



用 DPDK 和 VPP 优化 5G 网络用户面数据报文转发

朱宜斌(中兴通讯股份有限公司 江苏 南京 210012)

摘 要: 5G 作为各行业数字化转型的关键基础技术, 用户面需要提供高性能的数据转发能力。根据 5G 用户面数据转发的特点, 通过分析 DPDK 技术以及 VPP 技术的原理, 将两种技术相结合, 实现了 5G 用户面网元对数据报文高速转发, 并在商用网络上得到使用和验证。这种方法可以扩展到需要高速处理数据报文的其他网络设备中。

关键词: 5G; 数据报文转发; DPDK; VPP

0 引言

5G 作为产业数字化转型升级的基础性技术, 近年发展迅猛。据中国工信部数据, 截至 2022 年第三季度, 中国 5G 移动电话用户数已超过 5 亿, 而多家专业咨询机构对移动市场预测, 到 2022 年年底, 全球 5G 用户数将超过 10 亿。由于 5G 在各垂直行业应用的优势更加明显, 我国通信及工业产业链各相关单位对 5G 专网建设一直积极参与, 工业互联网产业联盟(AII) 及中国通信标准化协会(CCSA) 也为 5G 在各垂直领域应用制定了标准^[1], 以期 5G 服务于各个行业能有标准化和规模化的过程。

5G 技术追求在 eMBB(大带宽高速率)、uRLLC(低延时) 和 mMTC(海量接入) 三个场景的极致性能^[2]。5G 网络设备将以更高的吞吐速率来转发数据。5G 在服务各行业专网的应用时, 业界对它的 uRLLC 和传输性能提出了更高的要求^[3-4]。因此, 各 5G 设备生产厂家都在积极寻求提升网络设备单位硬件(CPU、内存、网卡等) 的数据报文转发能力。5G 网络在研制和部署时, 采用了 NFV(Network Function Virtualization) 和 SDN(Software Define Network) 技术^[5]。5G 网络设备多数基于通用硬件和 Linux 系统构建, 通用硬件和 Linux 内核协议栈在处理数据报文转发方面都存在性能不足

的问题, 这让各厂家在提升 5G 设备性能时必须考虑通用硬件的限制以及克服 Linux 内核协议栈性能较低的缺点。

本文在通用服务器(Commercial Off - The - Shelf , COTS) 上使用 DPDK(Data Plane Development Kit) 和 VPP(Vector Packet Processing) 技术, 从网卡高速接收数据报文, 再在用户空间中进行批量化处理, 绕过 Linux 内核协议栈, 满足 5G 网络设备的数据转发高性能要求。

1 5G 网络及用户平面

3GPP 将 5G 网络设计为基于服务化的架构, 如图 1 所示^[6]。网络中控制平面各网元以服务化接口挂在总线上互相配合, 完成终端设备接入 5G 网络、实施各种切片及控制策略、访问数据网络以及完成流量时长计费复杂流程。5G 网络还被设计成用户平面(User Plane , UP) 功能与控制平面(Control Plane , CP) 功能分离, 以便对两个平面进行独立的扩展和灵活的部署。

5G 网络的用户平面主要完成 UE 通过 RAN 和 UPF 访问 DN 的数据转发, 涉及到 UE、RAN、UPF 以及部署在 N6 口的各种设备, 比如 LoadBalance 网元、

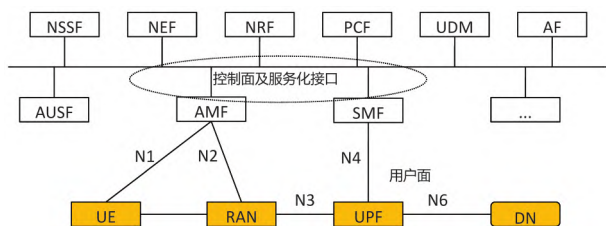


图1 5G核心网框架示意图

NAT/Firewall 网元、视频优化及加速网元等完成特定业务功能的网元。UPF 及这些业务功能网元在实现形态上基本都采用了在通用服务器上运行虚拟机或容器的模式。网络通信方面采用 25G 或 100G 高速网卡处理用户业务流数据。当业务功能网元较多时,业界采用编链的方式将每个特定业务流要经过的网元和 UPF 网元串接起来,或者采用将特定业务功能内置于 UPF 网元的方式部署。

2 DPDK 技术和 VPP 技术

2.1 DPDK 技术

DPDK 是 Intel 公司提供的数据包处理软件库,是为提升通用服务器的数据报文处理效能而专门开发的软件解决方案。

随着 NFV、SDN 等技术的演进,5G 时代的应用和网络设备更多的基于通用服务器构建。通用服务器的 CPU 需要兼顾各种应用的要求,不适合作为处理网络数据报文的网络处理器。同时为了处理数据报文而在通用服务器上增加专用的硬件加速器或 FPGA 也不经济。所以基于多核通用处理器这个硬件前提,如何提升其数据报文处理效率成为 5G 网络部署的一个关键技术点。DPDK 技术从软件的角度优化数据报文处理流程,不需要通用服务器在硬件做改动,而且对多核处理器有较好地适配,解决了通用服务器处理网络数据在硬件方面的后顾之忧。

5G 网元和应用多数基于 Linux 系统构建。Linux 系统下传统的网卡收发包处理流程在内核态和用户态之间切换,处理效率低。尤其 SDN 架构下,涉及到 Host OS 和 Guest OS 两个层面,Linux 操作系统处理后,报文效率下降更明显。DPDK 技术绕过 Linux 内核协议栈处理数据报文,全程运行在用户态空间,提升了虚拟化架构下数据报文处理效率^[7]。

基于 DPDK 的应用还可以充分利用 DPDK 框架特

性,极大减少数据报文的拷贝,进一步提升网络性能。

2.2 VPP 技术

矢量报文处理 (Vector Packet Processing, VPP) 技术是 Linux 基金会的开源项目 FD.io 中最核心的技术,其思想是通过将报文处理流程拆分成多个不同节点组成报文处理节点图,在每个节点上批量处理多个报文的方式来提高报文处理效率。这种节点图的处理方式同时让 VPP 技术具有很好的可扩展性,增加新类型业务时,只需在当前图中插入新的节点即可。相对于传统的一次处理一个报文, VPP 技术将底层硬件队列收到的多个包组成一个 Packet Vector 交给报文处理图中的对应节点来处理。这个包矢量被节点处理的过程中,指令和内存被第一个报文装载,后续报文可以直接使用,从而显著提升整个报文矢量的处理效率,整体上提升了单个报文处理的效率^[8-9]。

FD.io 项目描述的 VPP 技术框架提供了比较清晰的 4 个层次,如图 2 所示。图中 VPP Infra 为整个架构的基础框架层,提供内存管理、vector、ring、hash 等各种处理报文所要用到功能的原语。VLIB 为矢量处理库层,提供对报文矢量的处理函数及对应用的管理功能,比如 Graph Node 管理、多线程支持、数据包跟踪、计数器以及 debug CLI 功能等。VNET 为通用网络协议栈层,提供 I2/3/4 及其他网络协议的高效处理接口。Plugins 为应用插件层,使用者既可以复用 VPP 当前已经提供的插件,也可以遵循 VPP 规范加入自己开发的插件。

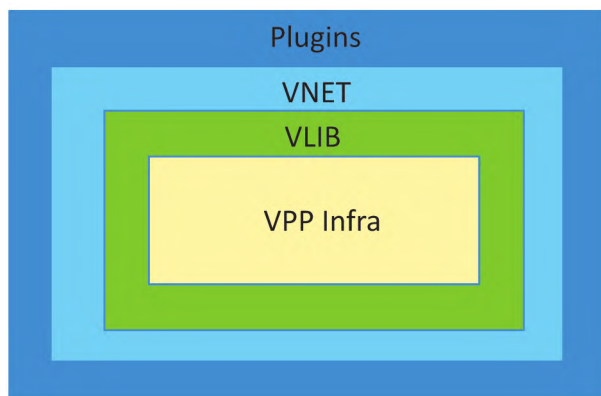


图2 VPP架构示意图

基于 DPDK 良好的性能和强大的底层处理能力, FD.io 项目经常将 DPDK 和 VPP 配合协作完成报文处理任务。在 DPDK 高性能收发报文能力支持下, VPP

的重点在优化 L2/3/4 报文上,旨在通过提高指令和缓存的命中数来增加转发平面吞吐量,同时使用模块化并行方式方便使用者扩展自己的应用。

FD.io 是开源项目,使用者遵循其开源协议可以对 VPP 架构源码进行定制化处理,以便适合具体项目的要求。

3 实现 5G 数据报文高速转发

3.1 5G 用户面报文处理特点和优化方案

5G 网络用户面承担了整个网络的用户数据转发,分析用户面路径上各个网元的报文处理工作流程,总结出其工作特点主要有:

(1) 各网元大多使用 COTS 硬件和 Linux 系统实现,采用 SDN 和 NFV 架构部署。

(2) 报文密度高。

(3) 大部分处理集中在 L2/3/4,UPF 网元对少量报文需要处理到 L7。

(4) UPF 网元对报文的处理最复杂,需要较好的扩展性。

针对 5G 网络用户面工作的以上特点,分析采用 DPDK 和 VPP 结合的方式能很好实现 5G 用户面的网络报文高性能转发,方案如图 3 所示。图中箭头代表数据报文出入方向。DPDK 在网卡初始化之后接管网卡基本的收发功能,作为 VPP 的入口和出口节点纳入 VPP 架构。DPDK 使用专用线程死循环接收网卡数据报文,然后通过 ring 送给 VPP 的 INPUT 节点。这样可以较好地转发高密度报文,并在 Linux 用户空间处理。VPP 在 INPUT 节点之后可以进一步根据报文的处理任务逐层分发到 L2/3/4 各个对应节点完成处理,处理完成之后交给 DPDK 发送出去。

网元将自身需要完成的流程组织成处理节点,再用插件的方式注册到 VPP 框架上,插入到节点图的合适位置,节点内逻辑完全由网元业务实现,与 VPP 框架能隔离。网元可以不使用 VPP 框架中的 VNET 层接口以及 Plugins 层的插件,而采用自身实现的插件,以进一步和 VPP 隔离,实现最大程度的灵活性。同时仅提供 VPP Infra 和 VLIB 两层的 VPP 架构将更为通用,可以更好地移植在各个网元上。

3.2 用 DPDK 和 VPP 优化 5G 用户面报文转发

5G 用户面在 N6 口必须部署 UPF 网元,可选部署 LoadBalance 网元、TCPO 网元、QUICO 网元和 FireWall/

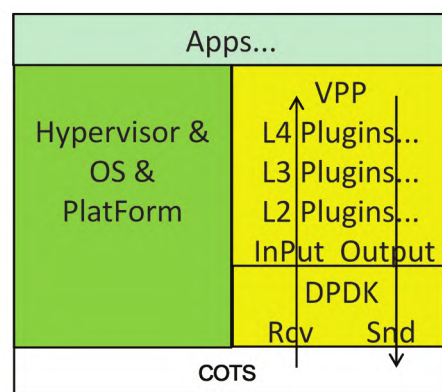


图3 DPDK 和 VPP 组合完成用户面报文转发 NAT 网元。

UPF 网元处理的主要报文类型为 IP 报文、GTP 隧道报文、L2TP 隧道和 IPSec 隧道报文等。这些报文以流的方式被归纳,并和对应的控制面处理策略关联起来,执行封装、解封装,执行数据包检测规则(PDR)和转发(FAR)规则,完成统计、控制、计费 and 识别等动作。所以 UPF 上的处理流程相对复杂,实现时需要将各 VPP 节点必需的控制面信息获取到,放置在 VPP 节点对应的处理内存表中,这样所有节点在处理对应流程时就能减少不必要的内存切换,做到处理流程指令和内存访问都能高效。

LoadBalance 网元主要为其他网元做负荷均衡,以实现按 L3 或 L4 的数据流的分担。LoadBalance 网元需要处理的数据报文量大,但处理逻辑相对简单,使用 DPDK + VPP 架构能获得非常好的转发性能。TCPO 网元、QUICO 网元和 NAT/FireWall 网元等业务功能网元都是针对某一特定协议报文的专项处理,处理逻辑相对集中,用 DPDK + VPP 架构能较好地完成扩展,且性能较高。也可以将 TCPO 等业务功能作为 UPF 网元内业务流程上的一个节点,减少 N6 口上网元间的流量提高整个网络的处理效率,但会带来 UPF 处理更复杂的代价,需要根据具体需求考虑。

以上实现方案经过实测,在网卡带宽不是瓶颈的情况下,在 LoadBalance 网元上少量 CPU 核就可以获得超过 20Gbps 的转发性能。而其他网元,尽管报文转发逻辑较为复杂,但相同数量的 CPU 也能获得 10 + Gbps 的转发性能。这些性能数据证明 DPDK 和 VPP 技术应用于 5G 网络用户面的可行性。



4 结束语

本文探讨了5G网络的部署模式和用户面对报文转发性能的要求,并实践了DPDK技术和VPP技术,实现了较好的转发性能,证明这两项技术可以促进5G网络更好地服务于各行业。随着5G网络在各垂直行业进一步深入部署,边缘计算的MEC、定制化的承载网络设备以及各种行业应用网关等,对网络性能有较高要求的设备也可以参考5G网络的方式,应用这两项技术以实现更好的转发性能。

参考文献

- [1] 黄颖,于青民,李宗祥,等. 5G在工业互联网领域的应用进展[J]. 移动通信, 2022, 46(1): 41-45
- [2] 黄珊. 5G关键技术及行业应用探索[J]. 数字通信世界, 2022(2): 30-32+35

(上接第42页)

参考文献

- [1] 刘昕,陈文华,吴汇波,等. 功放数字预失真线性化技术发展趋势与挑战[J]. 中国科学:信息科学, 2022, 52(04): 569-595
- [2] 马施榆,解程杰. 基于图卷积神经网络的射频功放数字预失真模型[J]. 数据通信, 2022(06): 12-15
- [3] 方俊. 基于增强型LSTM神经网络的5G射频功放线性化[J]. 无线通信技术, 2022, 31(02): 23-26
- [4] 杨宦. 基于Attention循环神经网络的功放数字预失真研究[D]. 电子科技大学, 2022. DOI: 10.27005/d.cnki.gdzku.2022.004252
- [5] 刘春秀,李军,许高明,等. 数字预失真器对5G功放记忆效应补偿能力的比较[J]. 数据通信, 2022(01): 9-13
- [6] 苗德华,解程杰. 5G宽带功放线性化平台开发与验证[J]. 无线通信技术, 2022, 31(01): 6-10
- [7] 周俊宇,许高明,韩栋,等. 基于数/模混合预失真技术的功放线性化[J]. 微波学报, 2022, 38(02): 91-94. DOI: 10.14183/j.cnki.1005-6122.202202019
- [8] 张强. 5G宽带功放的数字预失真技术研究[D]. 东南大学, 2021. DOI: 10.27014/d.cnki.gdnau.2021.000453

- [3] 范华敏. 5G时代行业专网的发展[J]. 邮电设计技术, 2022(2): 77-80
- [4] 杨芄,徐明伟,陈浩. 5G/后5G部署对互联网主干影响的分析与建模[J]. 通信学报, 2019, 40(8): 36-44
- [5] 张燕. 5G核心网虚拟化云资源池部署探讨[J]. 移动通信, 2019, 43(6): 10-15
- [6] 3GPP TS23.501 System Architecture for the 5G System; Stage 2 [17.3.0] [S]. 2021
- [7] 曾理,叶晓舟,王玲芳. DPDK技术应用研究综述[J]. 网络新媒体技术, 2020, 9(2): 1-8
- [8] 张宇巍,曾一,杨燕宁. 基于VPP的虚拟路由器数据平面加速方法[J]. 计算机系统应用, 2017, 26(10): 276-280
- [9] 潘恬,林兴晨,张娇,等. 基于高性能包处理架构VPP的带内网络遥测系统[J]. 通信学报, 2021, 42(3): 75-90

作者简介:朱宜斌,硕士研究生,中兴通讯股份有限公司工程师,主要从事电信云、移动核心网等方面的研发。■

- [9] 廖晓. 面向5G无线通信系统数字预失真设计与实现[D]. 湖南大学, 2021. DOI: 10.27135/d.cnki.ghudu.2021.002844
- [10] GOTTHANST, BAUDOUIN G, MBAYE A. Digital predistortion with advance/delay neural network and comparison with Volterra derived models [J]. 2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), Washington, DC, USA, 2014: 811-815. doi: 10.1109/PIMRC.2014.7136276
- [11] WATKINSB E, NORTH R. Predistortion of nonlinear amplifiers using neural networks [J]. Proceedings of MILCOM '96 IEEE Military Communications Conference, McLean, VA, USA, 1996, 1: 316-320. doi: 10.1109/MILCOM.1996.568636
- [12] LIU Z, HU X, XU L, et al. Low Computational Complexity Digital Predistortion Based on Convolutional Neural Network for Wideband Power Amplifiers [J]. in IEEE Transactions on Circuits and Systems II: Express Briefs, 2022, 69(3): 1702-1706, doi: 10.1109/TCSII.2021.3109973

作者简介:舒海燕(1998-),女,汉族,安徽六安人,就读于宁波大学计算机工程学院,通信工程专业硕士,研究方向为数字预失真、深度学习。■