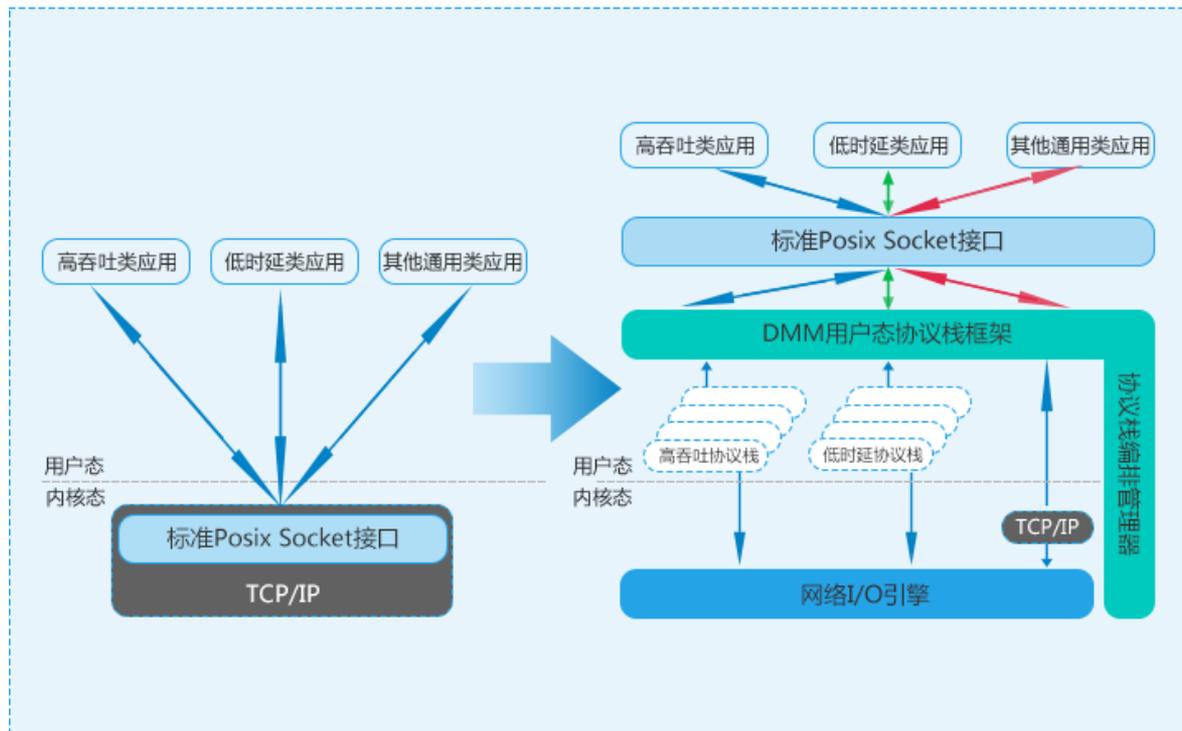
内核态TCP/IP协议栈问题：

- 无法满足不同应用的差异化性能需求
- Best-effort, 无SLA保障



DMM协议框架(双态多协议多实例)优势：

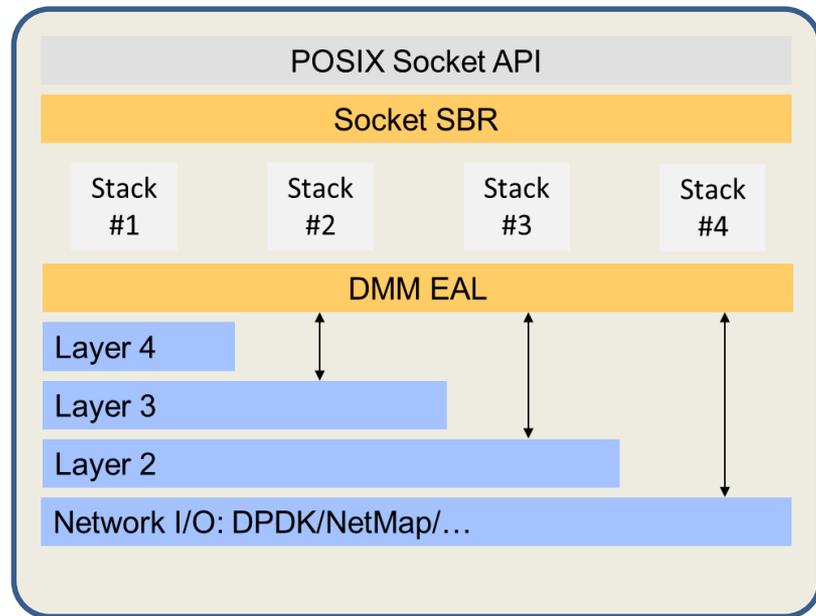
- 根据应用需求选择、组合不同协议栈，更好支持应用SLA
- 支持双端协商



DMM为协议栈开发者带来了什么



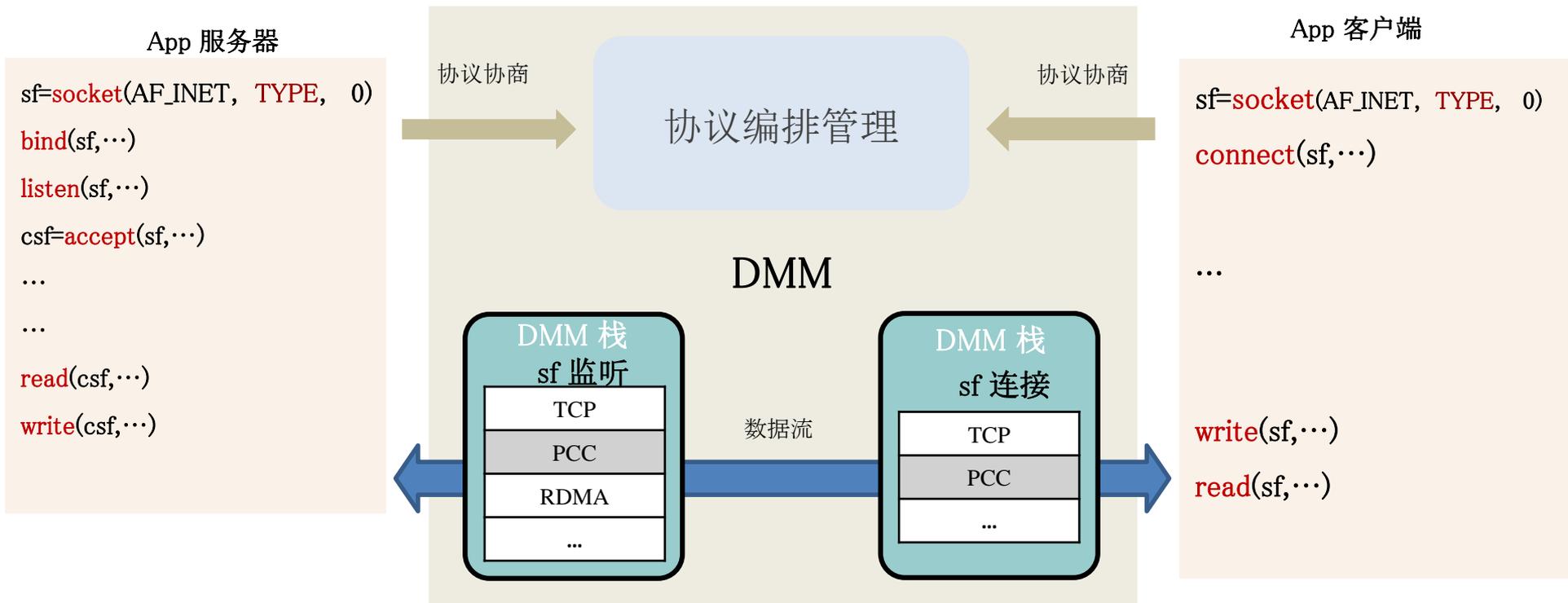
- 提供了一个用户友好的接口支持第三方协议栈接入
 - ✓ 灵活的 I/O APIs, including NIC/12/13/14 APIs
 - ✓ SBR (socket BRidge) 机制, 拦截socket 调用并重定向到相关协议栈接口
- 加速新协议栈的创新与开发
 - ✓ 允许协议栈开发者能够专注于协议的核心功能
 - ✓ 便于部署
 - ✓ 便于对接应用, 已经相关的测试





DMM为协议使用者带来什么

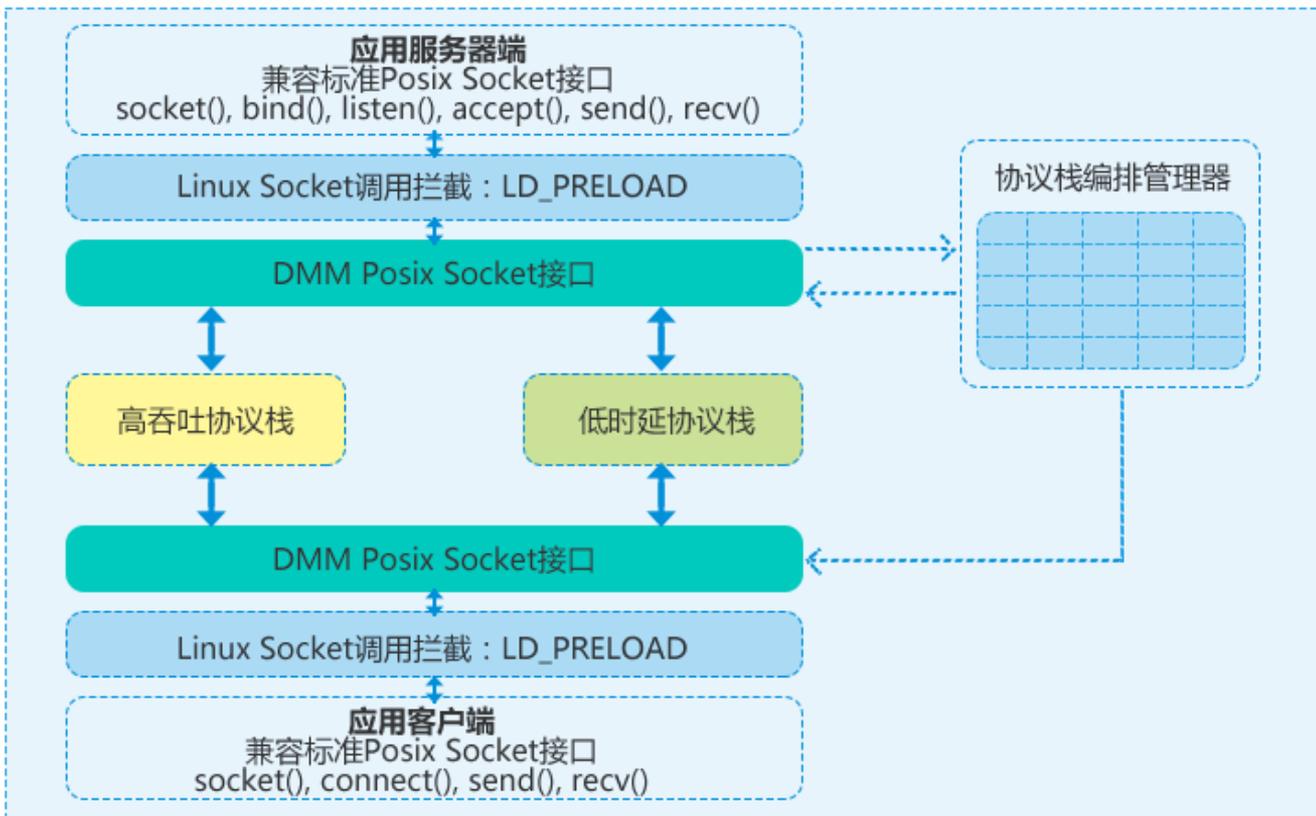
DMM 最大程度降低了对应用的影响



DMM为协议使用者带来什么



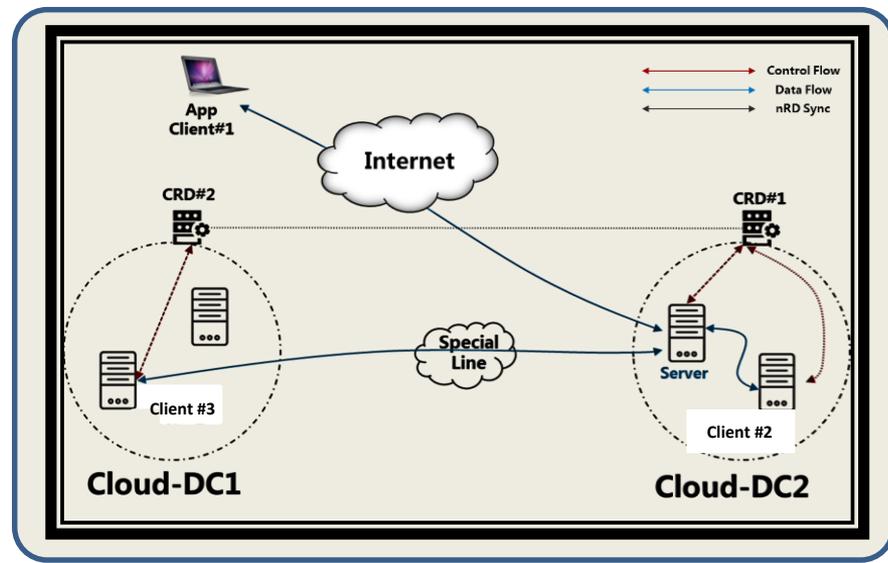
DMM 为用户提供协议路由的服务



实例一：面向云网络的最优协议路由



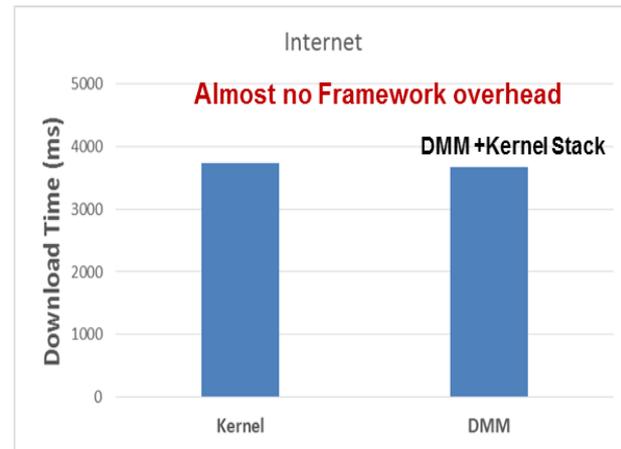
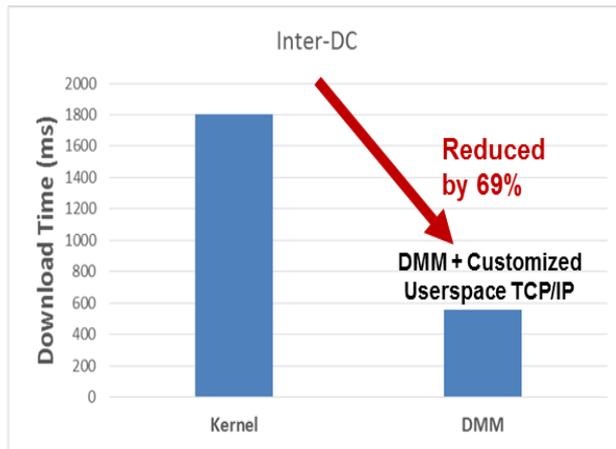
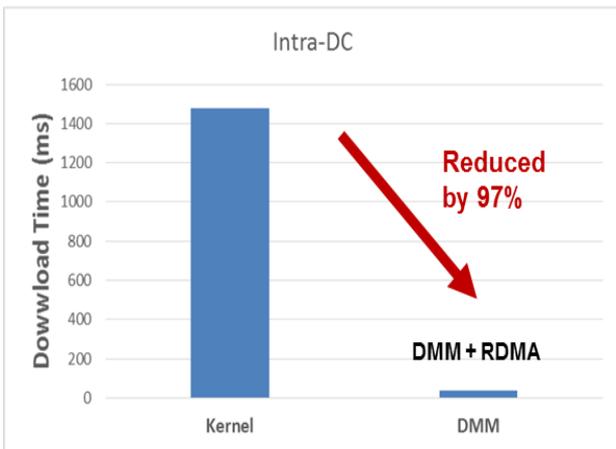
- 典型的场景：在不用的网络环境中使协议路由选择最优的网络协议。
- 文件同步应用
 - 3 个客户端
 - 分布部署在互联网(Client #1)、跨数据中心(Client #2)、同一数据中心内(Client#3)
 - 服务端（部署在数据中心）



实例一：面向云网络的最优协议路由



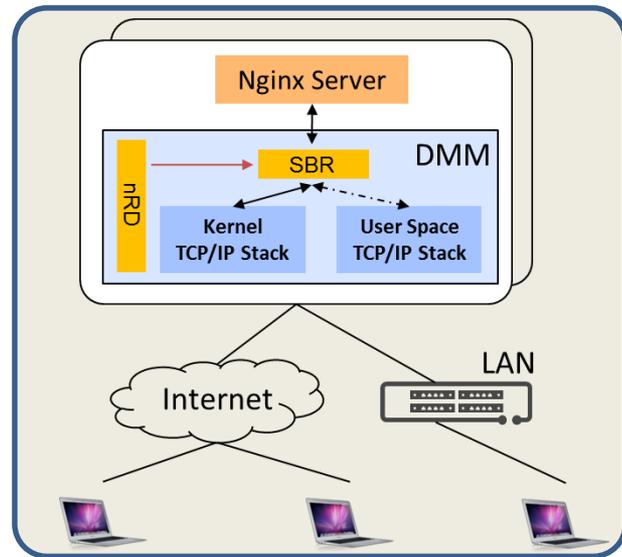
- 效果的比较：
 - Client #1: 无DMM: 只能使用内核提供的TCP/IP协议栈
 - Client#2 和 Client #3 部署DMM: 自动协商最合适的网络协议栈(可以分别选择 用户态TCP/IP协议栈, RDMA)



实例二: 支持双态协议栈的Nginx 服务器



- Nginx Application
 - ✓ 内核态协议栈 与 用户态协议栈
- DMM 协议路由策略:
 - ✓ Internet Connection --->内核态协议栈
 - ✓ LAN Link --->用户态协议栈

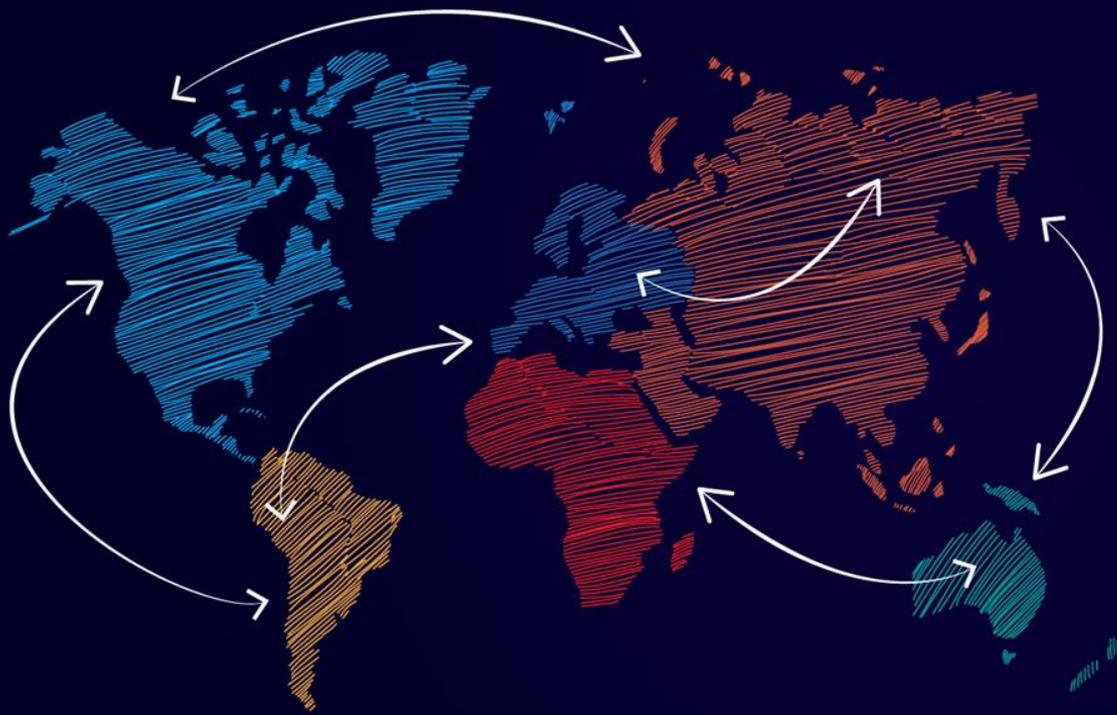


	Kernel	NSUE	DMM
Robustness	✓	×	✓
Performance	×	✓	✓
Customizability	×	✓	✓
Reliability	✓	×	✓

DMM 18.07 released



- 目前DMM 是open source software, licensed as Apache, 于2018年3月加入FD.io 社区。
- 最新发布18.07 版本
 - 该版本支持 pipeline 和 RTC模式的协议栈
 - 同时还内嵌两个协议栈LWIP stack (pipeline mode) 和rSocket stack (RTC mode) 支持RDMA 场景。
- 欢迎加入 DMM 社区, 您可以从下面获取更多信息:
- DMM 项目首页: <https://wiki.fd.io/view/DMM>
- DMM 代码: <https://git.fd.io/dmm> 或 <https://github.com/Huawei/DMM.git>
- 邮件列表: dmm-dev@lists.fd.io



Question ?

“欢迎加入DMM”

OPEN SOURCE NETWORKING DAYS

