

基于嵌入式系统的 PCE 硬件设计及实现

袁广前, 郑小平, 华楠, 金秋实, 刘海蛟

(清华大学 电子工程系, 北京 100084)

摘要:介绍一个基于嵌入式 Linux 的 PCE(路径计算单元)硬件系统设计及实现实例,其主要目的是用于光网络中“PCE+网管”架构的测试,也可用于多域多层异构光网络设备间的互联互通及路径计算。考虑到 PCE 开发需求及嵌入式系统的特点,为硬件系统的使用者预先载入板级支持包,完成硬件的启动引导。文中首先根据该 PCE 架构的需求比较了主流嵌入式 Linux 操作系统的异同及优缺点,然后就硬件系统的设计、实现和初始化做了详细的阐述,最后给出了一个可以成功运行的 PCE 测试实例及部分实验成果。

关键词:光网络;路径计算单元;嵌入式操作系统

中图分类号:TN915.11

文献标志码:A

文章编号:1005-8788(2014)01-0005-04

Design and implementation of embedded Linux-based PCE hardware

Yuan Guangqian, Zheng Xiaoping, Hua Nan, Jin Qishi, Liu Haijiao

(Department of Electronic Engineering, Tsinghua University, Beijing 100084, China)

Abstract: The paper will introduce a design and implementation of the PCE hardware based on embedded Linux system. This system is chiefly used in the tests of PCE+network management architecture in optical networks and in the interconnection/interworking between devices in multi-domain/multi-layer heterogeneous optical networks and path computations. In consideration of the requirements of PCE development and the characteristics of embedded systems, a board support package is preloaded for the system users to boot the hardware. Furthermore, in this paper, the similarities and differences and advantages and disadvantages are compared against the requirements of the PCE architecture and then the design, implementation and initiation are described in detail and an operable PCE test case and parts of the experimental results are given.

Key words: optical network; path computation element; embedded operating system

0 引言

随着通信产业的发展,用户端的通信业务种类和数据流量迅速增加,光传送网络规模不断扩大,网络中不同厂商的设备种类和数量也越来越多。多层多域异构光网络中,实现子网络和设备间的互联互通,并快速、可靠地完成整个光网络的连接调度、保护和恢复,是光网络动态管控机制的研究重点之一^[1]。IETF 工作组提出的基于 PCE(路径计算单元)的多层多域光网络动态控制机制,具有高效灵活的约束路径计算优势^[2,3]。研究人员结合具体网络结构,提出了多种基于 PCE 的光网络动态控制架构^[4-6]。不同的 PCE 架构均要对域内、域间的路径进行计算、连接建立和拆除等操作,需要进行大量的数据读写、分析、存储,一个独立并且简单易用的测试系统和应用程序开发平台是必不可少的。但是,这些架构大多通过仿真或承载于传统网元设备来实现,较少开发专用的 PCE 硬件设备。受限于现有网元设备的架设灵活性、可编程能力、CPU(中央处理器)负荷能力和兼容性等因素,不同 PCE 方案的网

络工作效率、运营成本和稳定可靠性难以保证。与此同时,Linux 逐渐成为嵌入式系统可选的操作系统,它以其内核精炼、开放源码、无需版权费用、占用系统容量小以及系统成熟稳定等优点得到了广泛的使用^[7]。因此,研究并开发能够支持嵌入式 Linux 的硬件系统,可以为 PCE 提供一个灵活的测试平台,也能为多域多层异构光网络中的路径计算和互联互通问题提供可选的解决方案。以下将介绍一种独立于现有网元设备的 PCE 硬件设计及实现方案。

1 适用于 PCE 的嵌入式系统

典型的嵌入式系统,以应用为中心,以计算机技术为基础,软硬件可裁剪,适应于特定的应用程序,对功能、可靠性、成本、体积和功耗都有严格要求。系统应具有稳定的性能、方便的开发环境和标准兼容性等。当前,较典型的系统有 VxWorks、QNX(实时操作系统)以及基于 Linux 的嵌入式系统等。

VxWorks 是专为实时嵌入式系统开发的,能够提供高效的实时任务调度、中断管理、系统资源管理和实时的任务间通信。在嵌入式系统领域中市场占

有率较高,使用较广。它支持 x86、PowerPC、Motorola68xx 和 Sun SParc 等多种处理器。其微内核仅提供进程调度、底层网络通信和中断处理服务,非常小巧(QNX4. x 大约为 12 kB),而且运行速度较快。但使用的环境与 UNIX 不兼容,大多数 Vx-Works 的 API(应用程序编程接口)是专有的^[8]。

嵌入式 Linux 的开发是目前操作系统领域的一个热点。对 Linux 经过小型化裁剪后,固化在容量只有几百 k 字节或几兆字节的存储器芯片或单片机中,应用于特定的嵌入式场合。系统的微内核体系结构简单,网络和文件系统以模块形式置于微内核的上层,驱动程序和其他部件可作为加载模块编译或者添加到内核。系统高度模块化的构件方法,可结合定制的驱动程序和应用程序提供附加功能。

PCE 应用程序测试过程中,需要进行多次修改和再编写,为简化应用程序编写过程,方便对软硬件资源进行的操作,保持整体系统稳定的同时消耗最少的资源,使用嵌入式 Linux 系统是较好的选择。本设计选用嵌入式 PowerPC Linux 内核,通过 freescale 提供的 LTIB(目标映像编译器)工具配置裁剪 Linux 内核并设置使用的根文件系统。

2 硬件系统的设计和实现

硬件设计过程采用模块化的方法进行,按功能需求分别设计电路原理图,导出网表文件后进行 PCB(印刷电路板)布线,生成光绘文件并制作样机。

2.1 PCE 硬件的功能模块划分

硬件系统需要具备既定的运算和处理能力,支持跨域的路径计算、连接的管理和维护、保护和恢复时的相应处理;具备 TE 数据库的功能,能动态地存储和读取资源状态信息;具备与光纤网管理系统以及控制平面设备相互通信的能力,能使用通用的网络通信接口接入到现有光网络中^[9]。对整个 PCE 硬件进行功能模块划分,如图 1 所示。

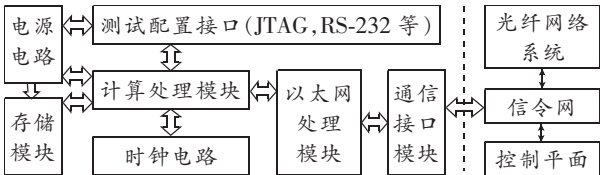


图 1 PCE 硬件系统功能模块划分

计算处理模块是 PCE 硬件系统的核心,通过嵌入式操作系统进行控制,实现 PCE 的软件加载、计算处理功能和整个硬件的管理控制。计算处理模块的运算和处理需要存储模块协助完成。以太网处理

和通信接口模块共同完成 PCE 硬件与其他设备的通信。电源电路为 PCE 内部芯片提供所需电压。时钟电路向计算处理模块提供时钟信号。PCE 硬件的调试等功能集中在测试配置模块。

2.2 具体功能电路的实现

(1) 计算处理模块

该模块完成跨域路径计算、跨域连接的建立、拆除、保护和恢复功能。采用 PowerPC 系列 MPC83XX 作为嵌入式系统的计算处理芯片,其工作电压为 3.3、2.5 和 1.2 V,所需时钟输入为 66.666 MHz,主频为 533 MHz,672 管脚 TBGA (载带型焊球阵列)封装。运算性能满足路径计算对处理器运行速度的要求^[10]。根据处理器内部功能的不同,划分为 DDR(双倍数据速率)控制模块、以太网控制模块以及接口模块等,便于其他模块与核心模块的连接设计。核心模块连接示意图如图 2 所示。

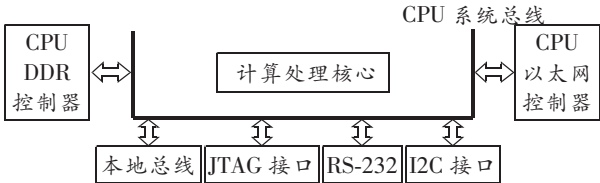


图 2 核心模块连接示意图

(2) 存储模块

该模块用来存储操作系统和程序运行时的各变量,临时存放待处理的通信消息内容。采用 184 管脚的 DDR 内存作为存储器,通过 DDR 总线与计算处理模块中的 DDR 控制器连接。由地址总线选择内存数据,通过 64 位数据总线读写数据。时钟由 CPU 的 DDR 控制器提供,保证同步的数据读写。通过配置控制寄存器设定其频率。连接关系如图 3 所示。

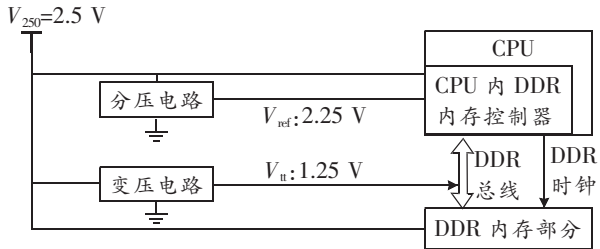


图 3 存储模块连接示意图

(3) 以太网处理模块和通信接口模块

该模块连接计算处理模块和外部设备,与外部设备通信。采用千兆以太网收发器,满足以太网通信的需要。核心芯片是 VSC82XX 以太网卡和 VSC73XX 以太网交换芯片,均使用 25 MHz 时钟信号。对内通过 GMII(千兆比特媒体独立接口)与

计算处理模块相连,对外通过网络变压器连接 RJ-45 接口。

以太网收发器收发物理层信号。发送时,对以太网控制器发来的 MAC(媒体访问控制)消息进行以下处理:编码格式转换,串/并转换,调制(采用 5 电平的 4D-PAM),加扰频,监听到物理层信道空闲时发送。接收时,将物理层信号解扰、解调和解码,所得数据通过 GMII 传送给 MAC 层,处理后交给核心处理单元。

设置以太网变压器,可以使信号的传输距离更远,且使以太网收发器芯片与外部隔离,抗干扰能力增强,同时,接到不同的网口时,不会对彼此设备造成影响。以太网变压器有 4 对输入,4 对输出,分别直接连接以太网收发器和 RJ-45 接口。

(4) 时钟电路

该电路为板上的独立模块提供时钟信号。由于晶体振荡器具有较高的振荡频率,频率稳定,对工作环境变化不敏感,因此本设计采用晶体振荡器作为

时钟源。晶体振荡器向计算处理模块提供其所需的时钟,在 CPU 内部经过倍频处理后用作 CPU 主频。以太网处理模块所需的主时钟也由外部晶振提供,利用此时钟,可产生千兆以太网传输所需的 125 MHz 时钟信号。时钟电路如图 4 所示。

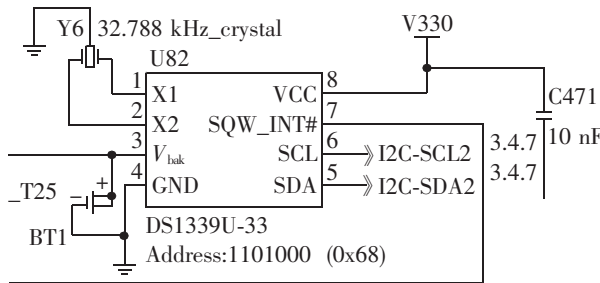


图 4 时钟电路连接图

(5) 电源电路

通过外接电源适配器,将外接的 220 V 交流电源转换为 5 V 的直流电源,为整个硬件系统供电。板上的各芯片所需电源由独立的电源芯片提供。图 5 所示为 3.3 V 电源的接入和滤波连接图。

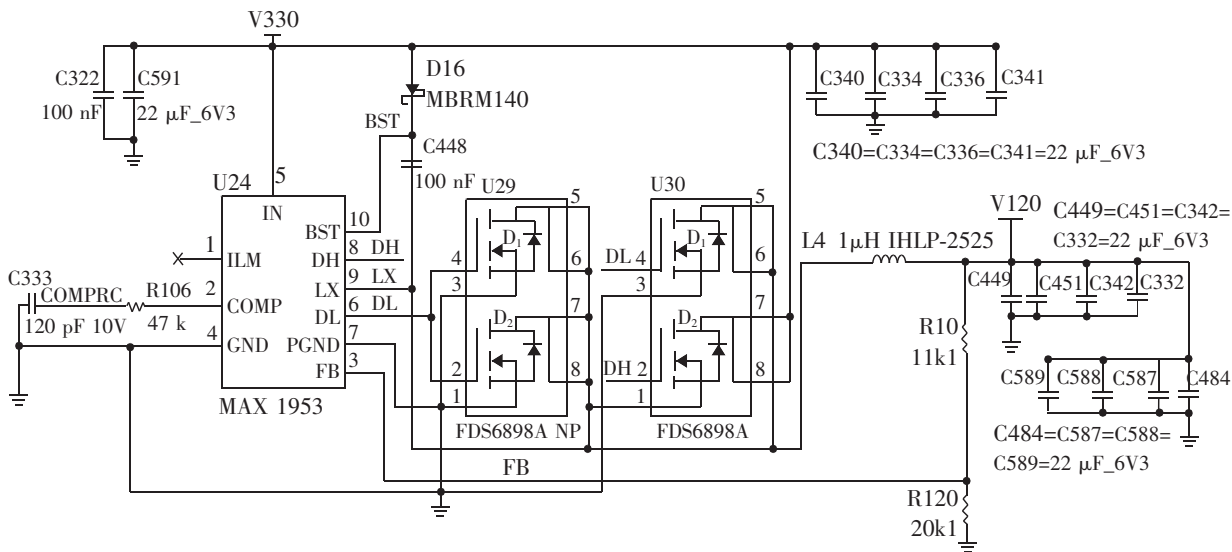


图 5 3.3 V 电源接入和滤波电路连接图

2.3 样机实现

依据 2.2 节的设计,使用 Cadence SPB 16.3 绘制出全部的原理图文件 PCE.DSN,导出网表后制作相应的 PCB 文件。在此过程中,考虑电路工作频率已接近微波频率,综合传输线理论,为避免电磁串扰,通过使用差分线对、圆角布线等措施,保证高频信号的板间传输。

制作出的样机具备 6 个 RJ-45 接口,4 个 USB 口,此外还留有 mini USB 和串口等扩展接口。机箱采用 1U 标准机箱,主板尺寸与标准服务器机箱电源尺寸相当。在通信接口配备和体积控制上有明

显的优势,便于不改变现有光网络结构下接入。

3 启动引导系统及初始化结果

嵌入式软件系统由 bootloader(启动引导程序)、操作系统内核以及根文件系统构成,它们组合成一个供 PCE 等应用程序运行的平台。

本系统使用 Freescale 提供的 Codewarrior 系列软件,通过 USB Multilink interface 连接至 J4 位置,进行 MCU 的初始化。使用的目标文件为 mtxs08qg.s19,主要命令为 SS Specify Object File、EM Erase Module、PM Program Module、PT Program

Trim Value。此时,CPU 及使用其输出电压、时钟信号的外设如风扇等已经可以正常工作,但需要独立芯片控制的外设如网口等还无法正常工作。

使用 USB TAP 连接至 P17 位置,通过 JTAG 接口来烧写 FLASH,将 init.cfg 文件分别写入到两片 FLASH 上。此时,整个硬件具有外设的驱动,网口和串口可用,可以通过 tftp 来进行进一步的配置和烧写。最终在超级终端下的启动界面如图 6 所示。此时,已经可以使用默认用户 root 登陆至 Linux 系统,进行 Apptk 等工具的安装,进而编译并运行已经编写好的 PCE 代码。

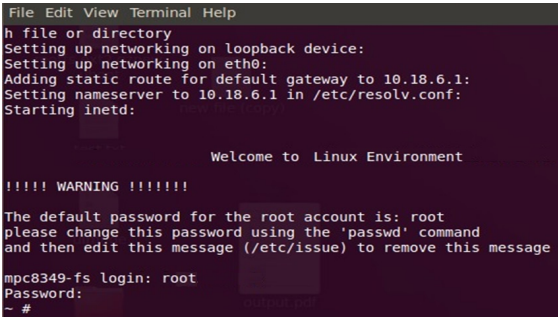


图 6 启动成功界面

4 PCE 应用程序测试实例

测试系统模拟 3 个不同的域:2 个 SDH(同步数字体系)设备域,支持 VC4(155 Mbit/s)粒度交换;1 个 OTN(光传送网)设备域,支持 ODU1(室外单元)(2.5 Gbit/s)、ODU2 (10 Gbit/s)以及波长(10 Gbit/s)3 种粒度交换。

图 7 给出了部分功能的实验成果。实验平台的网管系统承载在服务器上,路由及控制系统承载在

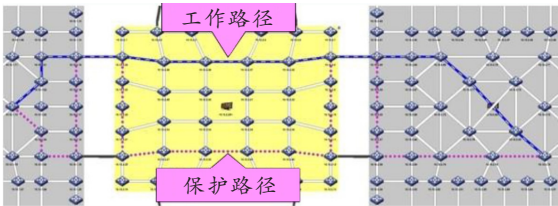


图 7 建立连接成功

基于嵌入式系统的 PCE 样机上,各个域内 GMPLS 控制平面以软件形式承载在节点设备的控制面板内。网络设备实体通过 VLAN(虚拟局域网)连接,通过 PCE 间的协同实现跨域或域内动态业务调度。跨域业务的路径计算采用反向递归的路径算法^[9]。连接建立成功后,综合网管界面显示如图中所示。分析实验统计数据可知,从网管前台发起连接请求到成功建立,全程用时 1 761. 919 ms。其中,源

PCE1 到源 SDH 设备的响应时间为 67. 347 ms,建立连接用时 732. 672 ms;目的 PCE2 到目的 OTN 设备的响应时间为 80. 175 ms,建立连接用时 1 733. 342 ms(98. 378%);PCE1 到 PCE2 的响应时间为 10. 845 ms(0. 616%)。实验数据表明,该硬件能够满足多域多层光网络中计算、建立和管控网络连接实时性要求。

5 结束语

本文给出了 PCE 硬件的分模块设计方案,制作出 PCE 硬件样机并安装板级支持包,为 PCE 应用程序提供了嵌入式 Linux 操作系统的运行环境。在“PCE+网管”结构的多域网络实验平台中,得到了初步的实验结果,验证了基于嵌入式系统的 PCE 硬件设计的正确性。后续的实验中,尚需对不同厂商设备同时接入多域多层光网络的情况进行验证。

参考文献:

[1] Chamania M, Jukan A. A Survey of Inter-Domain Peering and Provisioning Solutions for the Next Generation Optical Networks [J]. IEEE Communications & Tutorials, 2009, 11(1): 33-51.

[2] IETF RFC4655-2006, A Path Computation Element (PCE)-Based Architecture [S].

[3] IETF RFC5440-2009, Path Computation Element Communication Protocol (PCEP) [S].

[4] Tomonori Takeda, Sugiyama R, Oki E, et al. Implementation and experiments of path computation element-based inter-domain network control and management [A]. Optical Communication, ECOC 2008 34th European Conference [C]. Bruxelles Belgium: IEEE, 2008, 1-2.

[5] Cugini F, Paolucci F, Vacarengi L, et al. Implementing a Path Computation Element (PCE) to encompass physical impairments in transparent networks [A]. OFC 2007 [C]. Anaheim, CA, USA: OSA, 2007, 98-101.

[6] LURui, WANG Lei, ZHENG Xiaoping, et al. Implementation of PCE-based Management and Control Plane for Heterogeneous Optical Networks [A]. IEEE/OSA OFC/NFOEC 2011 [C]. Los Angeles, USA: IEEE/OSA OFC/NFOEC, 2011. NTuC.

[7] 王学龙. 嵌入式Linux系统设计与应用[M]. 北京: 清华大学出版社, 2003.

[8] Windriver System Inc. Tornado Training Workshop [R]. USA: Windriver System Inc, 1999.

[9] 金秋实. 支持带宽可变光网络的多域异构光网络管控机制研究和设计 [D]. 北京: 清华大学, 2013.

[10] 金建国. MPC8347嵌入式通信小系统设计参考[J]. 中国科技信息, 2008, (8): 100-101.