

基于 FPGA 的万兆以太网 UDP 协议通信接口设计

赵世超,左金印,魏 晓,赵 哲

(华北计算机系统工程研究所,北京 100083)

摘要: 针对嵌入式设备迅速增长的数据传输需求,介绍了一种依托于现场可编程门阵列(Field Programmable Gate Array, FPGA)平台,以光纤为传输媒介的万兆 UDP/IP 协议通信接口模块,并探究引入分支预测机制对通信接口模块发送时延产生的影响。通过对现有网络接口层、网际层、传输层和应用层典型传输模型的深入研究,使用硬件描述语言,模块化设计 ARP 控制器、IP 控制器和 UDP 控制器,实现完备的 UDP/IP 通信接口模块,并评估了引入分支预测机制对通信接口模块发送时延产生的影响。分析表明:该设计实现简单,适配嵌入式设备对高带宽、低延时、资源低占用的需求,具备自主维护 ARP 表的能力,支持多设备级联。该设计在高速数据采集、远距离信息传输、片上数据高速处理等应用场景具有积极的意义。

关键词: 现场可编程门阵列;光纤通信;万兆以太网;UDP/IP 协议;分支预测机制

中图分类号: TN919

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.222620

中文引用格式: 赵世超,左金印,魏晓,等. 基于 FPGA 的万兆以太网 UDP 协议通信接口设计[J]. 电子技术应用, 2022, 48(10): 113-117, 122.

英文引用格式: Zhao Shichao, Zuo Jinyin, Wei Xiao, et al. Design of 10 gigabit ethernet UDP communication module based on FPGA[J]. Application of Electronic Technique, 2022, 48(10): 113-117, 122.

Design of 10 gigabit ethernet UDP communication module based on FPGA

Zhao Shichao, Zuo Jinyin, Wei Xiao, Zhao Zhe

(National Computer System Engineering Research Institute of China, Beijing 100083, China)

Abstract: In order to solve the rapidly growing data transmission problem for embedded devices, this paper introduces a 10 gigabit UDP/IP protocol communication module, which relies on the Field Programmable Gate Array(FPGA) platform and uses optical fiber as the transmission medium, and explores the impact of introducing the branch prediction mechanism on the transmission delay of the communication interface module. Through in-depth research on the typical transmission model of the existing Open Systems Interconnection(OSI), using hardware description language, the ARP controller, IP controller and UDP controller are modularly designed to form a complete UDP/IP communication module. And the influence of introducing the branch prediction mechanism on the transmission delay of the communication module is evaluated. The analysis shows that the design is simple to implement and adapts to the requirements of embedded devices for high bandwidth, low latency and low resource occupation. It maintains the ARP table independently and supports multi-device cascading. It has positive significance in application scenarios such as high-speed data acquisition, long-distance information transmission, and high-speed processing of on-chip data.

Key words: FPGA; fiber-optic communication; 10 gigabit ethernet; UDP/IP protocol; branch prediction

0 引言

伴随万物互联时代的临近,计算机网络通信技术应用得到空前发展,嵌入式电子设备接入局域网的需求迅速增加。网路传输带宽不断提高也使得一些特定的领域对数据传输提出了更高的要求。远距离通信领域需要更高的数据传输带宽和更强的抗干扰能力,用于保障其通信的质量和可靠性。工业生产的数据采集系统需要更低的数据传输延迟和更大的数据吞吐容量,用于确保其数据的实时性和完整性。片上数据高速处理领域需要通用化的数据传输接口,用以增加其可扩展性。然而,嵌入式

设备上广泛使用的 CPU 和 MCU,碍于设计体积小、设计功耗低等原因,计算能力的提升速度并不像互联网带宽那样明显,并且其差距有进一步扩大的趋势^[1]。过高的以太网数据传输速率会过度消耗 CPU 宝贵的计算资源。在服务器平台上,国外学者提出的 TCP/IP 卸载引擎(TCP Offload Engine)技术是被广泛接受的解决方案,即使用 TOE 芯片硬件网卡实现 TCP/IP 协议接口硬件化的技术来减轻 CPU 运行负担^[2]。在嵌入式平台上, FPGA 从仿真、优化到在线调试都很便捷,在不改变外围电路的情况下可综合出不同的电路功能,具有集成度高、设计

灵活的优点,成为以太网协议硬件化平台的首选^[3]。

近年来,伴随数据信息采集和高速信息传输设备的升级迭代,众多学者在 UDP/IP 协议硬件化方向做出了很多努力。在震动采集监控系统 and 以太网视频传输系统中使用 FPGA 硬件化的 UDP/IP 协议,通过网线传输数据的设计已得到应用^[4-5],但这不能很好地满足在强烈干扰环境下的高速数据传输需求。光导纤维通信技术在远距离、高速数据传输和适应复杂电磁干扰环境方面拥有明显的优势。最近,有学者在新型 FPGA 平台上进行了使用 UDP/IP 协议的万兆光纤通信实验^[6-10],其中一些学者尝试进一步提升万兆 UDP/IP 接口传输性能。

本文旨在使用 FPGA 平台设计出用于实现万兆光纤数据传输的 UDP/IP 协议通信接口,并引入分支预测机制,评估其对系统性能的影响。

1 总体设计

使嵌入式设备 FPGA 平台具备 UDP/IP 通信能力,目前主流的设计方案有:FPGA+CPU 架构或 FPGA+MCU 架构,由 CPU 或 MCU 为 FPGA 解析处理 UDP/IP 协议,转交数据至 FPGA 进行数据处理;FPGA+MAC+PHY 架构,PHY 芯片用于实现物理层的相关功能,MAC 芯片实现数据链路层解析功能,FPGA 实现传输层数据传输功能^[10]。

在 FPGA+光纤通信平台上实现 UDP/IP 硬件化高速数据传输的众多研究中,有些采用第一种架构,此种设计使用 CPU 或 DSP 亦或是 MCU 来实现以太网协议的解析,网络数据传输会频繁地触发 CPU 中断,CPU 反复保护和还原任务现场会频繁访问存储器,这难以达到万兆以太网的传输带宽,且传输延迟不固定;有些采用第二种架构,此种设计依赖 OSI 传输模型中 MAC 层或 PHY 层芯片,如 AX88180、88E1111、W5300、W5500、RTL8211 等^[11-12],外置扩展芯片不利于进一步提升嵌入式设备系统的集成度;还有一些设计实现了 MAC 层协议硬件化^[13],或套用了 UDP 协议格式实现万兆光纤数据传输^[14],这种设计不能自动维护 ARP 表,无法便捷地实现多设备级联。故本文采用单独使用 FPGA 来实现完备的 UDP/IP 协议方案。

1.1 万兆以太网 OSI 参考模型

OSI 参考模型分为 7 层结构,分工实现网络通信工作。每层结构的逻辑单元模块相对独立,工作在不同层结构内的逻辑单元模块可各自独立设计,这种模块化的设计能够提升系统的可维护与可拓展性。

万兆光纤 UDP/IP 协议栈主要工作在传输层、网络层、数据链路层和物理层上。物理层主要功能为实现物理信号的接入,高速数据的串化、解串,以及 64B/66B 编码等。数据链路层主要功能为实现介质无关接口(MII)的连接,链路数据检错等功能。网络层的逻辑模块用于实现 IP、ARP 协议解析功能。传输层的逻辑模块实现 UDP 协议解析功能。本文采用的万兆以太网 OSI 参考模型如图 1 所示。

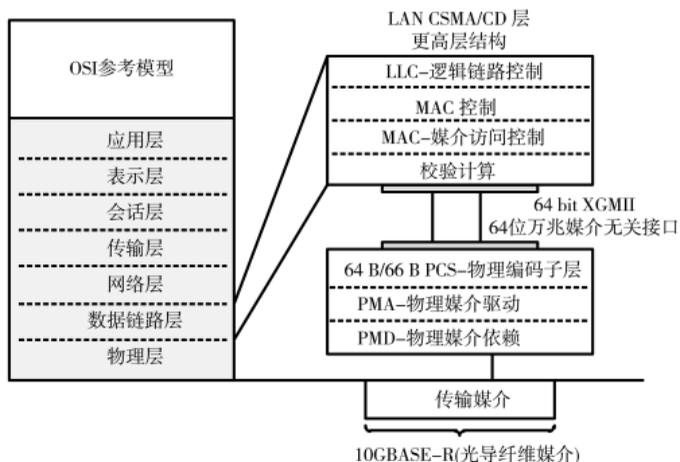


图 1 万兆以太网 OSI 参考模型

1.2 总体架构设计

FPGA 内通信接口系统各个功能模块的设计遵循 OSI 参考模型约束设计,传输层的 UDP 控制器模块、网络层的 IP 控制器模块和 ARP 控制模块、数据链路层的 MAC IP 核以及物理层的 PCS/PMA IP 核分工完成模块所在层的传输功能。各个控制器模块由内部的缓冲区、封装引擎、解析引擎、校验引擎、控制器等子模块组成,各子模块的连接关系如图 2 所示。

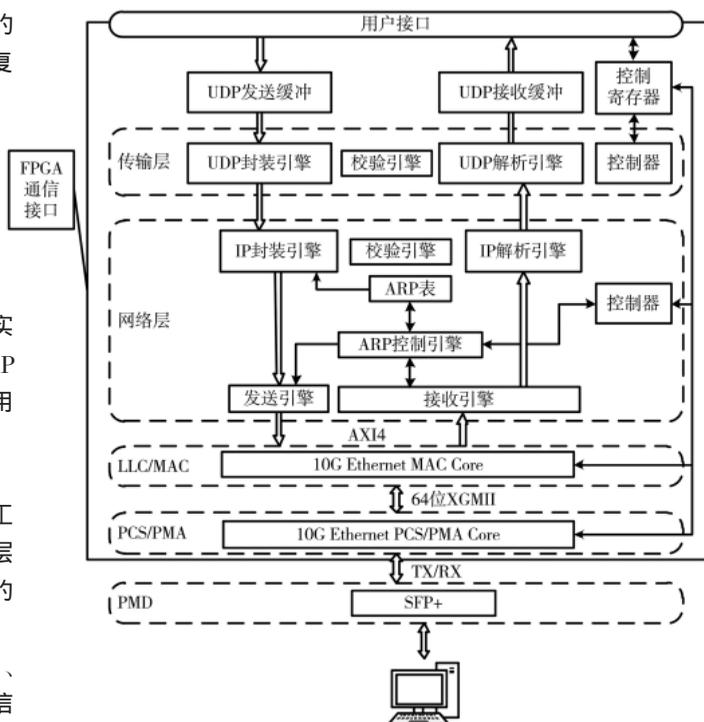


图 2 系统总体架构模型图

UDP 发送缓冲区为 128 KB 的 RAM,接收缓冲区为 256 KB 的 RAM。缓冲区中存放的用户数据格式为 2 字节同步头、4 字节源端口与目的端口、2 字节数据长度和 n 字节数据。控制寄存器用于存放各个逻辑模块的运行

状态字、复位控制标志位、本地 IP 地址、目的 IP 地址以及本地 MAC 地址等信息。

UDP 封装引擎和 IP 封装引擎以及 UDP 解析引擎和 IP 解析引擎之间均使用 FIFO 连接,用于传输 UDP 数据报,包含 8 字节 UDP 协议首部和 n 字节的数据。

网络层逻辑模块与 LLC/MAC 层通过 AXI4-stream 总线相连,用于传输 MAC 协议帧,包含 14 字节的 MAC 协议首部、20 字节的 IPv4 协议首部和 $(8+n)$ 字节的 IP 报文数据。

2 功能模块设计

物理层中负责 LLC 和 MAC 功能的逻辑子单元使用了开源的 10G Ethernet MAC Core IP 核来实现,负责 PCS 核 PMA 功能的逻辑子单元使用了开源的 10G Ethernet PCS/PMA Core IP 核实现,PM D 功能由标准的 SFP+ 光模块实现。物理层部分主要实现光电转化相关功能,这里不进行详细阐述。

2.1 校验引擎设计

校验引擎为其他逻辑模块提供校验和计算服务。该引擎会从校验请求发起的逻辑单元数据缓存区或 FIFO 中读取待校验数据,送入 16 位校验数据寄存器中,不足 16 位的数据末尾填充“0x00”。16 位的校验和寄存器与校验和进位寄存器初值为 0。将校验和寄存器、校验和进位寄存器与校验数据寄存器中的 16 位数据二进制求和,和填入校验和寄存器,进位填入校验和进位寄存器,逐轮累加至校验数据读取完毕。将校验和寄存器中的数据按位取反输出,得到 2 字节的校验和。校验引擎计算模型如图 3 所示。

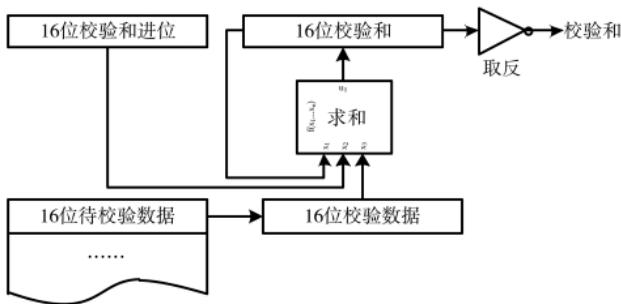


图 3 校验引擎校验和计算模型

UDP 协议处理中参与校验数据为:12 字节的 UDP 协议伪首部、8 字节的 UDP 协议首部和 n 字节的报文数据;IP 协议处理时参与校验的数据只有 20 字节的 IPv4 协议首部。

2.2 UDP 控制器设计

UDP 控制器用于生成和解析 UDP 报文,由 UDP 解析引擎、UDP 封装引擎、校验引擎以及控制器组成。UDP 解析引擎通过 FIFO 从网络层获取 UDP 数据报,校验其数

据正确性,还原出通信双方端口号、数据长度以及通信内容信息,写入 UDP 接收缓冲区供用户使用;UDP 封装引擎读取 UDP 发送缓冲区内的用户消息,为其添加 8 字节的 UDP 协议首部和 12 字节的 UDP 伪首部,计算数据报校验和,生成 UDP 数据报,使用 FIFO 传输至网络层;控制器用于维护模块的运行状态字和复位控制信息等。

UDP 解析引擎的主要工作状态有初始状态、就绪状态、接收状态、解析校验状态和写入状态,其状态转换关系如图 4 所示。

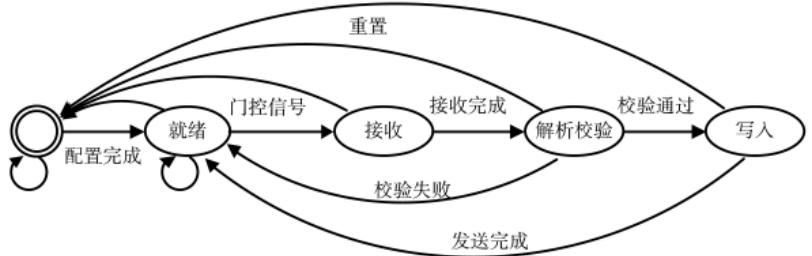


图 4 UDP 解析引擎状态转换图

UDP 封装引擎的主要工作状态有初始状态、就绪状态、读取状态、封装校验状态和发送状态,其状态转换关系如图 5 所示。

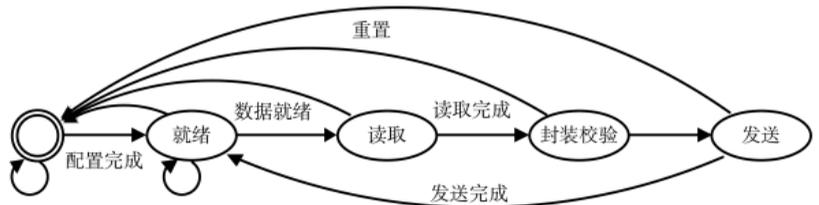


图 5 UDP 封装引擎状态转换图

2.3 IP 控制器设计

IP 控制器用于生成和解析 IP 数据报与 MAC 数据帧,由 IP 解析引擎、IP 封装引擎、校验引擎、接收引擎、发送引擎和控制器组成。接收引擎和发送引擎通过 AXI4-stream 总线与物理层相连,交互以太网 MAC 数据帧。发送引擎为 IP 数据报和 ARP 数据报添加 14 字节的 MAC 协议头生成 MAC 数据帧。接收引擎根据物理层传输的 MAC 数据帧还原出 IP 数据报和 ARP 数据报,过滤掉不相关消息。

接收引擎去掉 MAC 数据帧首部的 12 字节目的 MAC 地址和源 MAC 地址,对数据类型字段进行匹配。0x0800 为 IP 数据报,经由 FIFO 传输至 IP 解析引擎进行处理;0x0806 为 ARP 数据报,0x0835 为 RARP 数据报,通过 FIFO 交由 ARP 控制器进行解析;不能匹配的数据帧直接丢弃。IP 控制器接收引擎处理模型如图 6 所示。

IP 控制器的解析引擎对 IP 数据报首部的 20 字节首部进行正确性校验,还原出 UDP 报文通过 FIFO 发送至传输层使用。IP 封装引擎对传输层交付的 UDP 报文根

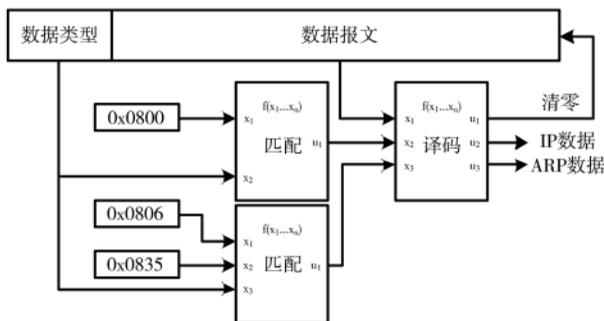


图6 IP控制器接收引擎数据处理模型

据版本号、首部长度、服务类型、总长度、生存周期、协议类型、源IP和目的IP等信息添加20字节的IPv4协议首部,计算出首部校验和,生成IP数据报文。

IP封装引擎与发送引擎组成的逻辑模块采用了分支预测执行处理机制。在大部分通信时间内,通信双方的IP地址与MAC地址在很大的概率上保持不变。每次传输数据时,预测机制根据上两次的MAC地址使用情况预测当前MAC地址是否需要变更,预测为不必变更时,使用缓存在寄存器中的MAC地址,完成MAC帧首部的封装,使得数据帧具备向下一层传输的条件。下层逻辑单元对MAC帧进行FCS计算的同时,ARP控制器对目的MAC地址进行查找确认。这样的处理机制可以尽快地将网络层的数据传输给物理层。具体的工作流程如图7所示。

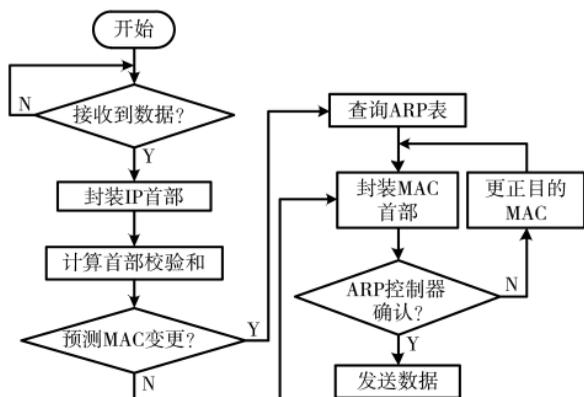


图7 网络层数据报封装工作流程图

2.4 ARP控制器设计

ARP控制器模块为系统提供物理地址解析服务,由ARP控制引擎和ARP表项组成。其中ARP表为占用6.4KB的RAM,存储用于当前通信的IP地址与MAC地址映射关系,除自身的MAC地址固定外,其余的内容由ARP控制引擎动态调整。

ARP控制引擎对外提供配置接口,模块自身IP地址可设定为自动获取模式和指定IP模式。ARP控制引擎根据控制寄存器内容,将通信对端的IP地址录入至ARP表,ARP控制引擎产生相应的ARP或RARP报文,从

网关或其他设备获取对应的MAC或IP信息。收到ARP请求报文自动产生ARP应答。ARP控制引擎工作流程如图8所示。

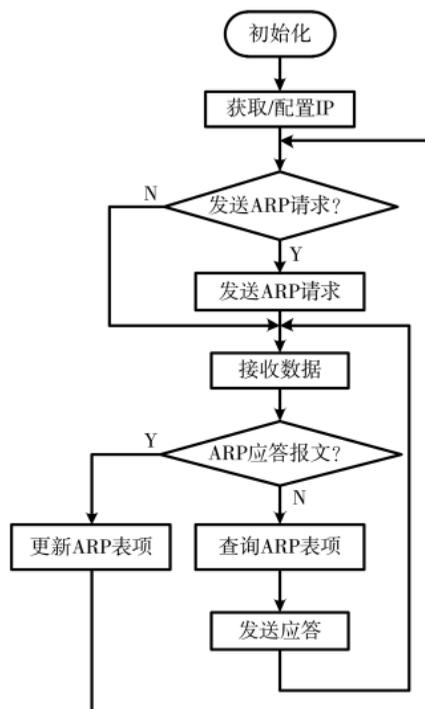


图8 ARP控制引擎工作流程图

3 实验分析

用于本次实验的国产化现场可编程门阵列平台为复旦微电子生产的JFM7K325T FPGA。使用Verilog HDL硬件描述语言生成文中设计,使用EDA工具对其进行编译、综合,统计综合后硬件平台的各种资源占用情况,并给出分析。

通信传输时延一般分为传播时延和处理时延,设计中引入的分支预测机制会对协议接口的发送时延产生影响。建立数学模型评估采用分支预测机制后对系统发送性能的影响。

3.1 资源占用分析

实验统计了PCS/PMA层IP核、LLC/MAC层IP核、网络层和传输层的其他逻辑电路模块对于LUTs资源、FFs资源、BRAMs资源等的占用情况。详细的资源占用数值见表1。本设计在复旦微电子的JFM7K325T FPGA上LUTs使用不超过3%,FFs使用低于2%,BRAMs占用低于3%,可以方便地应用到其他的设计中,提供UDP协议解析服务。

3.2 发送性能分析

对于不同主频的FPGA平台,运行时也不尽相同。这里用引入分支预测机制后的节省用时与原发送用时的比值来反映发送时延的变化。本节阐述的发送用时为从接收到用户数据开始到以太网帧封装完毕的时间跨

表 1 FPGA 平台资源占用统计

层级	LUTs	FFs	BRAMs
PCS/PMA	2 029	2 733	0
LLC/MAC	2 163	1 654	0
其他	约 2 700	约 2 900	约 12.20
总共	326 080	407 600	445
总计/%	2.11	1.79	2.74

度。假设应用层逻辑单元需要传输 n 字节的 UDP 报文，并规定在 FPGA 中拷贝单位字节的用时为 1 个单位。

传输层逻辑模块所使用的时间约为数据拷贝时间、UDP 报文首部建立时间与加入伪首部后校验计算时间之和，即：

$$t_{UDP}=(4+2+n)+1+(\lceil \frac{12+8+n}{2} \rceil+1+1) \quad (1)$$

网络层逻辑模块用时约为数据拷贝时间、IPv4 报文首部建立时间、首部校验计算时间、目的 MAC 查询时间和 MAC 帧首部建立时间之和，即：

$$t_{IP}=[4+(8+n)]+1+(\frac{20}{2}+1+1)+t_{ARP}+1 \quad (2)$$

式中 $t_{ARP}=m$ ，表示目的 MAC 在 ARP 表中的位置。考虑最坏的情况， m 也可用于表示当前所在局域网的规模。数据链路层所承载的数据载荷范围为 46~1 500，即：

$$46 \leq 20+8+n \leq 1\,500 \quad (3)$$

类似的，数据链路层逻辑模块用时约为数据拷贝时间与 FCS 校验计算时间之和，即：

$$t_{MAC}=(14+20+8+n)+(1+1) \quad (4)$$

采用分支预测机制后，实际上是将 ARP 查询功能和数据链路层功能并行化，所以节省掉的发送用时约为：

$$\Delta t = \begin{cases} t_{ARP}+1, & t_{ARP}+1 \leq t_{MAC} \\ t_{MAC}, & t_{ARP}+1 > t_{MAC} \end{cases} \quad (5)$$

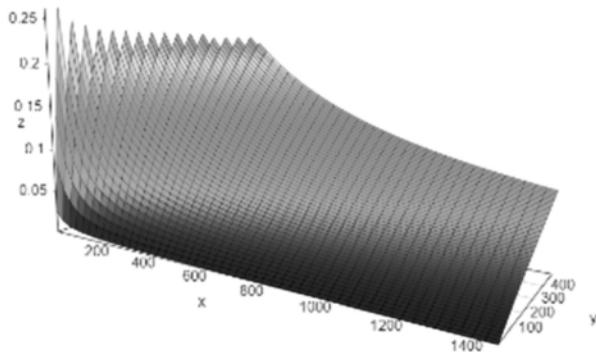
节省掉的发送用时与原发送用时的比率 γ 为：

$$\gamma = \frac{\Delta t}{t_{UDP}+t_{IP}+t_{MAC}} = \begin{cases} \frac{m+1}{89+3.5n+m}, & n-m+43 \geq 0 \text{ 且 } n \text{ 为偶数} \textcircled{1} \\ \frac{m+1}{89.5+3.5n+m}, & n-m+43 \geq 0 \text{ 且 } n \text{ 为奇数} \textcircled{2} \\ \frac{n+44}{89+3.5n+m}, & n-m+43 < 0 \text{ 且 } n \text{ 为偶数} \textcircled{3} \\ \frac{n+44}{89.5+3.5n+m}, & n-m+43 < 0 \text{ 且 } n \text{ 为奇数} \textcircled{4} \end{cases} \quad (6)$$

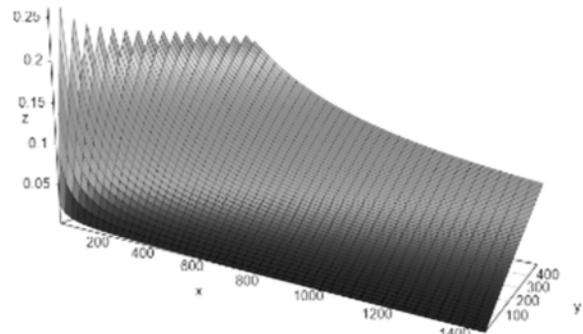
$$\text{约束条件为} \begin{cases} 3 \leq m \leq 500 \\ 18 \leq n \leq 1\,472 \\ n, m \in \mathbf{N}^* \end{cases}$$

式(6)的图像如图 9 所示。

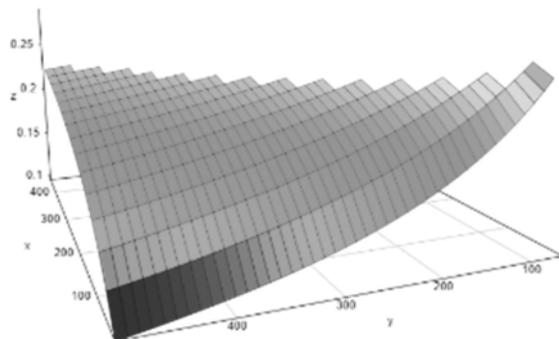
从式(6)的图像中可以看出，接近边界条件 $n-m+43=0$ 的邻近区域 γ 值较大。意为当前系统级联规模与 UDP 报文传输数据的字节数接近时，该设计的发送性能提升较为明显，发送用时可减少至原来的 80% 以下；系统级联规模与 UDP 报文传输数据的字节数相差越多，该设计的发送性能提升越不明显。该设计适合应用在频繁使



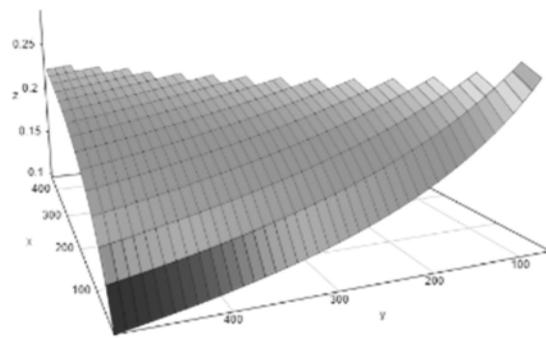
(a) 式(6)①图像



(b) 式(6)②图像



(c) 式(6)③图像



(d) 式(6)④图像

图 9 式(6)的图像

(下转第 122 页)



图 10 系统设计测试环境



图 11 多路视频测试显示图

操作简单方便、成本低、适用性广的优点,具有较高的推广和应用价值。

参考文献

[1] 杨念,李峰.基于 B/S 的嵌入式视频监控系统的设计与实现[J].计算机工程与设计,2008(21):5576-5579.
 [2] 徐大鹏.HD-SDI 视频嵌入式图像采集设计[J].计算机测量与控制,2015,23(9):3214.
 [3] 陈明伟,徐丹.球面坐标定位校正鱼眼图片并合成全景

(上接第 117 页)

用短数据报传输信息且级联规模不大的局域网系统。

4 结论

本文介绍了一种在国产化 FPGA 平台上使用光导纤维传输介质实现万兆通信的 UDP/IP 协议接口。该协议接口充分运用开源的万兆以太网 MAC IP 核与 PCS/PMA IP 核,使得开发过程便捷、简单。应用分支预测机制,进一步减少局域网内短数据报的发送时延。使用本文方法实现的 UDP/IP 协议接口占用资源低,能够灵活应用到其他的设计中。研究内容对高速数据采集、远距离信息传输、片上数据高速处理等应用场景有一定的借鉴意义。此外,对航天、能源、国防等使用国产化器件要求较高的特定领域具有一定的参考价值。

参考文献

[1] 冯一飞,丁楠,叶钧超,等.领域专用低延迟高带宽 TCP/IP 引擎设计与实现[J].计算机工程,2021.
 [2] 王祎辰.基于 FPGA 的万兆以太网 TCP 硬件协议栈设计[D].北京:北京理工大学,2016.
 [3] 熊雪均,谭力波,张俊杰,等.基于 FPGA 的低延迟 TCP 协议栈实现[J].电子测量技术,2020(10):43-48.
 [4] 于波,李建成,张强,等.基于 FPGA 的振动采集监控系统[J].电子测量技术,2021,44(14):33-37.
 [5] 于波,栾海鹏,韩玉斌.基于 FPGA 的以太网视频传输系统[J].电子测量技术,2020,43(8):138-142.
 [6] 柯洋.基于 FPGA 的高速数据传输板设计与开发[D].武汉:华中师范大学,2020.

图的方法[J].云南民族大学学报(自然科学版),2004(3):214-217.

[4] 华为海思半导体有限公司.Hi3559A/C V100 ultra-HD Mobile Camera SoC 用户指南[Z].2018.
 [5] 华为海思半导体有限公司.Hi3531DV200 H.265 编解码 AI 处理器用户指南[Z].2018.
 [6] GUALDI G,PRATI A,CUCCHIARA R.Video Streaming for mobile video surveillance[J].IEEE Transactions on Multimedia,

(下转第 128 页)

[7] 吴惑,刘一清.基于 FPGA 的万兆以太网 TCP/IP 协议处理架构[J].电子设计工程,2020,28(9):81-87.
 [8] Liu Nianyun,Xu Zhiqiang.The design of high-speed hardware UDP/IP stack based on FPGA for large-scale sensing systems[J].因特网技术学刊,2017,18(3):579-587.
 [9] 刘源,张刚.可靠 UDP 协议栈的 FPGA 实现[J].火力与指挥控制,2017,42(7):139-143.
 [10] 孔德伟,袁国顺,刘小强.基于 FPGA 的万兆以太网链路的设计与实现[J].微电子学与计算机,2019,36(12):21-25.
 [11] 李涛,韩鹏,候冠东,等.基于 FPGA 的 ORUDP 协议栈设计与实现[J].计算机工程,2020,46(6):155-163.
 [12] 王文,郑建生.基于 FPGA 的 TCP/IP 网络通信系统的设计与实现[J].现代电子技术,2018,41(8):5-9.
 [13] 王永超,刘超,王健,等.基于 FPGA 的多角度周转箱图像采集系统设计[J].自动化仪表,2019,40(10):35-38.
 [14] 许川佩,刘华颖.通用型芯片级光纤通信接口设计[J].微电子学与计算机,2019,36(11):24-29.

(收稿日期:2022-02-10)

作者简介:

赵世超(1995-),通信作者,男,硕士研究生,主要研究方向:通信保密,E-mail:zhaoshichao_cn@qq.com。

左金印(1967-),男,硕士,高级工程师,主要研究方向:通信保密。

魏骁(1990-),男,硕士,工程师,主要研究方向:通信保密。



扫码下载电子文档

版权声明

经作者授权，本论文版权和信息网络传播权归属于《电子技术应用》杂志，凡未经本刊书面同意任何机构、组织和个人不得擅自复印、汇编、翻译和进行信息网络传播。未经本刊书面同意，禁止一切互联网论文资源平台非法上传、收录本论文。

截至目前，本论文已经授权被中国期刊全文数据库（CNKI）、万方数据知识服务平台、中文科技期刊数据库（维普网）、DOAJ、美国《乌利希期刊指南》、JST 日本科技技术振兴机构数据库等数据库全文收录。

对于违反上述禁止行为并违法使用本论文的机构、组织和个人，本刊将采取一切必要法律行动来维护正当权益。

特此声明！

《电子技术应用》编辑部

中国电子信息产业集团有限公司第六研究所