

基于 DPDK 框架的多核实时通信系统架构

周晓玲^{1,2}, 俞春祥¹, 刘为^{1,2}

(1.中国电科新一代移动通信创新中心, 上海 200331; 2.中国电子科技集团公司第七研究所, 广东 广州 510310)

摘要: DPDK 是近期火热应用于通用处理器的数据面优化软件套件, 本文将分析其应用于实时通信系统, 基于 DPDK 的通用处理器下的实时通信系统架构在 5G 基站中的一种应用模式, 有别于传统 DSP、FPGA 的处理架构, 该架构也可借鉴应用于其他的实时通信系统。

关键词: DPDK; 5G; 多核; 实时通信处理架构

中图分类号: TP332

文献标识码: A

文章编号: 1004-7344(2021)39-0156-02

0 引言

传统通信系统中对于实时通信处理的部分一般是基于 DSP 和 FPGA 处理器来搭建系统完成实时信号处理。通用处理器更多是非实时系统的高层协议处理的应用选择。DPDK 最初是基于 x86 架构的通用处理器的平面开发套件, 其主要是为了加速优化网络包处理而生。本文将分析介绍引入 DPDK 技术后的通用处理器在实时通信系统架构中的应用; 首先介绍 DPDK 的主要技术原理、特点和基本应用, 其次介绍 5G ORAN 基站的前传接口, 最后介绍一种 5G 基站中基于 DPDK 的多核实时通信处理架构的框架和方法。

1 DPDK 简介和特点

什么是 DPDK? DPDK (Data Plane Development Kit) 是 Intel 的数据平面开发套件, 是为加速网络 I/O 所提出的高性能网络报文处理框架。当前火热的网络功能虚拟化, 将 DPDK 放在一个重要的基石位置。

DPDK 主要模块分解展示如图 1 所示, 其以基础软件库的形式为上层应用的开发提供一个高性能的基础 I/O 开发包。主要包括: 核心库、PMD 库、Classify 库、QoS 库等。它大量利用了有助于包处理的软硬件特性, 如大页、缓存行对齐、线程绑定、预取、NUMA、IA (Intel Architecture) 最新指令、Intel DDIO、内存交叉访问等。

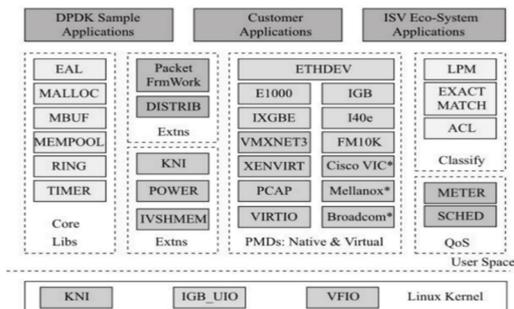


图 1 DPDK 的主要模块构成

核心库 (Core Libs): 提供环境抽象、大页内存、缓存池、定时器及无锁环等基础组件。

PMD 库: 提供全用户态的驱动, 以便通过轮询和线程绑定得到极高的网络吞吐量, 支持各种本地和虚拟网卡。

Classify 库: 支持精确匹配、最长匹配和通配符匹配, 提供常用包处理的查表操作。

QoS 库: 提供网络服务质量相关组件。

另外, DPDK 还提供如 PacketFramework 和 DISTRIB, 为搭建更复杂的多核流水线处理模型提供了基础的组件。

DPDK 主要思想和技术包括:

(1) 环境抽象层 EAL。

DPDK 的环境抽象层 EAL 负责对计算机底层资源 (如硬件和内存空间) 的访问, 并对提供给用户的接口实施了实现细节的封装。

(2) 并行计算机制。

DPDK 技术利用通用服务器平台的多核处理器来对网络数据包进行处理, 并行计算机制主要包含两种方式: ①提高处理器的指令并发度; ②多核并行处理网络数据包。

多核并行处理网络数据包是 DPDK 程序设计的基本思想。每个物理核都可以独立创建一个任务线程, 各个核之间的通信通过芯片内部总线来完成。多核并行处理会带来 Cache 一致性的问题。

DPDK 使用了 pthread 开发库, 若在系统中把相应的线程和 CPU 进行亲和性的绑定, 相应的线程尽可能使用独立的资源进行相关数据处理工作, 可减少或避免多核并行处理带来的 Cache 一致性问题。

(3) 内存池的管理。

Linux 使用 4kB 大小的分页来管理内存, 而 DPDK 使用 2MB 或 1GB 的大页来管理内存, 一个 TLB 表项可以指向更大的内存区域, 从而减少 TLB miss。

(4) 用户态网络驱动。

DPDK 抛弃了传统内核异步中断模式,采用 PMD 轮询或轮询混杂中断模式直接操作网卡的接收和发送队列,将报文直接拷贝到用户空间,不再经过内核协议栈。DPDK 用户态 I/O (UIO) 驱动技术为 PMD 提供支持。其主要功能是拦截中断,并重设中断回调行为,从而绕过内核协议栈后续的处理流程。

综上,DPDK 使用轮询为主来处理网络数据包,提供全用户态的驱动,拥有高效的内存管理机制,报文直接通过 DMA 传输到用户态处理,减少内存拷贝次数,节省大量的 CPU 中断时间和内存拷贝时间。

2 5G ORAN 基站的功能划分

4G LTE 基站的 BBU 和 RRU 的接口采用是 CPRI 协议。但 5G ORAN 基站的前传接口采用的是 eCPRI 协议^[1]。5G 具有大带宽、高吞吐量的特点,信道带宽要求从 LTE 的 20MHz 上升到 100MHz,甚至更宽 400MHz。且 5G 支持 Massive MIMO 技术,天线数和带宽的倍增对前传接口的吞吐量要求有了更高的要求。有别于 LTE 的基于 CPRI 协议的 Option 8 切分方式,5G 的 ORAN 基站可采用多种切分方式,降低系统对前传接口的吞吐量压力要求。比如采用 Option 7-2x 的切分方式,把 5G 协议栈的物理层的前端处理 (IFFT/FFT 等) 从 DU 端下沉到 RU 端。

5G 的前传接口采用的 eCPRI 协议是基于以太网的传输方式。随着网口速率的提升,传输速率已经从千兆以太网上升到 10G、25G/40G。这给前传接口的传输提供了更大的能力,也给 DPDK 的网络包加速优化有了更多的用武之地。

3 基于 DPDK 的实时通信系统处理框架

DPDK 作为网络报文的高性能处理工具,其主要研究应用方向为:加速实际网络应用、优化 NFV 中 x86 平台网络处理性能等。回顾 DPDK 的设计框架,可以把 DPDK 运用在高性能计算上。5G 基站的协议栈物理层的处理也具有高性能计算要求的特点。

利用 DPDK 作为通用处理 CPU 的前端网络性能的包加速处理,和前面介绍的 DPDK 的思想和技术特点:并行水平扩展机制等,一种基于 DPDK 框架的在通用服务器处理芯片上实现通信领域的实时通信系统处理的架构如图 2 所示。

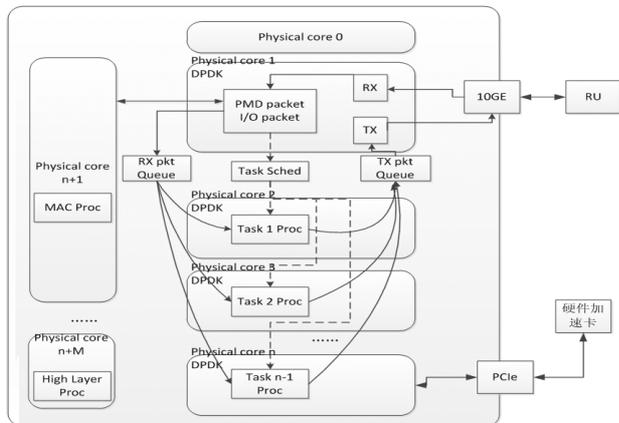


图 2 基于 DPDK 的多核实时通信处理架构

基于 DPDK 的多核实时通信处理架构的一些特性如下:

(1) 专属核专用于网络包的统一处理分发。

当多个核访问同一个内存池时,需要保证每次读写的数

据未被其他核修改,存在存取效率低的问题。由专属核统一接收网络数据包,解析 eCPRI 协议,后分发数据给其他的核做任务处理。可避免多核处理时所带来的 Cache 一致性的问题。

(2) 通信系统链路的功能切分并结合 DPDK 在多核处理器上的并行处理机制。

在收发网络数据包任务后,将通信系统链路的任务可理解扩展为实时包处理的一部分,链路处理按照逻辑功能模块,划分成若干子任务,以基站下行链路为例:按照不同的下行处理信道进行划分,不同的下行信道又可分为接收高层信令的调度、比特级处理、符号级处理、RE 映射等处理等。利用 CPU 亲和性,将各个子任务绑定到特定的核上,使单一任务尽量长时间地在单一核上进行处理,避免多个核共同处理任务而带来的一些切换开销。将任务以水平方向进行扩展,提高了 CPU 缓存的命中率和处理器性能、减少内存访问开销和每个核需要处理的负载。

软件还可以利用 Intel 提供的 AVX 指令集,提高指令的并发度。

基站链路的处理有时序的要求,各个子任务切分后的总时间要满足系统的要求,不满足则要么对任务进行进一步切分,采用多核并行处理提升效率,或者是采用硬件加速卡的功能降低 CPU 负载要求。

(3) 硬件加速处理卡。

DPDK 支持硬件卸载功能,硬件加速处理卡可作为可选项,通过 PCIe 接口和通用服务器芯片相连,根据系统/链路的实时处理性能要求,考虑是否需要作为关键时间敏感的处理任务的协同加速处理。对于基站而言,可以将物理层的编译码、加解密模块等模块在硬件加速处理卡上协助处理实现。

其他 DPDK 的软件优化策略:大页内存使用、锁机制、Cache 优化等的使用在此架构下均有发挥的空间,可进一步优化系统的实时处理能力。

4 结语

综上所述,从 DPDK 的技术原理看,DPDK 不仅可应用于网络包处理、网络虚拟化等应用。其也契合 5G ORAN 前传接口划分的方式,在 5G 基站的实时处理部分亦可有更多的发挥空间;随着 IT 和 CT 领域的技术融合,基于 DPDK 的实时处理系统框架打破传统以 DSP 和 FPGA 处理的一种思维模式,提供另一种的思路和方法,该方法亦可借鉴于不同的实时处理系统中。

参考文献

- [1] 李翠,陈庆奎.基于 DPDK 并行通信的动态监控模型[J].计算机应用,2020,40(2):335-341.
- [2] 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.
- [3] 朱河清,梁存铭,胡雪.深入浅出 DPDK[M].北京:机械工业出版社,2016.

收稿日期:2021-09-04

作者简介:周晓玲(1978—),女,汉族,福建福州人,硕士研究生,工程师,研究方向为宽带移动通信。