

基于 FPGA 的千兆网络数据采集系统设计与实现

朱 晴, 吴 宁, 顾薛平

(南京航空航天大学 电子信息工程学院, 江苏 南京 210016)

摘 要: 介绍了网络流量管理中的网络数据采集系统的设计与实现。以 Altera StratixGX 系列 FPGA 为平台, 结合 Marvell 88E1111 网络芯片, 完成了网络数据采集系统。本设计采用 SoPC 技术, 利用 Altera 提供的千兆以太网 IP 核, 完成 FPGA 系统无缝连接千兆以太网, 实现网络数据包采集。

关键词: 千兆以太网; SoPC; MAC; PCS; PMA

中图分类号: TP335

文献标识码: A

文章编号: 1674-7720(2011)21-0053-03

The design and realization of Gigabit network data acquisition system based on FPGA

Zhu Qing, Wu Ning, Gu Xueping

(Nanjing University of Aeronautics and Astronautics, College of Electronic and Information Engineering, Nanjing 210016, China)

Abstract: This paper introduces the design and realization of the system about network data acquisition. The paper takes the Altera StratixGX FPGA with Marvell 88E1111 as the platform to complete the network packet acquisition system. The design presents an approach for implementing the completed seamless connection between FPGA and Gigabit Ethernet using SoPC technology. And the gigabit Ethernet IP provided by Altera.

Key words: triple speed Ethernet; SoPC; MAC; PCS; PMA

随着互联网的迅猛发展, 对网络管理的需求日益迫切, 而流量监测和统计分析是整个网络管理的基础。为了实现网络进行综合管理, 有必要及时全面地收集、管理网络的流量信息, 准确获取网络流量数据, 对网络流量态势进行分析, 对网络健康状况及未来的发展趋势作出准确判断。然而, 随着网络带宽的增加和规模的扩大, 大规模网络流量采集面临着数据规模庞大和数据到达速度过快的挑战。为了克服上述困难, 通常采用硬件实现完成流量采集功能^[1]。当前采用硬件实现的流量采集方法(如 Cisco 的 NetFlow 或者 InMon 的 sFlow)都是基于对数据报文周期性地采样来进行流量采集, 这种方法存在处理速度缓慢、统计结果不精确、实现代价大、单个流处理访存次数多等缺点。

随着 FPGA 的发展, 其内部资源日益丰富, 速度和性能大大提高, 特别是内部嵌入了大量 IP 核^[2], 这些使得在 FPGA 上进行高速网络流数据的采集和分析成为可能。流量管理系统主要包含网络数据采集系统、网络数据统计与分析、网络数据分流三部分, 如图 1 所示。

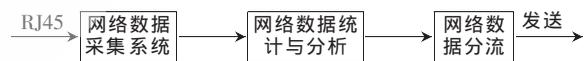


图 1 流量管理系统

本系统基于 FPGA, 获取 RJ45 端口的网络数据, 然后对数据进行流量统计和分析, 最后根据分析结果, 将网络数据包进行分流再重新转发回网络中。

1 网络数据采集系统总体设计

本文基于 FPGA, 重点研究了前端网络数据采集系统, 该系统总体框图如图 2 所示, 网络数据由 RJ45 端口接入系统, 经 Marvell 88E1111 和千兆以太网核的解析, 获取完整的数据包, 暂存在指定地址的片内 RAM 中进行统计和分析。

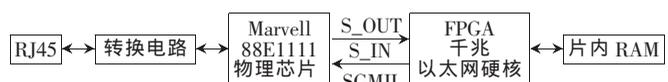


图 2 网络数据采集

网络数据采集系统主要通过 Marvell 88E1111 物理层芯片和 Altera 公司的千兆以太网 IP 硬核来实现, 两者通过标准的 SGMII 接口连接。

网络与通信 Network and Communication

在发送数据的过程中,88E1111 物理层芯片将 FPGA 的千兆以太网 IP 硬核产生的高速串行数据流经 RJ45 以太网口,由 5 类双绞线传递出去。在接收数据的过程中,88E1111 物理层芯片检测从 5 类双绞线传递进来的电信号,如果有有效电信号到来,将高速串行数据流传递给千兆以太网 MAC 核。

2 物理层器件 88E1111

Altera 公司的千兆以太网硬核默认支持的物理层器件有支持 10/100 Mb/s 的 National DP83848C,支持 10/100/1000 Mb/s 的 National DP83865 以及支持双物理层和 10/100/1000 Mb/s 的 Marvell 88E1111。

Marvell 公司的物理层芯片 88E1111 是一款高性能的芯片,支持包括 GMII、RGMII、SGMII、TBI、RTBI 等接口的 MAC 层和物理层的连接方式,可以直接接入千兆以太网 MAC 或交换机端口,10/100/1000 Mb/s 的物理传输层的媒介,光纤传输线连接方式以及双绞线的连接方式^[3]。88E1111 支持的最高速率为 1.25 GHz,I/O 电平标准为 3.3 V LVTTL。系统选择 Marvell 88E1111 作为 PHY 器件。

3 千兆以太网

以太网技术是一种广泛应用的网络技术,千兆以太网在兼容原有以太网技术的基础上进行了一系列改进,网络数据传输速率得到了大幅度提高。特别是 IEEE802.3ab(1 000Based-T)千兆网标准的出台,将 5 类非屏蔽双绞线应用在千兆网中,使得网络性能在原有布线基础上获得大幅度提高^[4]。如今,千兆网已经发展成为主流的网络技术。

千兆网兼容了原来以太网标准所规定的全部技术规范,其中包括 CSMA/CD 协议、以太网帧、全双工、流量控制等。千兆网的标准结构包括介质访问控制(MAC)、物理编解码(PCS)和物理介质接入(PMA)三个主要部分。

3.1 千兆以太网 FPGA 硬核

Altera 提供可参数化的千兆以太网 IP 核解决方案,该方案可在 Altera 的 CycloneII、CycloneIII、Stratix 等系列 FPGA 上工作,可配置使其包含 MAC、PCS 和 PMA 模块中的一种或多种,选择相应的接口标准^[5]。

构建千兆以太网系统必须先了解 MAC、PCS 和 PMA 三个模块的功能及其带来的影响。MAC 模块主要用于全双工模式下的流量控制及 MAC 帧的发送和接收,其主要操作有 MAC 帧的封装与解包以及错误检测。MAC 模块通过 MII/GMII/RGMII 接口与 PHY 器件或者 PCS 模块进行通信,同时为用户提供了基于 Aalon-ST 的 8 bit/32 bit 接口^[6]。

PCS 和 PMA 分别代表物理层的物理编解码子层和物理介质接入层,这两个模块在配置过程中是可选的。PCS 模块专门负责信道的编解码、纠错和扰码等,可以

通过 GMII 或 MII 接口访问,同时提供了 TBI 接口与物理介质接入层进行通信。PMA 模块主要功能是实现数据在物理线路上串\并或并\串转换,同时连接外部物理介质相关(PMD)的设备,可以驱动铜线或光纤网路进行工作,与 PMD 设备接口为 1.25 Gb/s 串行口或者 SGMII 接口。整个千兆以太网系统模块关系图如图 3 所示。

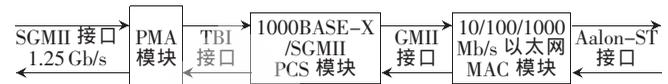


图3 千兆以太网模块

在 Altera 提供的千兆以太网 IP 核中,MAC、PCS 和 PMA 这三个模块都是可选和可配置的。不同的选择和配置,千兆以太网 IP 核与系统和物理芯片的接口是不同的,如表 1 所示,'√'表示选择了这个模块,'-'表示该模块未被选中。

表1 千兆以太网模块配置及接口

10/100/1000 以太网 MAC 模块	1000BASE -X/SGMII PCS 模块	PMA 模块	系统 接口	以太网 物理芯片接口
√	-	-	Avalon-ST GMI/RGMII/MII 和可选的 MDIO	
√	√	-	Avalon-ST TBI 和可选的 MDIO	
√	√	√	Avalon-ST 1.25 Gb/s 串行口	
-	√	√	MII/GMII 1.25 Gb/s 串行口	
-	√	-	MII/GMII TBI	

3.2 物理器件 88E1111 与千兆以太网的接口设计

FPGA 千兆以太网硬核与物理芯片 88E1111 的接口采用标准的 SGMII 接口。SGMII 为串行吉比特媒体独立接口,支持 10/100/1 000 Mb/s 的全双工 BASE-T 功能,支持 8 bit/10 bit 编码,速率为 1.25 Gb/s。此外,SGMII 是串行的,在连接千兆以太网硬核和物理芯片时,不需要提供额外的时钟,容易实现。

系统中 SGMII 接口如图 4 所示,RX 和 TX 为 1.25 Gb/s 的差分输入和输出信号,MDC/MDIO 为 MAC 控制接口对 88E1111 芯片的控制方式,以实现对其的管理。

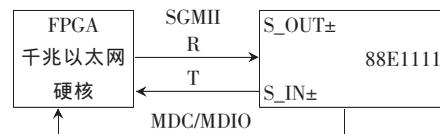


图4 物理芯片和千兆以太网的 SGMII 接口

SGMII 是 1.25 Gb/s 串行数据接口,根据表 1 内容选定系统中千兆以太网 IP 核包含 MAC、PCS 和 PMA 三个模块。

3.3 千兆以太网 IP 核的配置

Altera 提供千兆以太网控制器 IP 核,通过配置界面可将 IP 核配置为所需模式并进行 IP 核参数设置,配置界面分为以下 4 个配置页面^[6]。

(1)Core Configuration:核配置选项,配置以太网功能模块,是否包含 PCS 模块、FIFO 模块,配置接口类型、端

网络与通信 Network and Communication

口数等。系统配置包含 PCS 和 FIFO 模块,同时配置 PCS 接口类型为 LVDS I/O(SGMII 要求)。

(2)MAC Options:MAC 配置选项,配置 MAC 模块功能,系统配置包含 MDIO 模块。

(3)FIFO Options:FIFO 存储器选项,可设置 FIFO 存储器类型以及存储器数据长度。

(4)PCS/SGMII Options:物理介质接入层模块配置页面,配置物理层,系统配置包含 SGMII 桥接口。

4 网络数据采集系统设计

网络数据采集系统是基于 Altera 强大的 SoPC Builder 开发工具来完成系统设计。SoPC Builder 是一个软件工具,它属于一种基于 IP 或平台的设计方法。平台包括 Altera 的 Nios 处理器、Avalon 总线以及片内外存储器。利用 SoPC Builder,用户可以很方便地将各种 IP 模块通过 Avalon 总线连接起来,形成一个完整的系统^[7]。SoPC Builder 已包含一些常用的外设 IP 模块,用户也可自定义外设 IP。

Avalon 总线标准规定了 Avalon-MM(Memory Mapped)和 Avalon-ST(Strming)两种接口。千兆以太网接口逻辑采用 Avalon 总线作为与其他模块的互联总线^[8]。

4.1 系统介绍

为更好地实现数据链路和控制链路的分离,提高系统的执行效率,千兆以太网模块通过 Avalon-ST 接口和用户程序进行网络数据通信,以提升数据流传递过程中的效率。同时,系统自定义网络数据接收模块将千兆以太网解析的数据从 Avalon-ST 接口读出,并存入指定地址处暂存,等待用户程序处理。自定义网络数据发送模块将处理好的数据从 RAM 中读出千兆以太网的 Avalon-ST 接口。

系统自定义千兆以太网控制模块,通过 MAC 核的 Avalon-MM 接口实现对 MAC 核、PCS 和物理芯片的控制。系统详细结构图如图 5 所示。

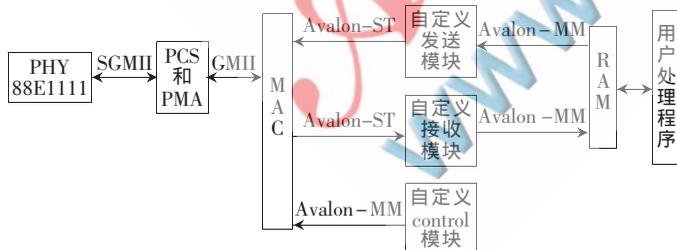


图5 网络数据包采集系统框图

4.2 控制模块设计

千兆以太网采用 Avalon-MM 从端口作为控制接口,接口提供了 8 bit 地址信号线,32 bit 的读写控制信号,同时还提供了一些状态信号。系统通过该控制接口,对 MAC、PCS 和外部物理器件进行寄存器初始化,实现对 MAC、PCS 和外部 PHY 的控制。

自定义 Control 模块是一个 Avalon-MM 主端口模块,

通过 Avalon-MM 总线接入 MAC 控制接口。在 Control 模块内部定义了一个状态机,如图 6 所示,依次实现对千兆以太网各模块的控制。

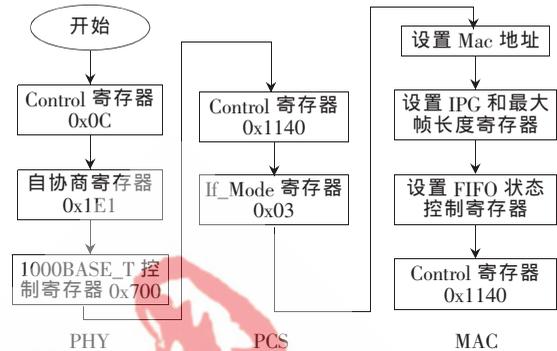


图6 自定义 control 模块状态转换

Control 模块首先初始化外部 PHY (Marvell 88E1111) 的寄存器,启动外部物理器件;然后,初始化 PCS 寄存器,实现 PCS 模块的帧同步;最后,对 MAC 寄存器初始化,设置其工作模式并启动 MAC 开始工作。

初始化 PHY 的寄存器是通过 MDC/MDIO 端口实现的,MDC/MDIO 是一个二线的控制端口,MDC 是控制时钟,MDIO 是串行的输入输出数据。自定义 Control 模块通过 MAC 的控制端口,对 PHY 寄存器初始化。初始化信号通过 MAC 模块的翻译,最终通过 MDC/MDIO 端口送入到 PHY 寄存器中。

初始化 PCS 和 MAC 的寄存器都是直接通过 Avalon-MM 从端口映射实现的。

4.3 自定义收发模块设计

千兆以太网解析模块通过 Avalon-ST 接口和用户程序进行通信。Avalon-ST 接口本质上已经不同于 Avalon-MM 接口,它是一种非总线形式的连接,给设备之间提供了总线之外的点对点专用连接,以提升数据流传递过程中的效率。

Avalon-ST 提供的点对点数据传输不能直接将网络数据流按网络数据包为单位传输到固定位置。将各网络数据以包为单位送到 RAM 指定地址存储,更方便用户对网络数据进行统计处理。

因此,系统自定义网络数据接收模块将千兆以太网解析的数据从 Avalon-ST 接口读出,并存入指定地址处暂存,等待用户程序处理。自定义网络数据发送模块将处理好的数据从 RAM 中读出,并送入千兆以太网的 Avalon-ST 接口。

5 测试及结果分析

为了测试千兆以太网的性能,本文采用 Stratix IV FPGA 与 PC 机通信的方式对千兆以太网的通信速率进行测试。PC 机的硬件平台为 Intel 奔腾 3.0 GHz CPU,2 GB 内存,Broadcom netxtreme gigabit ethernet 网卡,TCP/IP 最大传输单元大小为 1 500。测试方法为 PC 机

向 StratixIV 发送 4 GB 的测试数据,经千兆位解析后,重新发送返回至 PC 机。测试系统结构图如图 7 所示。



图7 测试系统结构图

自定义接收模块将接收的数据重新由发送模块送回 PC 机。数据传输率最高达 1 000 Mb/s。

本文以 Altera StratixGX 系列 FPGA 为平台,结合 Marvell 88E1111 网络芯片,实现了千兆网络数据采集系统。设计过程中采用 SoPC 技术,利用 Altera 提供的千兆以太网 IP 核,完成 FPGA 系统无缝连接千兆以太网,实现网络数据采集。测试表明,该系统能有效、高速地满足网络管理系统的要求。

参考文献

- [1] 王宏,龚正虎.Hits 和 Holds:识别大象流的两种算法[J].软件学报,2010(6):1392-1402.
- [2] 李璇,敖光良.基于 FPGA 的千兆以太网设计[J].网络安全技术与应用,2008(5):63-64.

[3] Marvell. 88E1111 Data Sheet Integrated 10/100/1000 Ultra Gigabit Ethernet Transceiver[Z]. 2004-11

[4] IEEE802.3.Part3:Carrier sense multiple access with collision detection (CSMA/CD) access method and physical layer specification[S], 2005.

[5] 柳利军,熊良芳.基于 FPGA 的千兆以太网交换芯片的设计[J].微电子学与计算机,2006,23(3):80-82.

[6] Altera Corportion. Triple speed ethernet megacore function user guide[Z]. 2008.

[7] 徐欣,于红旗.基于 FPGA 嵌入式系统设计[M].北京:机械工业出版社,2005.

[8] 李兰英.Nios II 嵌入式软核-SOPC 设计原理及应用[M].北京:航天航空出版社,2006.

(收稿日期:2011-06-01)

作者简介:

朱晴,女,1987年生,硕士研究生,主要研究方向:数字系统设计与计算机应用,边界扫描测试与可测性设计,FPGA 系统设计。